INSTITUTO
UNIVERSITÁRIO
DE LISBOA

Multi-period modeling for access to electricity in rural areas of Angola

Pedro Alexandre Castelão Subtil

PhD in management, with specialization in Quantitative Methods applied to Management

Supervisor:
Doctor Maria João Caldas Frazão Lopes, Assistant Professor
Iscte – Instituto Universitário de Lisboa

December, 2021

Department of Quantitative Methods for Management and Economics

Multi-period modeling for access to electricity in rural areas of Angola

Pedro Alexandre Castelão Subtil

PhD in management, with specialization in Quantitative Methods applied to Management

Supervisor:
Doctor Maria João Caldas Frazão Lopes, Assistant Professor
Iscte – Instituto Universitário de Lisboa

December, 2021

Department of Quantitative Methods for Management and Economics

Multi-period modeling for access to electricity in rural areas of Angola

Pedro Alexandre Castelão Subtil

PhD in management, with specialization in Quantitative Methods applied to Management

Jury:
Doctor Maria Teresa Delgado Calapez, Associate Professor, Iscte – Instituto Universitário de Lisboa (President of the Jury);
Doctor Pedro João Coimbra Martins, Coordinator Professor, ISCAC – Coimbra Business School;
Doctor Mónica Alexandra Vilar Ribeiro de Meireles, Assistant Professor, Iscte – Instituto Universitário de Lisboa;
Doctor Ana Maria Duarte Silva Alves Paias, Assistant Professor, Faculdade de Ciências da Universidade de Lisboa;
Doctor Maria João Caldas Frazão Lopes, Assistant Professor
Iscte – Instituto Universitário de Lisboa

December, 2021

Page intentionally left blank.

Acknowledgments

I would like to give a heartfelt thanks to my supervisor Professor Maria João Lopes for her guidance as well as for providing the necessary direction through this process. I would also like to express my recognition to Professor Maria João Carvalho e Cortinhal for her support and feedback in several key moments which made the conclusion of this work possible. It would not have been possible without her.

Many thanks also to Professor José G. Dias for his support and contribution through this process especially at the start of this journey when the direction to take was still unclear.

Finally, I would like to express my gratitude to my wife Sílvia and my children for understanding my absences when working on this thesis and to my parents for always supporting my desire to continue learning.

Page intentionally left blank.

Resumo

Embora a relação entre o acesso à eletricidade e o desenvolvimento seja conhecida, na África Subsaariana a taxa de ligação varia desde o medíocre (nas zonas urbanas) até à quase inexistência (nas zonas rurais). A literatura sobre planeamento energético para esta região depende fortemente de abordagens de países desenvolvidos. Angola é um país da costa ocidental da África Austral onde o investimento no sistema elétrico se concentrou na geração e intensificação da rede, enquanto que a expansão da rede tem estado, na sua maioria, estagnada. Em 2018 as zonas rurais do país tinham uma taxa de ligação de 2% devido à falta de know-how local, restrições orçamentais, dados pouco fiáveis e uma falta geral de uma visão holística da questão, por parte dos decisores. Propõe-se uma metodologia multi-período em duas fases para avaliar se a introdução de mini-redes híbridas no cabaz de abastecimento de Angola pode acelerar o acesso das comunidades rurais à eletricidade. A metodologia emprega uma abordagem heurística que integra sistemas de informação geográfica e teoria de grafos para produzir vários cenários com a evolução anual do sistema energético do país, para o período de 2022 a 2030. Estes cenários incluem soluções com e sem mini-redes híbridas e uma análise dos trade-offs envolvidos. A metodologia pode efetivamente apoiar a tomada de decisões e os resultados mostram que as mini-redes híbridas têm um papel a desempenhar na eletrificação de Angola, ligando mais 3.075.526 milhões de pessoas quando comparado com um cenário exclusivo de expansão da rede. Isto representa uma taxa de ligação de 86%, ultrapassando largamente os objetivos subscritos para 2030, e pode ter uma influência decisiva nestas comunidades rurais empobrecidas.

Palavras chave: Eletrificação rural, Planeamento, GIS, Teoria dos Grafos, Angola

Códigos JEL: C33, P25.

Page intentionally left blank.

Abstract

Although the relation between access to electricity and development is known, in Sub-Saharan Africa the rate of connection ranges from the mediocre (in urban areas) to the practically non-existence (in rural areas). The literature on energy planning for this region relies heavily on approaches from developed countries. Angola is a country on the west coast of southern Africa where the investment on the electric system has focused on generation and grid intensification while grid expansion has been mostly stagnant. In 2018 rural areas of the country had a rate of connection of 2% due to lack of local expertise, budgetary constraints, unreliable data and a general lack of a holist view of the issue, by the decision-makers. A two-stage multiperiod methodology is proposed to assess if the introduction of hybrid mini-grids in Angola's supply mix can speed up the access to electricity for rural communities. The methodology employs a heuristic approach that integrates Geographic information systems and graph theory to produce various scenarios with the yearly evolution of the country's energy system, for the period from 2022 to 2030. These scenarios include solutions with and without hybrid mini-grids and an analysis of the trade-offs involved. The methodology can effectively support decision-making and results show that hybrid mini-grids have a role to play in Angola's electrification connecting an additional 3,075,526 million people when compared to a mostly grid exclusive scenario. This represents an 86% connection rate, largely surpassing subscribed goals for 2030, and can have a decisive influence in these impoverished rural communities.

Keywords: Rural electrification, Energy Planning, Mini-grid, GIS, Graph Theory, Angola

JEL Codes: C33 - Spatio-temporal Models; P25 - Urban, Rural, and Regional Economics.

Page intentionally left blank.

Table of contents

Page intentionally left blank.

List of abbreviations and acronyms

AHP	-	Analytic Hierarchy Process
CPLEX	-	IBM ILOG CPLEX Optimization Studio
ECCAS	-	Economic Community of Central African States
EDL	-	Economic Distance Limit
ENDE	-	Empresa Nacional de Distribuição de Electricidade
GDP	-	Gross Domestic Product
GIS	-	Geographic Information System
GADM	-	Global Administrative Areas
HDI	-	Human Development Index
HV	-	High Voltage
IEA	-	International Energy Agency
GPw4	-	Gridded Population of the World version Four
IRENA	-	International Renewable Energy Agency
IRES	-	Integrated Renewable Energy Systems
IRSEA	-	Instituto Regulador dos Serviços de Electricidade e de Água
LCOE	-	Levelized Cost of Electricity
LP	-	Linear Programming
LV	-	Low voltage
MADM	-	Multi-attribute Decision Making
MCDA	-	Multi-criteria Decision Analysis
MCDM	-	Multi-criteria Decision Making
MILP	-	Mixed Integer Linear Programming
MODM	-	Multi-objective Decision Making
MPI	-	Multidimensional poverty index
MST	-	Minimum Spanning Tree
MV	-	Medium Voltage
MO	-	Multi-objective
NPC	-	Net Present Cost

NTL	-	Night Time Lights
POI	-	Point Of Interest
PRODEL	-	Empresa Pública de Produção de Electricidade
PV	-	Photovoltaic
REM	-	Regionalized Electricity Model
RNT	-	Rede Nacional de Transporte
RMSE	-	Root Mean Square Error
SE4all	-	Sustainable Energy for All
SHS	-	Solar Home Systems
SO	-	Single Objective
SSA	-	Sub-Saharan Africa
UNDP	-	United Nations Development Program
WACC	-	Weighted Average Cost of Capital

List of units

A - Ampere
GWh - Gigawatt hour
kWh - Kilowatt hour
M - Meters
USD - United States Dollar
V - Volt
W - Watt
W/h - Watt/hour
W/m^2 - Watt per Square Meter

Page intentionally left blank.

Glossary

Ampere	-	A unit of measurement of current flow in a circuit. One *ampere* (A) is the amount of current flow provided when one *volt* of pressure is applied to one *ohm* of resistance.
Circuit	-	The path that electricity follows.
Current	-	The rate of flow of electrical energy through a conductor wire. Measured in *amperes*.
Demand	-	The electrical energy use by all the various loads on the power system.
Distribution	-	Transmission of lower voltage energy over distribution power lines to its destination, where it is again, transformed for residential, commercial and industrial demand.
Electricity	-	The flow of electrons from atom to atom in a conductor.
Energy	-	The amount of time a load is on *(hours)* x the amount of power used (*Watt*). Measured in *Watt-hours* (Wh)
Fossil fuel	-	A fuel (such as coal, oil or natural gas) formed in the earth from plant or animal remains, over hundreds of millions of years.
Generation	-	Production of electrical energy through the transformation of other energy sources (heat, mechanical, hydraulic, chemical, solar, wind or others).
Hybrid systems	-	Consists of two or more energy sources used together for improved system efficiency and balance in energy supply.
Loads	-	Devices connected to the power system are referred to as electrical *loads*.
Mini-Grid	-	An off-grid electricity distribution network involving small-scale electricity generation
Off-grid systems	-	A system designed to help people function without the support of remote infrastructure (stand-alone power systems or mini-grids).
Ohm	-	Electrical unit of resistance.
Photo Voltaic	-	The photovoltaic type of solar power plant converts the sun's energy directly into electrical energy.
Pixel	-	The smallest addressable element in a raster image, or the smallest controllable element of a picture represented on the screen.
Power	-	The product of voltage x current. Measured in *watt* (W).

Power Plant	-	Power generation plants produce the electrical energy that is ultimately delivered to consumers through transmission lines, substations, and distribution lines.
QGIS	-	Quantum Geographic Information System (software)
Raster	-	A dot matrix data structure that represents a generally rectangular grid of pixels (points of color), viewable via a computer display, paper, or other display medium.
Renewable energy	-	A continuously recharged energy source.
Substation	-	A collection of equipment for the purpose of raising, lowering and regulating the voltage of electricity.
Transmission		High voltage (HV) power lines efficiently transport electrical energy over long distances to the demand locations.
Vector	-	Computer graphics images that are defined in terms of points on a Cartesian plane, which are connected by lines and curves to form polygons and other shapes.
Volt	-	A unit of electrical pressure that causes flow in a circuit. One *volt* is the amount of pressure required to cause the flow of one *ampere* against the resistance of one *ohm*.
Voltage	-	The potential energy source in an electric circuit. Measured in *volt.*
Watt	-	Unit of measurement indicating the electrical power applied in a circuit.
Watt-hour	-	Unit of electrical energy. It indicates the amount of work done, in an hour, by a circuit at a steady rate of one *watt*.

Adapted from (Blume, 2007)

List of tables

List figures

Introduction

Context

Energy has always played a crucial role in the evolution of human society (Smil, 2010) and electricity, especially, is a key contributor to human development since there is evidence of co-movement and a causality relationship between electricity consumption and human development[1] (Ouedraogo, 2017). However, the electric coverage rate for rural areas in Sub-Saharan Africa (SSA) was around 10% by the end of the second decade of the 21st century.

Historically, electricity tends to arrive first to urban areas where population density is very high and the cost of access to electricity per capita is lower meaning remote rural areas are often the last ones to be considered for grid extension, more so in SSA. In this region electrification is a challenge because, aside from budgetary constraints, decision-makers need to be aware of factors that help to define the problem of access to electricity, particularly in rural areas. These factors include cost, long-term sustainability, regulations, demographic and geographic factors, specific local constraints and social preferences. Additionally, there has been an excessive focus on grid extension whereas it has become more or less consensual that grid extension may not be solution for the supply mix of these areas. From an economic, environmental and social stance, off-grid solutions should be considered.

All of these challenges are also present in Angola, a country on the west coast of southern Africa with a total land area of 1,247,000 km^2. According to estimates for 2018 (INE - Instituto Nacional de Estatística (2018)), Angola had a population of 29,259,000 and an electric coverage rate of 30%, most of it in urban areas, below the electrification rate in SSA (43%), assessed by the International Energy Agency (Garcia & Hawes, 2021). In rural areas, the situation is even worse with a coverage rate of 2%.

To address this situation, Angola has subscribed ambitious goals (ECCAS - Economic Community of Central African States, 2014) for the year 2030 and beyond, namely achieving and electrification rate of at least 55% by 2025.

Nevertheless, progress has been slow and in the last 5 years investment has been more focused on generation and grid intensification while grid expansion has been almost stagnant. The national transmission public company (RNT) is the sector entity responsible for planning, supervising and managing grid expansion. They travel all through the country to collect data and assess the conditions for potential grid expansion often making studies and producing reports for decision makers on where to expand next. They try to focus on demographics, demand and budgetary

---

[1] *As measured by the United Nations Development Program's Human Development Index (HDI).*

constraints. Due to incorrect or disperse information, decisions are often made without taking into account all the variables involved and, consequently, missteps are made that lead to a slow expansion of the network. At current rate of expansion, remote rural areas are at risk of remaining isolated for an indefinite period, since little progress has been made since 2014. In fact, in some ways, the situation is regressing.

While the issue of access to electricity is vital for the development of any society, the literature regarding electricity planning and management in SSA presents significant gaps (Trotter et al., 2017). Studies are often conducted on a macro level, with moderate granularity, considering only a few aspects of the problem and a single time-step. This is a complex problem but often the literature oversimplifies the issue with low levels of integration of relevant criteria (economic, social, environmental, political, and technological). There are few studies, especially in the non-Anglo Saxon and non-Francophone countries like Angola. In Angola, the most complex studies are mainly qualitative studies.

This research will contribute to fill this gap by proposing an approach that is both scientifically sound and can have a significant impact on the current decision-making process for electric system expansion in Angola. A methodology where all relevant factors coalesce to form an effective decision support system for energy planning can positively influence the life of millions of people as well as the economic development of the country as a whole, particularly in impoverished rural communities.

Research problem

Taking into consideration the gaps that were previously identified and the current situation in Angola this study addresses the challenge of access to electricity in rural areas of Sub-Saharan Africa (SSA), in particular, the planning and decision-making process for access to electricity in Angola. It is a complex issue due to a series of factors:

- Data restrictions and dispersion, spread through different public and international institutions;
- Lack of local expertise on the use of tools for planning and decision-making purposes;
- Economic and budgetary constraints;
- Outdated internal practices of the national entities responsible for developing the grid.

In line with the Sustainable Energy for All initiative (United Nations Development Programme, 2015b) Angola has adopted specific targets regarding energy access, energy efficiency and renewable energies for the period until 2030 (ECCAS - Economic Community of Central African

States, 2014). Bearing in mind the timeframe of those targets this study covers the period between 2022 and 2030 for assessing the potential evolution of the coverage rate using a supply mix with grid expansion and hybrid mini-grids.

Research hypothesis

According to the literature devoted to developing countries like Ghana, Kenya, Ethiopia and Nigeria (Bertheau et al., 2017; Blechinger et al., 2019; Deichmann et al., 2011; Kemausuor et al., 2014; Moksnes et al., 2017; Parshall et al., 2009), mini-grids play an important role to accelerate electrification in unelectrified areas. However, this is not what is being done in Angola since the focus has been on the grid expansion. This apparent contradiction contributed to the formulation of the research hypothesis: "For some rural communities of Angola, hybrid mini-grids[2] are the fastest and most economically feasible solutions for electrification, for the period between 2022 and 2030".

Research objectives and contributions

The main objective of this study is to assess if hybrid mini-grids provide a solution for speeding up the electrification of Angolan rural communities, in the timeframe of this study.

To accomplish this objective the following middle-steps need to be performed:

- Collection of Angolan relevant data, such as population, national energy demand/supply mix and public and private spending in the electric sector;
- Characterization of the current state of the energy sector and national level strategy;
- Assessment of the state of art on electrification planning and decision making namely in developing countries;
- Review of free tools and data that may be relevant for the study;
- Definition of a conceptual framework for a model to test the study's hypothesis;
- Selection of assumptions and operationalization of the model;
- Production of maps depicting potential outcomes (scenarios) to aid decision making;
- Extraction of useful conclusions from the results;

This study integrates different areas of knowledge, allowing for a more holistic approach to this complex planning and management problem. This approach should maintain scientific validity and simultaneously have real-world application and impact. The integration of free Geographic

---

[2] *A mini-grid is an off-grid electricity distribution network involving small-scale electricity generation. Hybrid systems, in particular, consist of two or more energy sources used together to improve system efficiency and balance in energy supply.*

Information Systems (GIS) tools adds value from a scientific and practical point since it supports modeling different scenarios, and consequently facilitates the decision-making process.

The literature devoted to rural electrification in SSA is scarce (Trotter et al., 2017). This research aims to mitigate this gap by proposing a decision support model with scientific relevance and managerial application.

Moreover, the study's results will also be introduced to the non-academic community (local government entities and public in general) via seminars, public presentations and eventually book form, raising awareness for the choices that need to be made to accelerate the growth of the electricity coverage rate in Angola.

Thesis structure

This thesis is composed of five main chapters. In Chapter 1 a literature review was carried out on the state of the art on energy planning in Sub-Saharan Africa. This review focused not only on the approaches and complexity of different studies but also on the elements that compose the methodology of these studies. Chapter 0 presents the case study of this research with a demographic, economic and energy sector characterization of Angola. The data selection and methodology description, for the study, were carried in Chapter **Error! Reference source not found.**. Results are given in Chapter **Error! Reference source not found.**. In this chapter a geolocation and demographic characterization of unconnected rural settlements in Angola was produced followed by an estimation of demand, according to a series of assumptions connected to the level of poverty of the different regions. Then the supply mix was selected and the cost structure of each technology defined as well as the metrics that best measure the economic feasibility and sustainability of each solution. With these elements in place it was possible model grid growth and installation of different hybrid solutions. A sensitivity analysis of key cost structure components and scenario analysis was employed to compare different solutions. Finally, outputs for all scenarios were produced along with maps, with time steps, of the evolution of the system. A discussion of all results is carried out in Chapter **Error! Reference source not found.** while the conclusions, implications, limitations and avenues for future research are given in Chapter 0.

Literature on energy planning in Sub-Saharan Africa

Energy plays a crucial role in the evolution of human society. Nevertheless, access to basic energy services is extremely uneven across different regions of the globe and authors, like Smil (2010), have long warned that the long-term survival of humanity cannot be assured without setting limits on the planetary scale. This is perfectly illustrated by the insufficient levels of demand in SSA and the excesses of the richest western urban societies, where even the average per capita energy use is beyond what is reasonable, for healthy and comfortable living. The relation between energy use per capita and GDP was studied by Bhatia & Angelou (2015) and the results are given in Figure 1.

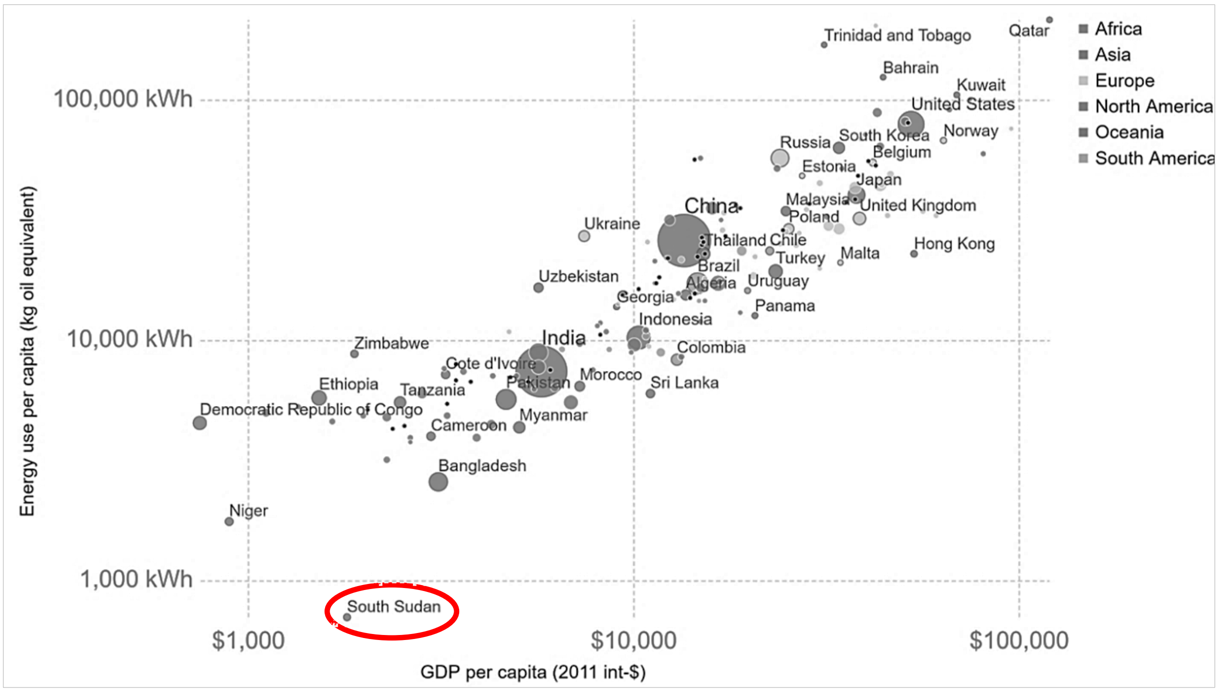


Figure 1 - Energy use per capita vs. (GDP) per capita in 2015 (Bhatia & Angelou, 2015)

Figure 1 shows not only how GDP and energy consumption are positively correlated, but also the contrasting situation between developing countries and western societies: western countries tend to appear nearer the top right corner of the graphic signifying higher energy consumption per capita associated with higher GDP. If Angola was represented in the figure, it would be placed near South Sudan, on the lower left corner, given that in 2018 the country had a per capita electricity consumption of approximately 322 kWh and a per capita Gross Domestic Product (GDP) of $3,432[3]. While the gap between developing countries and western society is beyond the scope of this research it would be useful to reflect on how other countries with similar per capita GPD have substantially superior levels of consumption and higher electric coverage rates, such as the neighboring countries of Congo and the Democratic Republic of Congo.

Long-term studies also show that beyond a direct effect on economic growth, an indirect effect also occurs, as energy can also influence other productivity factors (Halsnæs & Garg, 2011) that may also impact the development of nations.

Bhatia and Angelou (2015) studied the relation between energy and development and found that energy demand rises in correlation with the Human Development Index (HDI) but only up to a point (Figure 2).

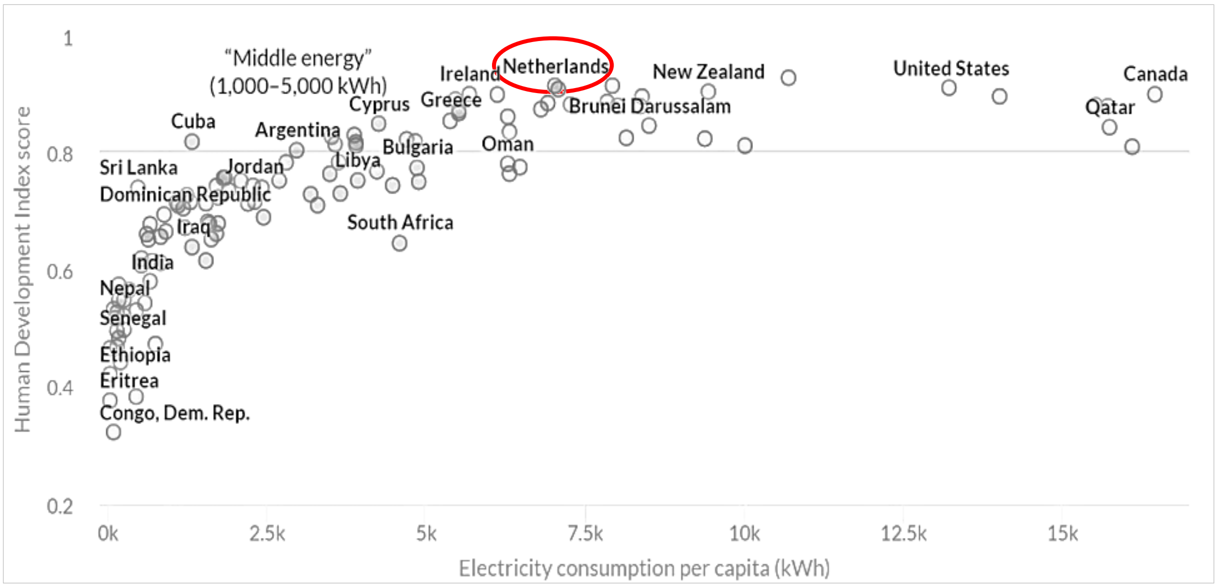


Figure 2 - Relation between electricity demand and the HDI (IRENA, 2018)

Figure 2 highlights that the relation between energy consumption and HDI is not linear meaning that some ways in which energy is consumed are pointless, wasteful and ineffective: countries on right have higher levels of electricity consumption and higher HDI scores but only up to a point. Levels of consumption like those of the United States or Qatar offer no increase in HDI when compared to the Netherlands, for example.

Additionally, countries on the bottom left of this graphic have low levels of electricity

---

[3] *https://data.worldbank.org*

consumption and low HDI score highlighting that it is imperative to increase the levels of access to energy in developing countries. However, literature has mostly focused on industrialized countries (Trotter et al., 2019), leaving ample room for contributions specifically directed at energy planning in Sub-Saharan Africa (SSA).

For developing countries, most studies focus on clean energy generation. However, Bhatia and Angelou (2015) suggest that, while moving away from fossil fuels is a necessity, those countries with high levels of demand should lead the change. For example, in SSA diesel could still be crucial for access to electricity. An intermediate solution needs to be found that permits the coexistence of these older technologies with more environmentally friendly ones. This is largely ignored in literature. Should millions of people, with no access to electricity, wait for a renewable, sustainable solution, while less environmentally friendly options are more readily available?

The literature regarding electricity planning and management in SSA is relatively scarce (Trotter et al., 2017), especially for non-Anglo Saxon and non-Francophone countries. For example, for Angola, officially a Portuguese speaking country, at the time of writing only one study existed on the topic of energy planning. In Table 1 the number of papers concerning electricity planning and management per country in SSA is listed.

Table 1– Papers pertaining electricity planning and management in SSA

Country	***Studies***	***Country***	***Studies***
Angola	*1*	Lesotho	4
Benin	1	Mali	4
Congo, Rep.	1	Mozambique	4
Djibouti	1	The Gambia	4
Seychelles	1	EAPP	5
Swaziland	1	SAPP	5
CAPP	1	Zambia	6
Burundi	2	Botswana	9
Cote d'Ivoire	2	Senegal	9
DRC	2	Mauritius	10
Malawi	2	Cameroon	12
Mauritania	2	Zimbabwe	13
Rwanda	2	Uganda	17
Sierra Leone	2	Ethiopia	> 20
WAPP	2	Ghana	> 20
Liberia	3	Kenya	> 20
Namibia	3	Nigeria	> 20
Sudan	3	South Africa	> 20
Burkina Faso	4	Tanzania	> 20
Cape Verde	4	-	-
Eritrea	4	-	-

Adapted from (Trotter et al., 2017)

In Angola the lack of affinity with internationally accepted languages by the scientific community might play a role in scientific output. Another study by Morrissey (2019) also highlights Nigeria and Kenya as dominant references in the literature on SSA, both English speaking countries.

Regarding SSA, Trotter et al. (2017) put forward that most studies present methodologies with limitations, such as relatively low levels of integration regarding the complexity of the approach, focus on single-objective methodologies and short-term analysis with one time step. This indicates gaps in literature that may be filled by more holistic approaches. Moreover, studies concerning alternative processes to achieve access to energy are mostly focused on urban areas, only a few consider the specificities of the electrification of rural areas.

Authors often take into account the goals of the United Nations Sustainable Energy for All initiative (United Nations Development Programme, 2015b) and Economic Community of Central African States (ECCAS - Economic Community of Central African States, 2014) in which electricity should reach the majority of the population of SSA by 2030. Although several countries in the region have subscribed these goals, less than ten years away from 2030, this goal is far from completion in Africa, and progress has slowed down.

An analysis regarding energy planning in developing countries (mostly in Africa and Latin America) was carried out in order to organize and compare different studies according to certain prevalent dimensions:

- The choice of supply mix;
- The country or region where the case study takes place;
- The complexity of the problem and resolution of the data for the case study;
- The general methodology, including number of research objectives;
- The type of tools used and their level of integration;
- The general approach to modeling and designing energy systems as well as the algorithms that were used.

These studies represented some of the most relevant entries in the literature and are referenced throughout this study.

Table 2 concerns the literature on choice of supply mix and countries where case studies have been carried out.

Table 2 - Overview of some relevant studies in the field of energy planning in developing countries (Part 1/3)

	Choice of supply mix										**Country of case study**									
Authors	Grid	Stand-alone PV	Stand-alone diesel	Stand-alone wind	Diesel mini-Grid	PV mini-grid	Wind mini-grid	Hydro mini-grid	Hybrid mini-Grid	Biodiesel mini-Grid	Kenya	Nigeria	Ghana	Ethiopia	Burkina Faso	Liberia	Malawi	Venezuela	Uganda	Colombia
Parshall et al. (2009)	●	●			●						●									
Deichmann et al. (2011)	●	●	●	●	●	●	●		●	●	●		●	●						
Modi et al. (2013)	●	●			●											●				
Kemausuor et al. (2014)	●	●			●								●							
Ohiare (2015)	●	●			●							●								
Mamaghani et al. (2016)		●	●	●					●											●
Moksnes et al. (2017)	●	●			●	●	●	●			●									
Mentis et al. (2016)	●	●			●	●	●	●						●						
Bertheau et al. (2016)	●	●										●								
Moner-Girona et al. (2019)	●	●			●	●			●		●									
*Blechinger et al. (2019)**	●	●							●			●								
López-González et al. (2019)		●		●		●			●									●		
Trotter et al. (2019)	●					●		●											●	
Korkovelos et al. (2019)	●	●	●		●	●	●	●									●			
*Ciller & Lumbreras (2020)**	●	●							●											

Source: Author compilation from several authors

Table 2 shows that the majority of studies include grid and standalone photovoltaic solutions, modeled with differing levels of detail. A large number includes diesel generation, still a cornerstone of energy solutions for developing countries. Hydro mini-grids are present in only a few studies, at a sub-national level, due to lack of reliable data according to Szabó et al. (2013) and Morrissey (2019). It is interesting to see that English-speaking countries like Kenya, Ghana and Nigeria are prevalent in literature featuring in several studies over the years. This is in line with the assertion posed by Trotter et al. (2017) and Morrissey (2019) that a few developing countries concentrate the majority of studies.

Table 3 is dedicated to the complexity, methodologies and tools used in some relevant studies in the field of energy planning in developing countries.

Table 3 - Overview of some relevant studies in the field of energy planning in developing countries

(Part 2/3)

	Complexity				Methodology				Tools			
Authors	Continent	Country	Province or municipality	Settlement level	Multi objective	Single objective	1 Time step	Multiple time-steps	GIS	Programing language	Software	Custom
Parshall et al. (2009)		●				●	●				Network Planner[1]	
Deichmann et al. (2011)		●				●			●			●
Modi et al. (2013)			●			●			●			●
Kemausuor et al. (2014)		●					●				Network Planner[1]	
Ohiare (2015)		●	●					●			Network Planner[1]	
Mamaghani et al. (2016)				●				●			HOMER[2]	
Moksnes et al. (2017)			●			●		●			OnSSET[3] and OSeMOSYS[4]	
Mentis et al. (2016)	●					●		●			OnSSET[3]	
Bertheau et al. (2016)	●					●	●		●	●		●
Moner-Girona et al. (2019)		●				●			●			●
*Blechinger et al. (2019)**			●			●	●		●		HOMER[2], Network Planner[1], OnSSET[3] and GeoSim[6]	
López-González et al. (2019)		●					●					●
Trotter et al. (2019)		●	●		●					●	IBM CPLEX	
Korkovelos et al. (2019)		●				●		●	●		OnSSET[3]	
*Ciller & Lumbreras (2020)**					●							

*Methodology only. No case study was carried out.

[1] https://qsel.columbia.edu/network-planner/

[2] https://www.homerenergy.com

[3] http://www.onsset.org/

[4] www.osemosys.org

[5] https://www.energyplan.eu

[6] https://www.geosim.de/en/

Analyzing Table 3, it becomes clear that literature is focused on single objective approaches with one time-step that use mostly pre-made modeling tools such as OnSSET, OSeMOSYS, Geo Sim and Network planner which may have contributed to the similarity of the approaches. Some authors rely on custom modeling although, in a few of these cases, the full methodologies are not clearly disclosed. Multi-objective approaches on a nationwide level are absent from the literature at the time of writing although they are used for smaller regions of countries (Moner-Girona et al., 2019; Trotter et al., 2019).

Algorithms and decision metrics prevalent in studies in the field of energy planning in developing countries are shown in Table 4.

Table 4 - Overview of some relevant studies in the field of energy planning in developing countries (Part 3/3)

	Algorithms			Decision metrics					
Authors	Planning approach	Grid extension algorithm	Objectives	Levelized cost of electricity (LCOE)	Net present cost (NPC)	Urban vs rural electrifcation inequality (custom	Regional electrification inequality (custom)	Social impact	CO$_2$ Emissions
Parshall et al. (2009)	Heuristic	Kruskal's Algorithm	Least-cost optimization	●	●				
Deichmann et al. (2011)	Heuristic	Prim's algorithm	Least-cost optimization	●					
Modi et al. (2013)	Heuristic	Prim's algorithm	Least-cost optimization		●				
Kemausuor et al. (2014)	Heuristic	Kruskal's Algorithm	Least-cost optimization	●	●				
Ohiare (2015)	Heuristic	Kruskal's Algorithm	Least-cost optimization	●	●				
Mamaghani et al. (2016)	Heuristic		Least-cost optimization	●	●				
Moksnes et al. (2017)	Linear Programming (LP)		Least-cost optimization	●	●				
Mentis et al. (2016)	Linear Programming (LP)		Least-cost optimization	●	●				
Bertheau et al. (2016)	Linear Programming (LP)		Least-cost optimization	●					
Moner-Girona et al. (2019)		Kruskal's Algorithm	Least-cost optimization	●					
*Blechinger et al. (2019)**	Linear Programming (LP)	Kruskal's Algorithm	Least-cost optimization	●	●				
López-González et al. (2019)	Analytic Hierarchy Process (AHP)		Economic, enviromental and social						
Trotter et al. (2019)	Mixed-integer linear programming (MILP)		Cost minimization, urban versus rural electrification inequality minimization, regional electrification inequality minmization			●	●		
Korkovelos et al. (2019)	Linear Programming (LP)		Least-cost optimization	●	●				
*Ciller & Lumbreras (2020)**	Mixed-integer linear programming (MILP)		Cost minimization, social impact maximization, CO2 emission minimization		●			●	●

Source: Author compilation from several authors

Table 4 shows a strong emphasis on the use of GIS tools and premade planning tools in smaller scale projects whereas custom tools are the most used at a country level. These custom tools can be built from the ground up or be the result of combining already available tools while decision metrics are mostly from the economic sphere.

The modeling approach for nationwide studies is mainly heuristic due to the complexity of the problem although some authors use linear programming and analytic hierarchy process. Finally, when grid extension growth is simulated, a minimum spanning tree (MST) related approach is used.

The complexity of the challenge of planning for electricity access, in rural areas of SSA, justifies the use of methodologies that help maximize the possibility that investments are made as efficiently as possible. To further understand the nuances of the relation between the energy planning problem in SSA and the different methodologies that are available, a deeper analysis was carried out. Ciller and Lumbreras (2020) suggest that, in a general sense, the level of complexity of a study is function of the scope of analysis. These authors conducted a literature review in this field and proposed a classification of the research, according to the level of complexity, as summarized in Table 5.

Table 5 – Planning literature in SSA according to complexity and computational load

-	***1. Pre-feasibility studies*** Initial feasibility assessments that are usually based on the Levelized cost of Electricity (LCOE) of several energy supply alternatives. Resolution of the solutions ranges from villages to nationwide. Network designs not calculated. Most use GIS technologies.	+
modeling complexity	***2. Intermediate analysis*** Network layouts are calculated using geometric considerations, rules of thumb or analytic expressions. They allow for planning on a national level but do not replace engineering studies.	**Computational speed**
+	***3. Detailed generation and network designs*** Accurate generation and network designs (mini-grids, grid extension and stand-alone systems) incorporating detailed engineering models and equations. Due to high computational load and technical detail, these designs are more appropriate for subnational problems.	-

Adapted from (Ciller & Lumbreras, 2020) by author.

A deeper description of each type of study is now given.

Prefeasibility studies

These types of studies offer preliminary information, before more complex models are employed (Ciller & Lumbreras, 2020; Corigliano et al., 2020). Some exploit GIS technologies but only as input data, which can include the location of power plants, the layout of the power grid, solar irradiance, wind speed, hydro potential, the layout of the roads and population density, among others but do not produce a grid extension layout or network designs for off-grid solutions. This is a key distinction from other types of studies. In prefeasibility studies algorithms are often employed to firstly group consumers in demand clusters. Then, for each demand cluster, the technology is usually chosen using least-cost approaches in which off-grid solutions are sized using rules of thumb rather than analytic expressions. These studies consider only a solution which is given instantaneously (instantaneous models). Prefeasibility studies have as an advantage the low computational requirements, useful for nationwide problems, and as a drawback are made with significant simplifications and assumptions in order to tackle large-scale problems. Some examples of these types of studies are Modi et al. (2013), Mentis et al. (2017), Moksnes et al. (2017) and Korkovelos et al. (2018).

Intermediate analysis

Intermediate analysis models also exploit GIS technologies as inputs to access the necessary data but apply more complicated calculations to network designs. The layout of the grid extension and off-grid solutions is usually built with heuristic methods and the use of minimum-spanning-tree related algorithms. This is a significant improvement over prefeasibility studies and increases the complexity of the modeling approach. Often, these algorithms are limited to distance-based constraints resulting in layouts that while not technically accurate, are still useful for decision-making. Grid extension and off-grid solutions, and their components (high, medium and low voltage, transformers, etc.), are sized exogenously using, for example, peak demand as reference. Most approaches in this category focus on least-cost optimization and disregard environmental and social factors. This category includes methods that go beyond the use of rules of thumb for their assumptions, as well as those approaches that combine different methods in the pre-feasibility group. Cuisinier et al. (2020) reaffirmed that, usually, models with characteristics similar to Intermediate analysis studies use Linear Programming (LP), Mixed-Integer Linear Programming (MILP) or Meta-heuristic approaches. The outputs produced using these approaches are mostly instantaneous growth solutions when dealing with rural electrification problems in a large scale. Examples are Parshall et al. (2009), Deichmann et al. (2011), Kemausuor et al. (2014), Ohiare (2015), Bertheau et al. (2016), Blechinger et al. (2019) and Ciller & Lumbreras (2020). Additionally, these models usually only allow for the connection of one location with each new grid extension. Approaches with similar traits to Intermediate analysis models have received other designations in literature, such as Comprehensive Models (Corigliano et al., 2020).

Detailed network design

Finally, there are detailed network design studies (Ciller & Lumbreras, 2020) that use models with a very high level of detail, usually not available, for developing countries. They are used for detailed consumer-by-consumer designs typical of engineering applications. These models are essentially what Cuisinier et al. (2020) classified as Operational MILP approaches, that are quite accurate, employ distribution network constraints and generation constraints and allow the connection of several locations with one grid extension. This last aspect is an important distinction relative to other types of studies. Additionally, network extension and off-grid solution layouts and sizing are very detailed and fit for real world implementation. This approach is outside of the scope of a study at a national level, due to computational limitations. An example of this approach is the Regionalized Electricity Model (REM) developed by the Massachusetts Institute of Technology, as commercial tool (Ellman, 2015). In spite of its complexity, this approach often uses a single objective (cost minimization). Other objectives tend to appear as exogenous variables, such as considering a penalty for CO_2 emissions. This level of detail comes at the expense of intensive computation placing these models outside of the scope of this research since they are not directed at nationwide planning studies.

Since Pre-feasibility studies, Intermediate Analysis studies and Detailed Network Design studies often share common premade tools (like OSeMOSYS, OnSSET or Network Planner), the models that drive energy planning in different studies present similarities. This facilitates comparison but may be somewhat restricting in terms of flexibility and innovation. These tools have been reviewed by Blechinger et al. (2019) and Morrissey (2019) in some detail and, in effect, they employ black box models with steep learning curves carrying the risk of being used but not fully understood. For example, these tools focus mostly on using a least-cost approach and share a set of common assumptions, like the use of a buffer zone around grid, where all locations are assumed to be grid connected. This is often not the case since small population clusters near the grid are rarely grid connected due to economic reasons. Nevertheless, the analysis of studies that employ these tools is invaluable and allows the determination of the common elements and methodological stages for dealing with the problem of access to electricity in rural communities of developing countries:

i) Most studies begin with a data collection stage. Regardless of the resolution and complexity involved in attaining these elements, the ancillary data usually pertains to: GIS data with the spatial distribution of population, the demand for each of these vertices, the definition of an adequate supply mix using data on local resources, the definition of the cost structure for each technology in the supply mix as well as the selection suitable decision-making metrics. These metrics often require carrying out economic analyses over the expected life of equipment (often decades);

ii) The methodology stage involves proposing an approach, that usually involves algorithms, to select the best electrification option for each demand vertex. This approach can integrate typical elements of decision making with LP, MILP, Meta-heuristics and Graph theory methods combined with uncertainty mitigation measures, such as sensibility analysis and scenarios.

These elements constitute the basis of energy planning for developing countries and are present across the majority of studies in this research’s literature review. A detailed review of these topics and their integration in the overall methodology, was carried out in Subchapters 1.1 and 1.2.

1.1 Data collection – literature review

The majority of studies classifiable as Pre-Feasibility and Intermediate Studies require data on the different dimensions of the rural electrification problem in developing countries: population distribution, energy demand, supply, the cost of energy solutions and associated decision-making metrics. This process of data collection is sometimes complex due to the absence of quality information in developing countries which forces the researcher to use indirect methods to extract the required information. The way in which this is carried out is described from subchapter 1.1.1 to 1.1.5.

1.1.1 Spatial distribution of the population

In planning, one of the guiding principles and questions for research is “Where do the people without electricity live?”. The existence of people, as consumers of a product or service, is one of the driving forces for the development of several sectors: energy, water, health and education among others. This means population mapping is an important input for planning for development.

Frequently, population mapping uses GIS tools and there are two main types of formats for population distribution maps:

- Maps in raster format, in which each pixel can have several attributes, namely population size;
- Maps in polygon format (point or vector) where each point or vector shape can have several attributes, namely population size or density.

While both formats are of great use for planning and decision-making, they require quality ancillary data. Governments and national statistics institutes periodically produce censuses with demographic characterizations of the population. However, populations change rapidly making population datasets rapidly outdated. Additionally, census data often lacks the resolution required for the purposes of planning in rural areas where the smallest administrative division is still quite large and population density is often relatively low. This is currently a challenge in many SSA countries, but there is a real need for this type of data and the potential applications are diverse. For example, according to Zhao et al. (2019), high-resolution gridded population maps are essential for effective urban planning, disaster prevention and rescue, environmental and ecological protection and public health monitoring. Several authors support the assessment that the use of accurate geospatial data on global population distribution for analysis and planning is increasing (Gaughan et al., 2015; Freire et al., 2020). Over the years, research has shown the need for more detailed spatial information on human populations in SSA, for the development of the energy sector (Deichmann et al., 2011; Florczyk et al., 2020; Hay et al., 2005; Leyk et al., 2019; Linard et al., 2011; Stevens et al., 2020; Tatem et al., 2007; Zhao et al., 2019). As described by Hay et al. (2005), the aforementioned authors distribute census polygonal data into continuous surfaces using the following methods:

- *Areal weighting* involves overlaying a raster surface (a grid) on administrative unit vector data and assigning population, to each raster cell, according to the proportion of the polygon area. This procedure was used to generate the gridded population of the world (Deichmann et al., 2011). This method assumes that population is uniformly distributed;
- *Pycnophylactic interpolation* has the same basis as the previous method. Raster values are iteratively smoothed using the weighted average of nearest neighbors and then, the total is adjusted to maintain the original vectors count. This method assumes there are no significant boundaries in space, when it comes to population distribution;
- *Dasymetric mapping* allows the incorporation of different information such as high-resolution polygon population data to address the issue of land use (e.g., urban, rural, forest, protected areas) in order to distribute population differently according to the type of land use;
- *Smart interpolation* has similarities to dasymetric mapping but uses ancillary data, such as the location of roads, bodies of water and/or settlements as well as the existence of protected, non-habitable areas, to name a few. This approach allows for the distribution of the population in a non-uniform and non-random way. It has been used, for example, in the United Nations Environment Program - Global Resource Information (Tobler et al., 1997).

A comparison of several freely available population distribution products is given in Table 6. At the time of writing there were no yearly maps for the period of this study (from 2022 to 2030).

Table 6 - Commonly used population distribution products

Data	**Years**	**Format**	**Price**	**Resolution**	**Approach**	**Source**
Gridded population of the world, version 4 (GPWV4)[4]	2000; 2005; 2010; 2015; 2020	Raster	Free	1 Km^2 per pixel	dasymetric	CIESIN - Columbia University
Oak Ridge national laboratory (ORNL) Landscan[5]	2000 to 2016	Raster	Free	1 Km^2 per pixel	smart interpolation	Oak Ridge national laboratory
Worldpop gridded population tiles[6]	2000 to 2020	Raster	Free	100 m per pixel	dasymetric	-
World Population Estimates (WPE)[7]	2013; 2015; 2016	Raster	Free	150 m per pixel	dasymetric	ESRI

Adapted from Leyk et al. (2019)

Statistics produced concerning the spatial distribution of population are usually based on administrative divisions. However, administrative divisions are not the same in all countries. For example, a country like Angola has three divisions or levels (Province, municipality and commune). Other countries, like Kenya, have as many as five divisions of the territory, increasing the detail that is available (also called the resolution of the data).

For the Gridded Population of the World map - version four (GPW4), different methodologies were initially used. The dasymetric mapping approach was the one that produced lower error. The Afripop.org and Worldpop.org initiatives, both using the dasymetric approach, produced some of the highest resolution gridded population datasets currently available. For a detailed explanation see Linard et al. (2011) and Leyk et al. (2019). These authors suggest that, for all commonly used population distribution products, regardless of the approach, urban areas are usually better characterized than rural areas. This is due to the use of ancillary data pertaining to the existence of

---

[4] https://sedac.ciesin.columbia.edu/data/collection/gpw-v4
[5] https://landscan.ornl.gov/
[6] https://www.worldpop.org/methods/populations
[7] https://www.esri.com/en-us/arcgis/products/data/data-portfolio/demographics

human settlements. For place of residence, for example, the use of Nightlights[8] (NTL) data can be useful. Grippa et al. (2019) propose a methodology to improve urban population models with very high-resolution data. According to the authors, one of the main challenges of dasymetric mapping, in particular, is that it uses the available ancillary data to find spatial reallocation "weights" that allow for the redistribution of census administrative data into finer resolution subunits. These weights can be extracted with expert knowledge or machine learning. Built settlement (from small huts to skyscrapers) layers are usually extracted from satellite images and are considered the most important predictor for human population distribution. However, several types of ancillary data can also be used as given in Table 7.

Table 7 – Ancillary data comparison

Gridded population datasets	**Ancillary data layers**									
	Pop. data	Roads	Land Cover	Built Structures	Cities or urban areas	Night time lights	Infra structure	Environmental data	Protected areas	Water bodies
GPWV4	•									•
Landscan	•	•	•	•	•			•	•	•
Worldpop	•	•	•		•					•
WPE	•	•	•	•	•	•	•	•	•	•

Produced by the author

According to Table 7 WPE is the product that uses the largest number of ancillary datasets in opposition to GPWV4 that only takes into account population data and water bodies which are the two data layers that are present across all datasets.

In a broader sense, there are model related issues and validation challenges that are relevant for any population distribution mapping approach (Leyk et al., 2019). A list of these issues is presented and described next:

- *Issues of Scale*: differences in the resolution of input data are relevant across regions as well as in the same region, at different moments in time. Several studies on this topic have been conducted. One study (Bakillah et al., 2014) was conducted using point of interest (POI)

---

[8] https://earthobservatory.nasa.gov/features/NightLights

data to refine population estimates in the city of Hamburg. An extensive review of methods for control of the quality of volunteered data was conducted by Senaratne et al. (2017) focusing on positional accuracy, topological consistency, thematic and semantic accuracy, completeness and temporal accuracy, lineage, usage and purpose. The study highlights that lower population density positively correlates with fewer points of interest, thus affecting data completeness or positional accuracy. This was confirmed by other authors (Girres & Touya, 2010; Haklay, 2010; Mullen et al., 2015; Neis et al., 2013). These findings are consistent with the difficulty in finding data for rural SSA. Recently Herfort et al. (2019) suggested that combining crowdsourcing with deep learning can improve the quality of human settlement maps. The study also suggests that the creation of these maps should rely on automated procedures, where possible, and rely on human skills when needed.

- *Issues of Currency and temporal agreement:* Ancillary data may be time variant. For Africa, census information is often over 10 years old and the relations between census data and administrative boundaries may not be clear. It is important to take into account these aspects of temporal agreement among several layers of input data, especially when projecting forward/backward from census data. For census data, the completeness of coverage and margin of error are crucial.
- *Issues of Semantics*: Data provided by censuses is a convention, whose distribution almost never occurs at one moment in time (de jure[9] or legal distribution). Most methodologies use a place of residence or Nightlights (NTL) (de jure) concept.
- *Model related* issues: Most approaches use data conversion, and data conversion propagates uncertainty. The relation between ancillary data and population for the purposes of population reallocation has some error. This is compounded by the error in the census and ancillary data itself. It is ideal to test the relation between population and ancillary variables. For example, the relation between NTL and population or built structures and population has been well established in the literature. Sutton et al. (2001) analyzed the relation between city population and city areal extent, for 1.383 cities. The study used nighttime satellite imagery comprised of bright pixels (illuminated areas) and dark pixels (areas without light) along with the location of each city as well as its population size and area. When several cities fell into one 'lit' area, their populations were summed. By using the known percent of population in urban areas for every nation, a total national population was also estimated. Results were just 6.8% over the best estimates and census results available at the time. This is true in urban areas since they have more illuminated

---

[9] *A De Jure count of the total population is the enumeration of persons who usually reside in a given place. A De Facto count is the enumeration of persons physically present at specified place on census night.*

areas. More recently, a systematic review of studies on urban mapping using NTL was carried out by Li & Zhou (2017).

- *Validation* issues: Validation requires high-resolution data that is rarely available in SSA countries. Some authors propose that it may be possible to perform an internal validation of the model (Gaughan et al., 2015; Reed et al., 2018; Sorichetta et al., 2015). In general, the procedure implies comparing the gridded maps with finer resolution census data. Grippa et al. (2019) use an internal validation procedure that can give a measure of the suitability of the population distribution approach. One metric used to test the performance of the smart interpolation procedure was the root mean square error (RMSE) which can be interpreted as average error (distance) between a predicted value and the observed value. The procedure has three main steps:
    - Stage one - A set of coarser administrative units is created by aggregating some of the smallest ones.
    - Stage two – A smart interpolation technique is applied to these new coarser areas.
    - Stage three –The aggregated (predicted) values for the sample areas are compared to existing (reference) data from the census. The metric used to test the performance of the smart interpolation procedure is the normalized metric root mean square error (%RMSE).

With this study it was possible to attain a global %RMSE of 44% across several individual countries.

The %RMSE tends to improve with higher-resolution data that is usually only available for smaller areas. Patela et al. (2015) worked with a subnational area and were able to produce lower %RMSE (12%) as did Thomson et al. (2021) who attained similar results but these authors worked with very high-resolution data from household location in the slums of Nigeria. Also, according to the literature (Grippa et al., 2019; Leyk et al., 2019; Stevens et al., 2020), the key issue in the differences in results is more related to the quality and availability of the ancillary data (nightlights, land cover and land use data and its resolution) than the process itself. This leads to national level studies generally having higher levels of error than smaller areas. There is no very high-quality ancillary data, at a country level, for SSA. It is only available for some urban areas. Leyk et al. (2019) concluded that it is up to the analyst to decide if the data product is adequate for the purposes of any given study. There is currently no systematic procedure to decide this.

Population is one of the most important factors for design of the energy systems, so it should be estimated, due to its close relation to demand. Subchapter 1.1.2 reviews literature on demand estimation.

1.1.2 Demand estimation

In long-term planning, energy demand can be internalized in the modeling or treated as an exogenous variable. Energy demand models have extensive literature and Suganthi and Samuel (2012) carried out a review of models used to forecast demand. The authors presented a list of methods for energy demand estimation, but no application for rural areas is mentioned. Other authors have provided valuable references regarding the economic aspects of renewable energy expansion in SSA, like Parshall et al. (2009) and Deichmann et al. (2011). Both studies classified population as urban or rural, for demand estimation. Yearly demand was estimated for households, hospitals, schools and other public institutions. Settlements were classified according to four demand profiles and household and productive demands were estimated separately. The authors suggested that demand estimation with statistical models is to be avoided, in most SSA countries, due to lack of data. A similar study, by Bhattacharyya and Timilsina (2010) focuses on rural demand analysis and found very few applications on modeling demand estimation for rural areas. The authors argue that key issues, such as the poor–rich and urban–rural divides, are often poorly reflected in these models. They consider energy accounting models more reliable than the econometric approach but conclude that both suffer from data deficiencies in developing countries. Generally speaking, research on electricity demand estimation (Mentis et al., 2017; Moksnes et al., 2017; Riva et al., 2018) has converged on the limitations of an econometric approach for Sub-Saharan Africa (SSA) since, usually, generation that is not lost is consumed. Factors like demography must also be considered. In a review of integrated renewable energy systems (IRES), Riva et al., (2018) point out the importance of considering future energy demand in the research, including for education and health purposes. According to these authors, this makes bottom-up approaches that are less dependent on historical data, more reliable for the purposes of demand estimation.

1.1.2.1 Estimating demand using a bottom-up approach

Moner-Girona et al., (2019) propose a bottom-up approach for projecting energy demand by using three different types of electricity consumption patterns, based on the allocated population, index of poverty and the social and productive infrastructure. Mentis et al. (2016) also tackled the issue of demand estimation. They conducted a study in Nigeria for the time period from 2022 to 2030 and used the current population as well as growth rates and their own criteria for classifying settlements as rural or urban. The authors recommend that this classification should be done on a country-by-country basis since it implies differences in the way of life of the populations, and their energy demand patterns. Moksnes et al. (2017) presented an approach to optimize residential electricity demand for two levels of demand, urban and rural, in line with World Bank tier framework (Bhatia & Angelou, 2015). These tiers of access, also used by Nerini et al. (2016), are a simplification of the energy progress report produced by the International Renewable Energy Agency (IRENA, 2018) in its yearly editions. These tiers try to adjust the estimated demand according to the potential applications for electricity (for example urban vs. rural uses) and relative poverty, in each location.

In order to classify demand, Mentis et al. (2017) adjusted the population, for several countries, to reflect the numbers for 2030 and employed demand patterns, representing different levels of access to electricity services, from basic lighting to services that provide comfort such as air conditioning. These levels or tiers were adapted from official SE4all goals (United Nations Development Programme, 2015b). The process for estimating demand used geospatial data for administrative boundaries and population density was assigned to point locations, representing one squared kilometer (km^2) of territory. Data on the existing infrastructure (transmission lines) and general national access to electricity was also used to establish a baseline for the study. The population for 2030 was split into different tiers of demand pattern to produce demand for each point. Demand per point was determined multiplying 2030 population in a location by access tier yearly consumption level.

For efficient estimation of demand in a large scale, in a Nigerian case study, Blechinger et al., (2019) used automated algorithms based on GIS and statistical tools. Load profiles and socioeconomic data were derived from literature and expert opinions.

Regardless of the approach there is consensus in literature on the importance of considering future growth and changes in demand patterns.

1.1.2.2 Multi-period demand estimation

In a review of IRES, Riva et al. (2018) point out the importance of considering future energy demand. They propose the following taxonomy to classify the case studies concerning rural applications:

i) *Fixed demand* category – no evolution of demand in the horizon of the study;

ii) *Arbitrary trend* – Assumption that demand will grow at a fixed rate, in each year of planning. The trend is often taken from national plans and goals for the energy sector;

iii) *System dynamics (SD)* models – try to capture nonlinear behavior in demand growth over time;

iv) *Input/output models (I/O)* – used for macro-economic top-down analysis. With limited applications for rural applications since it cannot model non-formal and non-monetary transactions.

The authors found that only 25% of long-term studies consider variations of demand in the period of the study. Among these, the most often used method is the arbitrary trend. However, this method also presents limitations in rural settings, namely, the determination of the adequate growth rate of demand at each time-step. In order to deal with this uncertainty, arbitrary trend methods are often used in combination with scenario analysis. The authors pose that planning in rural settings should consider firstly socio-economic dynamics caused by access to energy and cost reduction due to technology changes. Secondly, the demand model must take into account the specific functions and appliances used in a rural setting.

1.1.2.3 A classification for demand modeling approaches

Morrissey (2019) reviewed several usual low-cost electrification modeling (LCEM) approaches and modeling inputs are enumerated, demand being one of them. According to the research, LCEMs have three ways of estimating demand:

i) One approach estimates future demand from current demand considering factors like population density, income/poverty level and urban/rural location. Some of these models often account for household demand while others account for the needs of markets, schools and health centers in unconnected areas;

ii) Another approach is to ignore demand using instead a predetermined Levelized Cost of Electricity (LCOE) or using population density as a proxy for demand (Bertheau et al., 2017);

iii) A third approach assumes a constant level of demand, for each time step, for all households. It is also possible to distinguish rural and urban demand but maintaining both constant over time. Demand level can be derived from using access goals, like IEA level of basic access or the World Bank’s Tiered Energy access Framework (Bhatia & Angelou, 2015).

All of these approaches emerge as valid, although there are no clear criteria to evaluate which one is the best one (Morrissey, 2019). It should be selected on a case-by-case basis, considering context.

1.1.2.4 Clustering of the population according to demand patterns (Tiers)

Beyond establishing demand tiers, Ciller and Lumbreras (2020) also found that clustering is necessary to group the consumers and determine the vertices that may be candidates for mini-grids, grid extension and isolated systems. A proper clustering procedure should consider all costs associated with the mini-grid, which are related to generation sizing and network design.

Population settlements can also be joined into clusters according to certain criteria (geographical distance, cultural proximity or consumer patterns, for example). In a World Bank study, Bhatia &

Angelou (2015) propose a set of tiers of access to energy, each one representing a different demand pattern. These tiers are often used in literature. Each one comes with implied trade-offs in terms of access, usefulness and cost. Rural needs for access to electricity may never be identical to those of a large city, but some level of access is needed, and none is unacceptable. The energy tiers, and corresponding technologies as proposed by these authors, are given in the Table 8 and Table 9, respectively.

Table 8 – Access to Household Electricity Supply (capacity and technology by tier)

Tiers	Tier 0	Tier 1	Tier 2	Tier 3	Tier 4	Tier 5
	No capacity	**Very low capacity**	**Low capacity**	**Medium capacity**	**High capacity**	**Very high capacity**
Minimum daily supply capacity	-	*5 watts*	*70 watts*	*200 watts*	*800 watts*	*2000 watts*
	-	*20 watt-hours*	*275 watt-hours*	*1 kilowatt-hour*	*3.4 kilowatt-hour*	*8.2 kilowatt-hour*
Typical supply technologies	-	Solar lantern	Rechargeable battery	Medium solar home system	Large solar home system	Large fossil-based generator
	-	-	-	Fossil fuel-based generator	Fossil fuel-based generator	Central grid
	-	-	Solar home system	Mini-grid	Mini-grid	-
	-	-	-	-	Central grid	-

Adapted from Bhatia and Angelou (2015)

Table 9 – Access to Household Electricity Services (appliances by tier)

Tiers	Tier 0	Tier 1	Tier 2	Tier 3	Tier 4	Tier 5
Typical appliances by power load						
Lighting	-	Task light	Multipoint general lighting	-	-	-
Communication and entertainment	-	phone charging, radio	Television, computer printer	-	-	-
Cooling and heating	-	-	Fan	Air cooler	-	Air conditioner, space heater
Refrigeration	-	-	-	Refrigerator, freezer	-	-
Mechanical loads	-	-	-	Food processor, water pump	Washing machine	Vacuum cleaner
Product heating	-	-	-		Iron, hair dryer	Water heater
Cooking	-	-	-	Rice cooker	Toaster, microwave	Cooker

Adapted from Bhatia and Angelou (2015)

IRENA has similar, if not exactly identical, set of tiers (IRENA, 2018). The World Bank (Bhatia & Angelou, 2015) and IRENA (IRENA, 2018) are the two main references that are used to established demand patterns, for planning purposes, when real world data is not present.

1.1.2.5 Demand evolution over time

In a literature review, Morrissey (2019) stated that some studies take into account that as more people have access to electricity their demand patterns increase. For that reason, these studies consider how demand has historically increased after connection, if the data is available.

Demand growth estimation is an often-ignored aspect of demand modeling. When it is considered, it is usually in one of two ways (Morrissey, 2019):

- Using historical data from demand increase, over time;
- Using a standard estimate of how demand might grow.

In the case of SSA, demand growth is limited by supply, and consequently, the second avenue is often the one taken. For example, Moksnes et al. (2017) used a linear growth rate for residential demand from 2022 to 2030.

Since demand needs to be supplied, a literature review on supply mix selection was carried out. It is given in Subchapter 1.1.3.

1.1.3 Supply mix selection and estimation

Beyond economics, there are other reasons for the inclusion of technologies in the *supply mix*, usually related to common factors identified in the most recent studies on energy planning in SSA. Most studies include grid extension and photovoltaic energy, in different configurations. The most common alternatives are stand-alone photovoltaic (PV) energy, diesel energy and PV mini-grids.

In these studies, according to the literature, the main challenges for the selection of technologies in the supply mix are:

i) The availability of the resources: for example, wind energy is not always available, in all countries. At least not at a level that enables its inclusion in the supply mix. There some case studies on wind energy for Latin America (Colombia and Venezuela) (López-González et al., 2019; Mamaghani et al., 2016), but it is rarely used in national level studies for SSA.

ii) The concern with prioritizing the inclusion of renewable and/or sustainable energies in the mix, due to environmental considerations;

iii) The availability of data (technologic, economic and geospatial) that allows for the analysis of the resource and its inclusion in the energy supply mix. Some authors are enthusiastic about the large hydro potential in SSA (Korkovelos et al., 2018) while others (Szabó et al., 2013; Morrissey, 2019) propose that hydro energy, in particular, will not be a part of the supply mix in SSA due to inaccurate or inexistent data, in most countries;

iv) The difficulty of computing the cost of diesel. Despite being usually incorporated in research and representing a large recurring cost, it is difficult to compute the cost of diesel in terms of pump price and transportation price, especially in remote areas. Some authors (Szabó et al., 2013) compute the transportation cost as a function of time of travel. For example, the time it takes from the closest vertex of access to the grid to a given settlement. Other cost components are linked to the operation and maintenance of generators. All of these elements need to be considered. Oil prices are difficult to predict, and fixed prices and scenario analysis with high and low prices are commonly used (Morrissey, 2019).

These factors need to be considered when proposing a supply mix for any country.

After the selection of the supply mix, there is a need to quantify the supply load. It is useful to consider that in any supply system (grid or mini-grid, renewable or not), energy is lost (Blechinger et al., 2019; Korkovelos et al., 2019; Morrissey, 2019). Thus, to estimate the required supply level it is important to consider not only the demand, but also losses. This effect is given in the literature, in its simplest form (Korkovelos et al., 2019) as Equation (1):

$$\boldsymbol{Supply\ (kWh)} = \frac{\boldsymbol{Demand\ (kWh)}}{\boldsymbol{(1-loss\%)}} \tag{1}$$

Literature, manufacturers and suppliers of energy solutions are good sources of information regarding the percentage of energy lost in generation and transmission. This means that if demand values are available, estimating a value for supply, given the characteristics of a settlement, is possible. Diesel systems are generally considered the most inefficient, with the highest loss rates. On the other side of the spectrum are PV systems with almost negligible levels of loss. The loss of a national grid is a function of the system itself and has to be locally attained.

Often not all resources are available in all locations and there is scarcity of data regarding the application of mini-grid solutions. This impacts the supply mix selection and estimation.

Information for analyzing the global cost of investment in technology is crucial for decision makers. For large scale planning purposes, researchers often rely on the use of an approximated model that retains the main cost elements of the technology but avoids excessive complexity. The way in which the global cost of technologies can be incorporated in these models is discussed in Subchapter 1.1.4.

1.1.4 Cost structure and assumptions for energy solutions

Although supply mix selection and estimation are crucial, decision makers in government use public funds to address the needs of the population and the country. In order to be effective, planning in the energy sector should account for the cost of access to electricity, whether it is achieved by grid or mini-grid solutions. The choice of supply mix needs to take into account the characteristics of each country, as not all energy resources are equally available in every location. The main factors and assumptions, listed in the literature, about the cost of different technologies is now carried out.

1.1.4.1 The cost of grid extension in literature

Due to practical and computational reasons, in a nationwide study, there is a need for some simplification of technical and economic elements. This is common in the literature on energy planning. In its simplest form, an electric grid can be summarized in three stages: generation, transmission, and distribution (Figure 3) all of them involving costs.

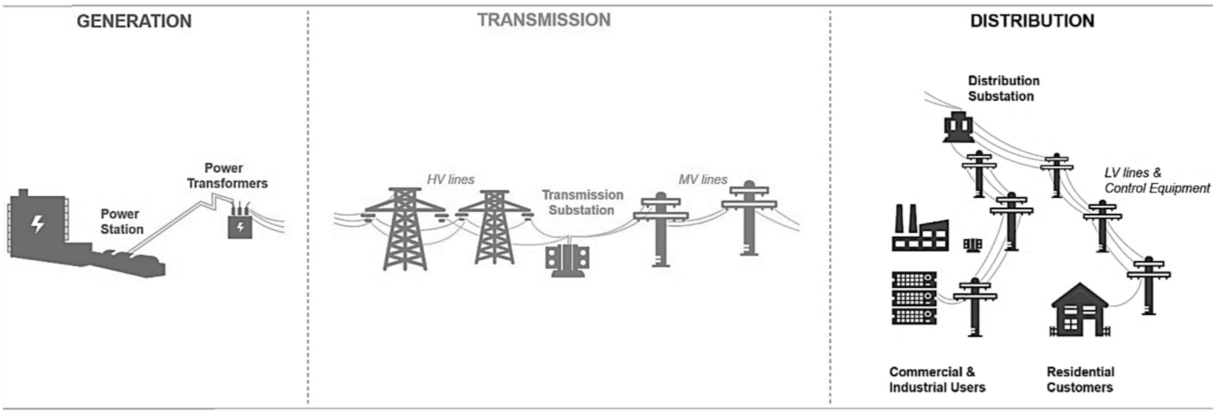


Figure 3 – Basic grid structure

Adapted from (Blume, 2007) using art from https://www.pinterest.com/

In a broad sense the cost of grid extension usually includes all costs incurred during the equipment's lifetime (measured in decades, in the case of grid technology). Generation cost is typically either not considered or treated as an exogenous variable. This was the case for all national level studies in the literature review. One aspect that is salient in the literature is that the calculations themselves are not always presented with absolute transparency. One reason may be that calculations are carried on in pre-made tools like Network planner, OSeMOSYS and ONSSET, and only the outputs are given.

Some notable exceptions are Deichmann et al. (2011) and Korkovelos et al. (2019) who presented their calculations more clearly.

Bearing in mind economic and budgetary reasons, one of the objectives of planning for electricity is usually the minimization of the cost of access. For grid expansion, the main inputs and assumptions, for the studies that were reviewed, are aggregated in Table 10.

Table 10- Main inputs and assumptions for studies on Grid extension

Cost structure element / input	**(Deichmann et al., 2011)**	**(Modi et al., 2013)**	**(Mentis et al., 2016)**	**(Nerini et al., 2016)**	**(Moksnes et al., 2017)**	**(Dagnachew et al., 2017) *(low cost assumption)***	**(Blechinger et al., 2019)**	**(Korkovelos et al., 2019)**
Line lifetime (years)	NS	30	30	30	NS	NS	20	30
Transformer lifetime (years)	NS	10	NS	NS	NS	NS	NS	NS
High-voltage line cost ($/km)	192,000 (220kV), 90,000 (132kV)	NS	53,000 (108kV), 28,000 69kV)	53,000 (108kV), 28,000 (69kV)	92,823 (kV not stated)	28,000 (132kV)	NS	28,000 (69 kV)
Medium-voltage line cost ($/km)	106,154 (132kV), 23,000 (33kV), 20,000 (11kV)	40,000 (kV not stated)	9,000 (33Kv)	9,000 (33kV)	9,000 (kV not stated)	9,000 (33kV)	20,000 (kV not stated)	13,000 (33 kV)
Low-voltage line cost ($/km)	10,611 (kV not stated)	40,000 (kV not stated	5,000 (0.2kV)	5,000 (0.2kV)	5,000 (kV not stated)	5,000 (kV not stated)		10,000 (0.2 kV)
Transformers	$21,818 – $60,000/unit	$105/ grid kW	$5,000 / 50 kVA	$5,000 / 50 kVA	NS	$5,000 / km	2 transf. per cluster. 100 USD / kW	3,500 (kVA)
Connection cost for grid ($/hh)	NS	25	125	125	NS	100	400/customer	150
Connection cost for mini-grid ($/hh)	NS	100	100	100	NS	100	400/customer	125
T&D losses (%)	NS	19.2%	7–29%	10%	NS	NS	11%	10%
Distribution loss (%)	NS	12%	NS	NS	NS	NS	NS	2%
Distribution O&M cost (% installation	NS	3%	2%	2%	NS	NS	NS	2%
Transformer O&M cost (% installation cost)	NS	3%	NS	NS	NS	NS	NS	NS
Generic Grid cost	Szabó et al. (2013) assume a generic grid construction cost (i.e., for all technologies) of €0.025/kWh/km.							

NS – not stated

Produced by the author

These elements can be used to produce a cost structure for the national grid. In the case of so-called intermediate studies (Ciller & Lumbreras, 2020) the grid is represented using its main cost elements in a non-exhaustive, but useful, way.

1.1.4.2 National grid design – simulating grid growth for planning purposes

The main cost factors of renewable technologies, like solar and wind, are well documented, either in the literature or through local suppliers. However, this is not the case for grid extension since the future design of the grid is unknown. Intermediate level studies (Ciller & Lumbreras, 2020) often use a simplified prediction of the grid. This involves having a simulation that represents grid extension, and its cost, in a non-technically accurate way. To model grid extension, algorithms are used in literature. At a national level, MILP, heuristic and graph theory approaches are the most frequent. Examples are Deichmann et al. (2011) who used the Kruskal's algorithm where at each stage the highest payoff is chosen. Other authors, like van Ruijven et al. (2012), emphasize the importance of using GIS tools for computing distances, since data is usually given with coordinates, in degrees, and their actual size (in kilometers) varies around the globe (smaller at the poles and larger at the equator).

The exact way and tools used to this end were not disclosed. Blechinger et al. (2019) propose a network in which an MST is used to connect locations by minimizing global cost. Solutions to these problems often use large amounts of memory, meaning some algorithms can only be applied to a limited amount of grid vertices. To overcome these limits Corigliano et al. (2020) used an iterative procedure that combined Kruskal's and Dijkstra's algorithms where the MST tree was initially used, using cost as function of inter vertex distance. Many studies in the literature use ready-made tools like Network planner, OSeMOSYS and ONSSET, which means they tend to have similarities and it can be relatively difficult to grasp the more technical aspects, often implying reading software documentation. There is a steep learning curve to this type of approach and the results tend to be partly formatted by the tool, which is undesirable. This is one of the reasons why ready-made tools are not always employed in literature.

1.1.4.3 The cost structure of photovoltaic energy in literature

Photovoltaic (PV) devices convert light into electricity (Blume, 2007). The key components of a PV power system are the solar panel (photovoltaic cells), the inverter (required for most mini-grid systems), the storage battery and the charge controller (Figure 4).

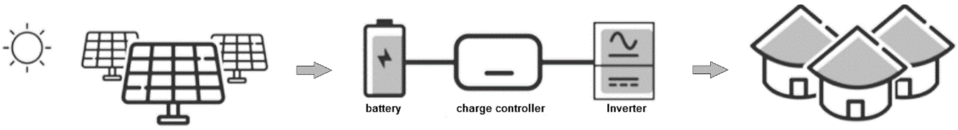


Figure 4 – Simplified scheme of a PV installation[10]

Adapted from Mamaghani et al., (2016)

The fundamental data assumptions, in terms of cost, from the literature are not always transparent. Often it is necessary to carry out additional reading to understand the full scope of the research. Some examples of the assumptions used in literature, are given in Table 11. Naturally, values vary over time and across locations.

Table 11- Main inputs and assumptions for studies on PV energy solutions

Cost structure component	**Modi et al. (2013)**	**Szabó et al. (2013)**	**Mentis et al. (2016)**	**Moksnes et al. (2017)**	**Moner-Girona et al. (2019)**	**Blechinger et al. (2019)**	**Korkovelos et al. (2019)**
PV lifetime (years)	20	NS	15-20	NS	15	20	15
Battery lifetime (years)	3	4	NS	NS	NS	NS	NS
PV module - panels (stand-alone) capital cost	$1,000/kW	1,000 Euro/kWp	NS	NS	NS	NS	NS
PV system - panels + charger + inverter (stand-alone) capital cost	NS	NS	$5,500/kW	$1,633/kW	5,500/kW	1,250/kW	5,500/kW
PV system mini-grid capital cost	NS	NS	NS	$1,363/kW	4,300/kW	1,250/kW	4,300/kW
Installation and other components	50% of PV price	800 Euro/kWp	NS	NS	NS*	NS	NS*
Battery price	$213/kW	1.5 Euro/Ah	NS	$1,688/kW	NS*	Capacity: 250/kWh	NS*
PV O&M	NS	2.5% of PV + battery price	NS	$10/kW	NS*	NS*	NS*

* All costs included in capital cost. NS – not stated

---

[10] Art from: https://medium.com/i-dev-insights/electrifying-africa-a-brief-introduction-to-solar-the-opportunities-article-1-of-3-9604dc450301.

Produced by the author from several sources

Different types of PV solutions are suitable for rural areas, for both household and agricultural applications (like water-pumping stations) but all suffer from seasonal fluctuations. To mitigate these fluctuations they are often coupled with alternative sources of energy (Mamaghani et al., 2016). It is also feasible to implement these types of solutions in communities that are close to the grid, but have relatively low population density. The grid will likely not reach these communities, within a reasonable period.

In very undemanding communities, it is even possible to implement standalone PV solutions without battery or inverter (Figure 5). The output is limited and unreliable. It can be applied for Tier 1 (IRENA, 2018) applications like powering a task light, phone charging and radio, for a few hour per day.

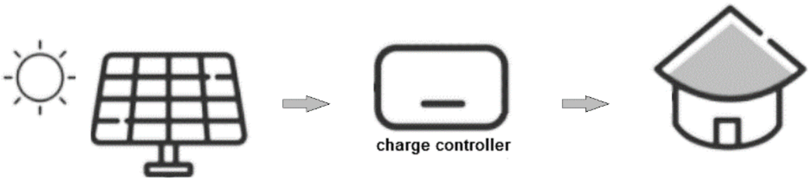


Figure 5 – Simplified scheme of a PV installation[11]

Adapted from Mamaghani et al. (2016) using art from https://www.pinterest.com/

These standalone home systems (SHS) are necessarily simpler than PV for mini-grid. Assumptions are based in Tier 1 (IRENA, 2018), 5Wh peak and are listed in Table 12.

---

[11] *Art from https://medium.com/i-dev-insights/electrifying-africa-a-brief-introduction-to-solar-the-opportunities-article-1-of-3-9604dc450301.*

Table 12 - Inputs and assumptions from literature on solar home systems (SHS)

Capital Cost for PV systems (aerial lines)	**USD**	**source**
Solar Home system (10W)	**80**	*http://d-lab.mit.edu/solar-lighting/*
Lifetime (years)	2	
residual value	0%	
Recurring Costs	**Input**	
O&M – PV system	(8 USD) 10% of investment cost	
Brands / Models	Pico Home System PSHS 7000 (http://www.fosera.com/)	
WACC	-	Public debt rate of return

Produced by the author from several sources

For energy planning purposes in SSA, PV energy is the most prevalent alternative to grid extension, by itself or coupled with another source of energy as remarked in the literature. The main reason is the availability of data on solar power for most countries, and the relative ease of sizing and cost calculation.

1.1.4.4 The cost structure of diesel energy in literature

One of the main criticisms leveled at solar and wind energy is their unpredictability. This can be mitigated with the addition of battery sets and/or a diesel generator to the system. Although the environmental and health hazardous effects of diesel generation are documented, they are essential to support services such as health services in isolated areas. In the literature they are often included, as backup, in mini-grids, paired with more clean solutions, like PV or Wind, since for some communities, power may be unreliable without the inclusion of a diesel generator.

The cost of diesel electricity, for planning purposes is connected to a few factors, as given in Table 13.

Table 13 - Main inputs and assumptions for studies on Diesel mini-grids from literature

	(Parshall et al., 2009)	(Szabó et al., 2013)	**(Mamaghani et al., 2016)**	**(Moksnes et al., 2017)**	**(Moner-Girona et al., 2019)**	**(Blechinger et al., 2019)**
Generator for mini-grid	1000 USD/kVA	NS	6,000 USD (10 kW)	721 USD/kW	The same methodology and assumptions as Szabó et al. (2013)	820 USD/kW
Brand	NS	NS	Generac	NS		simulation software
Indicative capacity	NS	NS	NS	100 kW		NS
Replacement cost	NS	NS	5,000 USD (10 kW)	NS		NS
Capacity factor	NS	NS	NS	0.70		NS
Efficiency	NS	NS	NS	33.00%		30% - 35%
Scaling factor	0.64	NS	NS	-		NS
Installation	25% of generator cost	NS		NS		NS
Civil engineering	1,667	NS	NS	NS		NS
Fuel tank	1,741	NS	NS	NS		NS
Lifetime	5 years	NS	85,000 hours	15		NS
O&M USD/year	5.00%	0.01 Euro / kWh	0.075 USD/hour	10.00%		0.05 USD/kWh
Diesel consumption (liters/kWh)	0.40	pump price + transport cost	NS	NS		
Fuel (USD/liter)	1.05		1.10	1.20		0.68 USD/liter

NS – not stated

Produced by the author from several sources

Most authors agree that diesel generators should not be seen as the only solution to ensure access to electricity in isolated communities, but they are historically essential as backup given the unreliability of some renewable sources.

Considering the excesses in electricity consumption in other regions of the globe (Smil, 2010) perhaps, for these impoverished communities, environmental factors should not be central to the decision-making process.

1.1.4.5 The cost structure of hydro powered mini-grids in literature

Hydroelectricity accounts for over 20% of the world's electricity consumption (Smil, 2010). The potential energy from falling water is converted into kinetic energy, which turns the blades, or vanes, in a hydraulic turbine. The turbine turns the generator rotor producing electricity (Blume, 2007). Research on mini-hydro potential is not as common as other resources. Some authors exclude the option due to lack of data regarding the cost structure of hydropower (Deichmann et al., 2011). Other authors present the final total cost of the hydropower mini-grid. Few studies detail its cost components. Morrissey (2019) carried out a literature review on the economics of mini-grid solutions for electrifying SSA, while (Korkovelos et al., 2018) focused solely on hydro potential in SSA. The authors noted that given the absence of data on micro-hydro their role is limited in the process of electrification, even though it has a relatively low cost (Nerini et al., 2016). Significant generation potential from micro-hydro has been identified across SSA, including Angola. The challenge is in gathering reliable data for identifying small scale hydroelectric potential as this source of energy requires the existence of water, ideally in elevation, so that it can be channeled to a lower area (Figure 6).

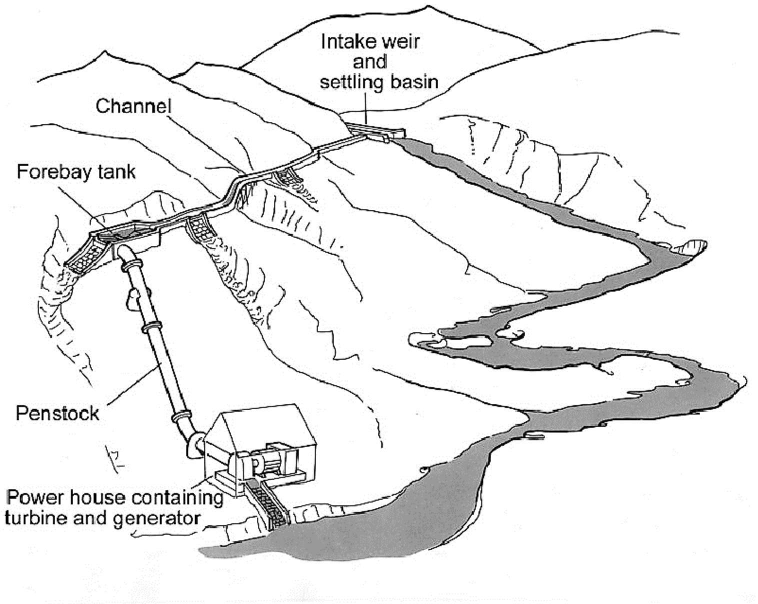


Figure 6 – How Mini-hydro installation

Adapted from https://www.intechopen.com/books/energy-conservation/hydro-power

Even with these challenges and limitations, some of the cost assumptions that have been used in literature are presented in Table 14.

Table 14 - Main inputs and assumptions for studies on hydro mini-grids from literature

	Mentis et al. (2016)	**Nerini et al. (2016)**	**Moksnes et al. (2017)**	**Moner-Girona et al. (2019)**	**Korkovelos et al. (2019)**
Mini-Grid Hydro (Investment - USD/kW)	Assumptions extracted from (Nerini et al., 2016)	5,000	***< 10 MW*** 2,902 USD (invest.) + 2.05 USD/kW (fixed cost)+4.464 USD/MWh (variable cost)	Assumptions extracted from Szabó et al. (2013)	5,000
O&M cost		2%	2%		2%
Capacity factor		NS	50%		0.5
Life (years)		30	NS		30

NS – not stated

Produced by the author from several sources

Metrics are required to measure and compare the total cost of these different technologies. In Subchapter 1.1.5 some of the most useful and frequently used metrics are discussed.

1.1.5 Decision making metrics in energy planning

A significant number of studies on energy planning have cost minimization as the objective. The more prevalent economic decision-making metric is the Levelized Cost of Electricity (LCOE) often paired with other metrics as listed in Table 15.

Table 15 – Recurring economic analysis tool in literature

	NPC - net present cost	LCOE - Levelized cost of electricity	Sensitivity analysis and Scenarios
(Parshall et al., 2009)	•		•
(Szabó et al., 2013)	•		•
(Mentis et al., 2016)		•	•
(Nerini et al., 2016)		•	•
(Mamaghani et al., 2016)	•	•	•
(Moksnes et al., 2017)		•	•
(Mahbub et al., 2017)	•		•
(IRENA, 2018)		•	
(Trotter et al., 2019)	•		
(Moner-Girona et al., 2019)	•		•
(Blechinger et al., 2019)		•	
(Korkovelos et al., 2019)		•	•

Produced by the author from several sources

Very few nationwide studies on energy planning for SSA consider the environmental effects (e.g. CO_2 emission). Most studies propose the introduction of environmentally friendly technologies, but do not quantify their benefit. In SSA, the main decision criterion is cost. This is expected considering the stage of development of the countries in this region. Thus it is essential to choose adequate metrics for measuring the economic impact of access to electricity.

1.1.5.1 Net Present Cost (NPC)

Public investment is constrained by the national budget, so a measure of global cost of investment over time is essential.

One measure that is widely used to measure the economic feasibility of projects is the Net Present Cost (NPC). It represents the sum of a stream of discounted cash flows (exclusively outflows in this case), often occurring at different moments in time (Khan et al., 2007). The present value of a cash flow depends on the interval of time between a given moment in time and the moment the cash flow occurs. It also depends on the discount rate. The discount rate can be interpreted as the adequate average return that the entities investing or financing a project should demand for taking risk. It provides a method for evaluating and comparing capital projects over time. Its general mathematical formulation is given in Equation (2).

$$NPC = \sum_{t=0}^{n} \frac{CF_t}{(1+d)^t} \quad (2)$$

where,
CF_t is a Cash Flow occurring at moment t
d is the discount rate
n is the last year of the period of analysis

This metric is a standard in project economic feasibility studies (where it is generally called Net Present Value) and it is a useful reference when comparing alternative investment decisions. Like so many of these metrics it is less useful in a vacuum, or without context, so it should be used along with other metrics and methods, such as, the Levelized Cost of Electricity (LCOE), sensitivity analysis and scenario analysis.

1.1.5.2 Levelized cost of electricity (LCOE)

For energy planning, it is essential to have a metric that allows for the cost comparison of different generation and delivery technologies. The LCOE is used in almost all relevant studies in this field that include economic analysis (including engineering studies). It is calculated by adding all the costs associated with an energy generation source (costs of capital, fuel, operation and maintenance (O&M)), in currency, such as USD, and dividing them by the amount of electricity produced (in kWh), in its lifetime. This calculation results in an LCOE expressed as a currency-specific cost per kilowatt-hour (USD/kWh). A discount rate is also applied to account for the time value of money. LCOE takes into account the expected demand for energy, not the maximum generation capacity. This is important since each user reaches peak demand at slightly different moments in time. LCOE can be found explicitly (Branker et al., 2011) assuming a constant value per year in Equation (3).

$$LCOE = \frac{\sum_{t=0}^{n} \frac{C_t}{(1+r)^t}}{\sum_{t=0}^{n} \frac{E_t}{(1+r)^t}} \tag{3}$$

where,

C_t - Cost of the energy generating system and the energy generated over its lifetime ($);
E_t - Energy generated in the same period of time (kWh);
t – Time of the cash flow and energy generation;
r – Discount rate (%).

For solar energy, it is usual to consider a degradation factor for solar panels. Meaning they produce less energy over time. For solar systems, this expression can be further detailed as given in Equation (4):

$$LCOE = \frac{\sum_{t=0}^{n} \frac{(I_t+O_t+M_t+F_t)}{(1+r)^t}}{\sum_{t=0}^{n} \frac{S_t(1-d)^t}{(1+r)^t}} \tag{4}$$

where,

I_t – Investment costs ($);
O_t – Operational costs ($);
M_t – Maintenance costs ($);
F_t – Interest expenditures for t;
S_t – Rated energy output per year;
d – Degradation factor
r – Discount rate (%).
t – Time of the cash flow and energy generation.

This definition is identical to the one proposed by Aldersey-Williams and Rubert (2019) and can be interpreted as the minimum constant real price required for a project to achieve a target return. It is often compared with market tariffs, to assess the sustainability of the solution that is being analyzed. Grid parity happens when the electricity produced by an alternative energy source has the same price as that of traditional energy (electrical grid tariffs). It is achieved, in different areas and countries, at different moments in time. It depends on factors like the amount of sunlight or wind. Research has shown that, in SSA, even in extremely poor areas, people are willing to spend over 10% of their total income to have access to electricity (Sievert & Steinbuks, 2020). For example, if the LCOE for a PV solution is close to market tariffs, that would mean that the solution is competitive when compared to grid extension, since it would represent identical cost.

LCOE is an international standard for comparison of the cost of technologies and features prominently in scientific and even non-scientific documents (IRENA, 2018; Sievert & Steinbuks, 2020). It makes sense that it has relevant strengths:

- Simplicity – one metric allows for the comparison among different projects;
- Adoption – widely adopted, which is important for comparison purposes, among project and countries;
- Interpretation – it represents the minimum constant real price required for a project to achieve a target return. It is sensitive to the variability of the discount rate. This stems from using established concepts from economics and finance (discounted cash flows and the time value/ cost of capital).

LCOE, like many metrics, has many critics and limitations as well as some solutions to mitigate each criticism:

- LCOE does not address issues of *intermittency* associated with renewable energy. Ideally, electricity supply should be available 24 hours a day, every day, without blackouts. In this sense, LCOE does not account for the variability of renewables when compared to uninterrupted service from a stable electric grid. Dispatchable and non-dispatchable technologies should be evaluated separately and compared with caution. This criticism can be mitigated with the introduction of batteries or another alternative that compensates potential blackouts. This would allow for the approximation to true grid parity (Branker et al., 2011).
- There is not clear consensus on the best *discount rate* to be used, so it varies in literature, between the risk free rate and the weighted average cost of capital (WACC) (Aldersey-Williams & Rubert, 2019). The discount rate should reflect different technologies with differing risks and distinct financing solutions. However, it is useful to keep in mind that WACC can vary over the period of analysis.
- Another point of contention is the handling of inflation (Sklar-Chik et al., 2016). The incorporation of inflation can generate distortions of comparison between technologies over long periods, as they are not implemented in the same year. LCOE can include nominal

costs as long as a nominal discount rate is considered. Incremental analysis with time-steps can also mitigate this problem.

- One aspect that is often overlooked is the *variability of fuel costs* (Szabó et al., 2013). These can represent a large part of the cost of some projects. One way to deal with these issues are the use of sensitivity analysis, scenarios and the introduction of time-steps in the model.

The interpretation for LCOE hinges on the fact that the price target of energy needs to be constant over time. An "adjusted LCOE" that accounts for price variation over time has been introduced (Nissen & Harfst, 2019), and is given in Equation (5).

$$LCOE = \frac{\sum_{t=0}^{n} \frac{(I_t + O_t + M_t + F_t)}{(1+r)^t}}{\sum_{t=0}^{n} \frac{S_t (1-d)^t (1+epr)^t}{(1+r)^t}} \qquad (5)$$

where,

I_t – Investment costs ($);
O_t – Operational costs ($);
M_t – Maintenance costs ($);
F_t – Interest expenditures for t;
S_t – Rated energy output per year;
d – Degradation factor
r – Discount rate (%).
t – Time of the cash flow and energy generation;
epr –Expected rate of evolution for energy price.

According to the authors, this "Energy Price Adjusted LCOE" can now correctly indicate if Grid parity has been achieved. This approach requires a forecast of energy price evolution, expressed by a rate ("epr").

The last stage of the literature review is given in Subchapter 1.2 and concerns the methodology of planning for access to electricity, since this methodology is essential to model the support system for decision making. With this in mind the relevant literature was reviewed.

1.2 Methodology of planning for access to electricity – literature review

In literature, the methodologies used to solve access to electricity problems in developing countries employ diverse mathematical optimization techniques including Linear Programming (LP), Mixed Integer Linear Programming (MILP), Meta-Heuristic approaches (Ciller & Lumbreras, 2020; Trotter et al., 2019) and graph theory (Blechinger et al., 2019; Corigliano et al., 2020; Kemausuor et al., 2014). For the grid growth process, in planning studies, there is some consensus on the use of least cost approaches associated with algorithms that build what are generally called minimum spanning trees (Kruskal, 1956). In some studies, least cost does not refer to minimizing the total cost of investment, measured by the Net Present Cost (NPC) but instead they take into account the Levelized Cost of Electricity (LCOE) which is more associated with the long-term economic sustainability of energy systems. Solutions vary according to the use of different metrics.

Parshall et al. (2009) use Kruskal's algorithm in order to connect settlements within a given radius and with a preset maximum distance to the nearest grid access. Similarly, Deichmann et al. (2011) also used the Kruskal's algorithm to ensure only settlements within a technically feasible threshold distance are selected for grid connection. In both aforementioned approaches, the unconnected settlements will be selected for mini-grid connection. On the other hand, Mentis et al. (2016; 2017) use an analytic hierarchy process (AHP) with weights assigned to several types of geospatial data in order to build a decision metric, in order to drive the spanning tree's growth. Dagnachew et al. (2017) propose the Economic Distance Limit (EDL), a cost metric based on the Levelized Cost of Electricity (LCOE) as main driver for grid extension. Moksnes et al. (2017) use the Open-Source Spatial Electrification Tools (OnSSET[12]) to simulate grid extension in SSA. The metric that drives grid extension in OnSSET is the LCOE. This is equally the case for Blechinger et al. (2019) and Korkovelos et al. (2019) because all use the OnSSET tool for their research.

Despite the complexity of energy planning in SSA, few multi-objective models could be found in literature. One exception was found for the case of Uganda (Trotter et al., 2019). This study has a highly detailed generation model for the grid, usually not seen in research. These authors propose a model with three objectives: minimizing cost, minimizing regional asymmetries and minimizing urban/rural electrification inequality. The trade-off in this approach seems to be the granularity of the model which was at a district level. A district, in Uganda, is a third level administrative division, potentially containing many settlements. The problem, for 112 districts of Uganda, was solved in ILOG CPLEX Optimization Studio[13] (CPlex) and a Pareto-front was produced to aid decision makers in visualizing the alternatives. The complexity of the problem did not allow for a higher resolution

---

[12] *This tool is available at http://www.onsset.org/.*

[13] https://www.ibm.com/docs/en/icos/12.9.0?topic=mc-what-is-cplex

analysis (for example at settlement level). For this type of approach, applications with larger samples are absent from literature on energy planning. Another multi-objective approach in literature is proposed by Ciller and Lumbreras (2020) but only in conceptual terms since no real world application was performed.

The large majority of models consider few time steps, often just one (instantaneous models). This is a limitation for planning purposes since the implementation process takes years and little room is left for yearly readjustments.

It is important to take into account the trade-offs between the quality of the outputs (how much detail is given) and the complexity of the approach as well as the mathematical technics that best serve the purpose of the study and research objectives. Reviewing literature for this type of study is complex since it usually encompasses a large diversity of areas of knowledge and varied modeling technics across different studies. In order to facilitate the understanding on how these different dimensions may be integrated in research, a short review on elements of decision-making, modeling grid and off-grid layouts as well as the most frequent limitations and mitigation techniques is given in Subchapters 1.2.1 through 1.2.5.

1.2.1 Fundamental elements of decision making

Decision making often involves managerial and engineering dimensions and the priorities of decision makers can change over time. These priorities can lead to substantially different solutions, which can be controversial (Corigliano et al., 2020).

What is often presented to the decision maker is a large set of options, with diverse cases, in order to help the decision-making process. The process of making a decision implies weighting relative utilities from a set of alternatives and choosing the one that best suits an objective, for the purposes of implementation. The final decision is left to the decision maker.

According to the literature (Kumar et al., 2017; Malczewski & Rinner, 2015) in order for decision making to be possible some elements must be present:

i) *Decision makers*, who have inherent characteristics such as rationality, autonomy, reactivity, preferences and beliefs. These traits can be modeled in a computer to produce decision-making agents that can act in simulated real-world environments. This type of simulated approach goes beyond the scope of this study.

ii) *Criteria* (objectives, attributes). An objective can be seen as a statement on the desired state of a system under analysis. It gives information on the direction attributes should take to “improve”. This originates a maximization or minimization objective function. The attributes themselves are measures of a certain property that a location or a set of locations share. For example, the distance of a set of locations to the nearest water source.

iii) *Hierarchical structure*. The relation between objectives and attributes has a hierarchical structure. One typical structure, as proposed by Malczewski and Rinner (2015), and consists of four levels: goal, objectives, attributes and alternatives (Figure 7).

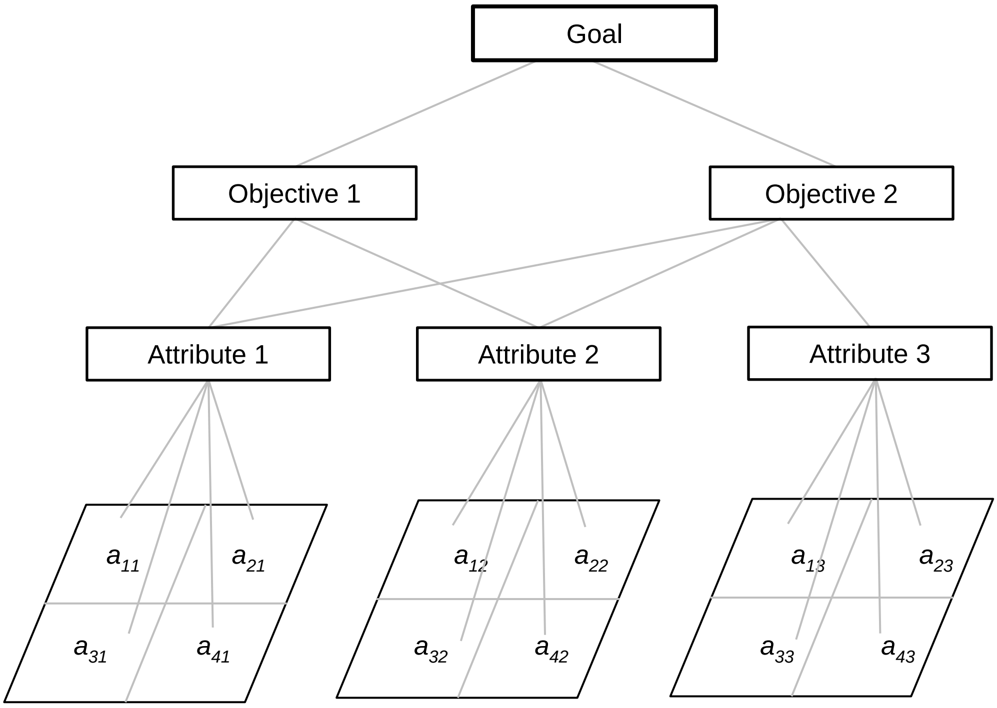


Figure 7 - Hierarchical structure of a decision problem: a_{ik} is the value of the *k*-th attribute (criterion) associated with the *i*-th alternative (*k* = 1, 2, 3 and *i* = 1, 2, 3, 4).
Adapted from Malczewski and Rinner (2015)

The decision alternatives (lowest level) are evaluated against the attribute level, so a decision can be made.

iv) Decision alternatives are evaluated in terms of their attribute and are classified as *feasible* or unfeasible. The feasible alternatives can themselves be divided into two categories: dominated and non-dominated solutions.

To mitigate the uncertainty associated with energy planning problems, scenario analysis can also be employed to present a series of different, non-exhaustive non-dominated solutions. It is often worthwhile to test the extreme cases to assess the reasonableness of a given approach in terms of global cost and the layouts for grid growth and off-grid energy solutions.

1.2.2 Modeling grid growth and off-grid layouts

Regardless of the number of scenarios employed, one of the more significant challenges of planning is the modeling aspect for grid extension growth and off-grid solutions layouts. Some studies sidestep this issue and only present numbers while others present final solutions (instantaneous models). Few present intermediate solutions over time. The modeling approaches for nationwide problems are necessarily a simplification of reality due to computational and practical limitations. Naturally, the simplification should not mean the models cannot serve their purpose of supporting decision making in planning activities. Graph theory has been extensively used for representing and solving grid extension problems. Informally, a graph is a set of vertices connected by links, which can be edges (undirected graph) or arcs (directed graph). Links connecting two vertices can be weighted, expressing a certain attribute between the two vertices (e.g., distance). The grid extension problem can be solved by looking for a minimum spanning tree in a graph. A spanning tree is a connected graph without cycles meaning that the number of links must be n-1, where n is number of vertices in the graph. A graph can contain several spanning trees, and the minimum spanning tree is the one that minimizes the total weight among all the existing spanning trees.

The most well-known algorithms for solving the minimum spanning tree problem are Kruskal’s algorithm (Kruskal, 1956) and Prim’s algorithm (Prim, 1957).

In the literature on energy planning Kruskal’s algorithm is prevalent. It selects edges sorted by non-decreasing weight, which do not form cycles, until a connected graph is attained. The steps required to build an MST with this algorithm are given in Figure 8.

Kruskal's Algorithm - steps

Step 1: Sort all edges in increasing order of their edge weights.

Step 2: Pick the smallest edge.

Step 3: Check if the new edge creates a cycle or loop in a spanning tree.

Step 4: If it doesn't form the cycle, then include that edge in MST. Otherwise, discard it.

Step 5: Repeat from step 2 until it includes |V| - 1 edges in MST.

Figure 8 – Steps involved in Kruskal's Algorithm to generate a minimum spanning tree

Adapted from Kruskal (1956)

Prim's algorithm spans a sub-tree starting with a randomly chosen vertex. In each iteration, it connects the sub-tree to the external vertex (a vertex not belonging to the current sub-tree) with the lowest connection weight. The steps required to build an MST Prim's algorithm are given in Figure 9:

Prim's Algorithm - steps

Step 1: Determine the arbitrary starting vertex.

Step 2: Keep repeating steps 3 and 4 until the fringe vertices (vertices not included in MST).

Step 3: Select an edge connecting the tree vertex and fringe vertex having the minimum weight.

Step 4: Add the chosen edge to MST if it doesn't form any closed cycle.

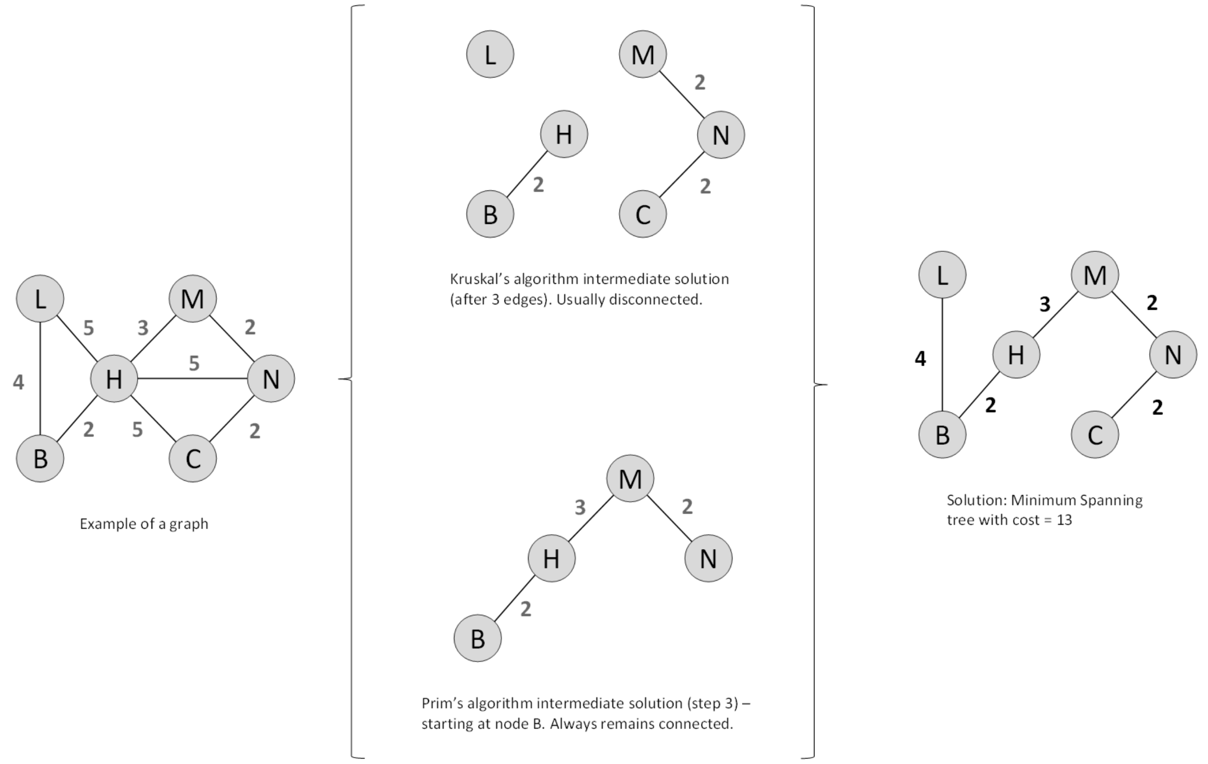
Step 5: Exit Return T

Figure 9 – Steps involved in Prim's Algorithm to generate a minimum spanning tree

Adapted from Prim (1957)

The main difference between these two algorithms is that Prim's algorithm current solution is always a connected graph, containing a sub-set of vertices whereas Kruskal's algorithm current solution is a set of connected graphs.

The differences between these two algorithms are further detailed in Figure 10.


Example of a graph

Kruskal's algorithm intermediate solution (after 3 edges). Usually disconnected.

Prim's algorithm intermediate solution (step 3) – starting at node B. Always remains connected.

Solution: Minimum Spanning tree with cost = 13

Figure 10 – Illustration of solution of Spanning tree problem using different algorithms. Produced by the author

Both algorithms provide an optimum solution to the problem. However, Prim's algorithm allows the visualization of the order by which the network grows which is well suited for multi-period approaches.

1.2.3 Challenges and energy planning in SSA

Energy planning in SSA presents numerous challenges beyond the modeling of layouts for grid and off-grid solutions. In the literature, it was possible to identify some fundamental challenges and specificities of energy planning in developing countries:

(i) The current crisis of reliable data, in many developing countries, means the location and size of the population is unknown, outside large urban areas. Often, there is little or no data on demand and availability of technologies and resources in different regions of SSA (Grippa et al., 2019; Linard et al., 2011; Moner-Girona et al., 2019; Stevens et al., 2020; Szabó et al., 2013). This makes planning for development (and people) extremely difficult.

(ii) Including future trends in demographics, technology and economics, for example, at national and international levels. Planning for future, not current demand means considering growth in the population and demand itself (Dagnachew et al., 2017; Morrissey, 2019).

(iii) The need to consider alternatives beyond just grid extension, including all available possibilities in any given location. Diesel should be included since it still is a cornerstone of energy supply in SSA. Research on the economics of renewable energy expansion in SSA often looks for a solution that reconciles accelerating the electrification of SSA's with the need to keep CO_2 emissions in check (Blechinger et al., 2019; Korkovelos et al., 2018; Trotter et al., 2019).

(iv) The need to translate results into costs – decision making implies the use of limited economic resources. Planners usually spend large amounts of time to obtain reasonable estimates of the cost of electrification in administrative areas. This holds for the majority, if not all, of the literature (Blechinger et al., 2019; Cuisinier et al., 2020; de Roo et al., 2017; Deichmann et al., 2011; Korkovelos et al., 2018, 2019; Mentis et al., 2016; Modi et al., 2013; Parshall et al., 2009).

(v) Finding suitable methods and metrics and the importance of including all cost that occur during the lifetime of the equipment (Bertheau et al., 2017; Blechinger et al., 2019; Ciller & Lumbreras, 2020; Corigliano et al., 2020; Deichmann et al., 2011; Moner-Girona et al., 2019; Parshall et al., 2009; Trotter et al., 2019).

(vi) The importance of producing or using tools that generate outputs that can be quickly employed by decision makers to select the best electrification options, for each settlement or location (Corigliano et al., 2020; Kumar et al., 2017).

(vii) These tools can integrate different methods and types of information, often making comparisons difficult (Corigliano et al., 2020; Kumar et al., 2017).

Due to all these challenges, many of the studies in developing countries present some limitations.

1.2.4 Limitations of current models

Most of the models in the literature consider only a single time step, called overnight-build models (Morrissey, 2019). These models ignore that grid extension takes longer than mini-grid solutions which implies people will have to be without access to electricity for longer periods. They also do not take into account process dynamics such as:

i. Population growth – due to natural growth and migration;
ii. Increasing demand - newly connected households tend to see their demand increasing over time. According to Morrissey (2019) this can be assessed through data about the usage of utilities (Kemausuor et al., 2014; Parshall et al., 2009; Sanoh et al., 2012) or by using a standard estimate of how demand will increase (Moner-Girona et al., 2019);
iii. Future diesel prices – Volatility of prices is a reality that is rarely incorporated in models. Usually, scenarios with high or low prices are presented;
iv. Decreasing capital costs – usually costs at the time of writing are used while the cost of renewable components is expected to fall and, consequently the role for renewable generation is underestimated. This implies models have to be updated. One example on how to mitigate this issue is given by Dagnachew et al. (2017) who applied a historically derived learning rate, over every time-step, in the proposed model.

Minimizing the aforementioned issues leads to more useful decision-making support models.

1.2.5 Approaches to the mitigation of the inherent limitations of models

In spite of being a fundamental tool for policy and decision makers electrification models should be analyzed taking into account their limitation, their assumptions and the data that is available. After all, planning models are simplifications of reality and data regarding SSA is scarce and often its accuracy is poor (Morrissey, 2019). Across the board, different studies present similar trends, for the same country, but significantly different results. This difference can be attributed to the mathematical formulation of the problems, the solutions used for missing data, the simplification employed to ease computational load and the use of different technologic supply mixes. Additionally, the inclusion of dynamic parameters for inputs like population and demand growth, horizon of analysis, discount rate, methods for estimating fuel price evolution or the decreasing capital cost associated with renewable technologies can also be reasons for these differences.

Some measures can be used by modelers to mitigate the risks associated with these challenges:

i. Referencing only models that have full disclosure of their methodology, as some models proposed in literature are not fully disclosed.
ii. Clearly stating and referencing all the assumptions used for the modeling approach;
iii. Clearly stating the sources and potential quality of the data as well as the processes of filling its gaps;
iv. Suiting the approach to the problem;
v. Employing some sort of scenario analysis with extreme cases. This allows the decision maker to clearly understand the general trade-off some options will imply and facilitates

the choice of a general direction, regardless of the exactness of the numbers.

vi. The model should be easy to iterate upon (update and improve).

This set of measures also contributed to address the uncertainty associated with these types of models.

CHAPTER 2

The case study: Angola

Angola is located in the western region of Southern Africa, occupying an area of approximately 1,246,700 km^2. This makes Angola the sixth largest country of Africa (INE - Instituto Nacional de Estatística, 2016). The extent of its coastline is of more than 1,600 kilometers, bordering the Atlantic Ocean. Angola has land borders to the North with the Republic of Congo and the Democratic Republic of the Congo, to the East with the Democratic Republic of Congo and Republic of Zambia, and to the South with the Republic of Namibia, with an extension of more than 1,400 km (Figure 11).

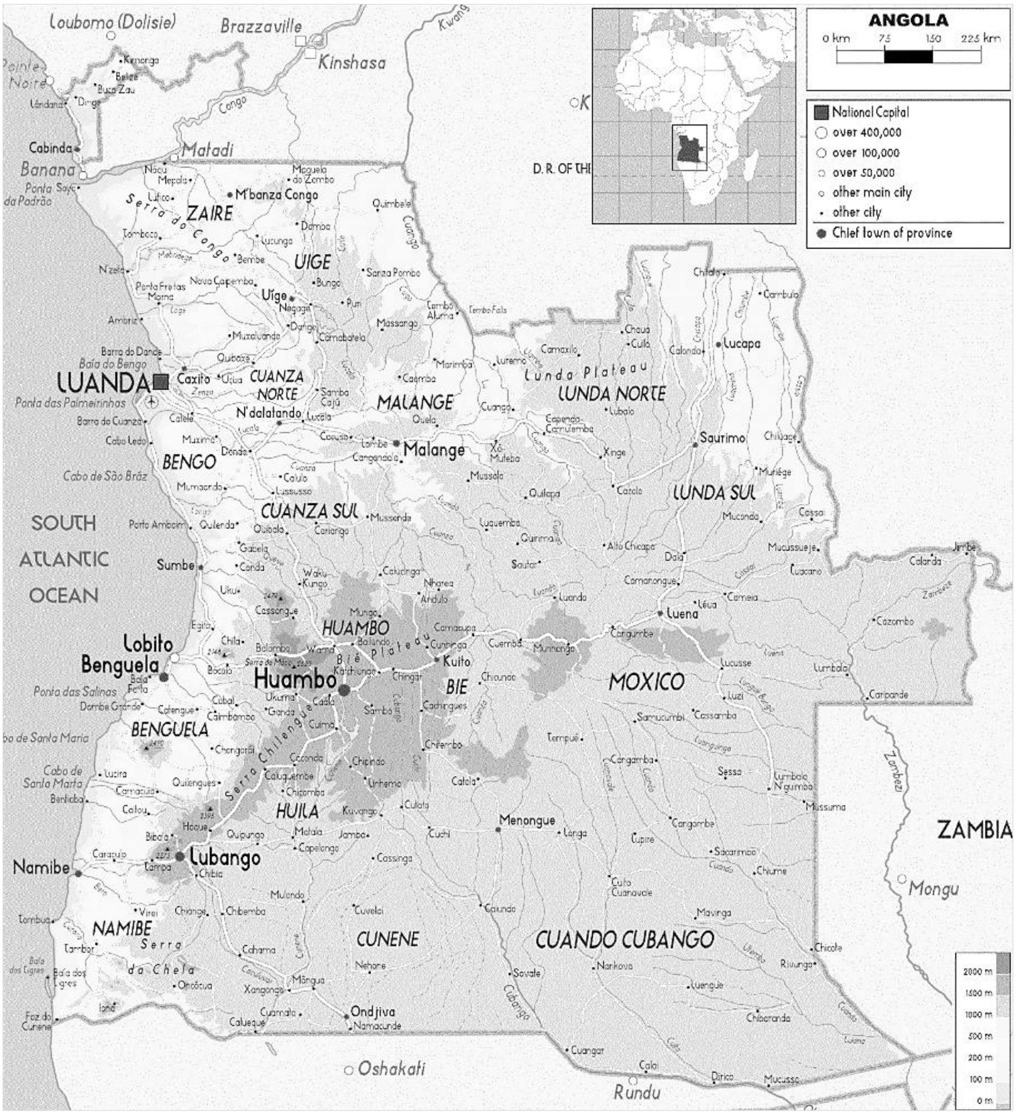


Figure 11 – Map of Angola

Extracted from: https://www.gif-map.com/maps/africa/angola/physical-map-of-angola-2.gif

Angola is divided into three administrative levels (Provinces – level 1; Municipalities – level 2; Commune – level 3). At the time of the census of 2014, the political and administrative division of Angola was composed of 18 provinces, 162 municipalities and 559 communes.

The capital city is Luanda with 2.5 million people in its urban core and an additional 4 million people in the surrounding territories. The country has six other large cities with populations ranging from 300.000 to 1.000.000 people. These cities are N'dalatando, Huambo, Lobito, Benguela, Cuito and Lubango.

The official language is Portuguese and the currency is the kwanza. The language in particular is a legacy from being a Portuguese colony until 1975 and the country has recently celebrated 40 years of independence.

Despite being a young nation, Angola is an important oil producer and is also rich in minerals, diamonds and iron in particular. The fishing industry, beer production, cement and timber are also highlights of the country's economy. In geopolitical terms the country's role is important in pacifying the region since it has one of the largest armies in the continent.

Despite the large hydro potential in Luanda, access to running water, sanitation and electricity is still restricted, and outside the city perimeter things are much worse. The country's economic growth until 2015 did not translate into social development and a real improvement of conditions of living of the population and as a result the population is very poor and there is social unrest. This is illustrated by the results from a study on multidimensional poverty that showed that 9 out of 10 people are poor in Angola (INE - Instituto Nacional de Estatística, 2019). This reality was one of the driving forces for choosing Angola for this study's case.

It is important to make a general characterization of the country in different dimensions: demographic, the energy sector and a general characterization of the electric system. All of these elements will be presented in more detail in Subchapters 2.1 through 2.3.

2.1 Angolan demographics

2.1.1 Population

The 2014 census reports that the total population of Angola is 25,789,025 people (INE - Instituto Nacional de Estatística, 2016). The census produced urban and rural population numbers, for each commune in the country. The largest portion of the population reside in urban areas (63%), where the average household has 4.8 people and the remaining (37%) reside in rural areas, where the average household has 4.4 people (Table 16).

Table 16 - Indicators from 2014 Census

Indicators	**Angola**	**Area**		**Gender**	
		Urban	*Rural*	*Men*	*Women*
Total population	25,789,025	16,153,987	9,635,037	12,499,041	13,289,983
Average number of people per household	4.6	4.8	4.4	-	-

Adapted from (INE - Instituto Nacional de Estatística, 2016)

According to INE (INE - Instituto Nacional de Estatística, 2018), the Angolan population has been growing at a rate above 3.0% since 1975 but in 2014 the population's growth rate decreased slightly to 2.7%. In 2014, the average life expectancy was of 60.2 years (57.5 year for men and 63 years for women) and the average age was 20.6 years. The country had diverse population densities through the territory ranging from an average 2.7 people per km^2 in interior provinces like Moxico and Cuando Cubango to the northeast to more than 100 people per km^2 in Luanda, Benguela and Cabinda along the country's shoreline (Figure 12).

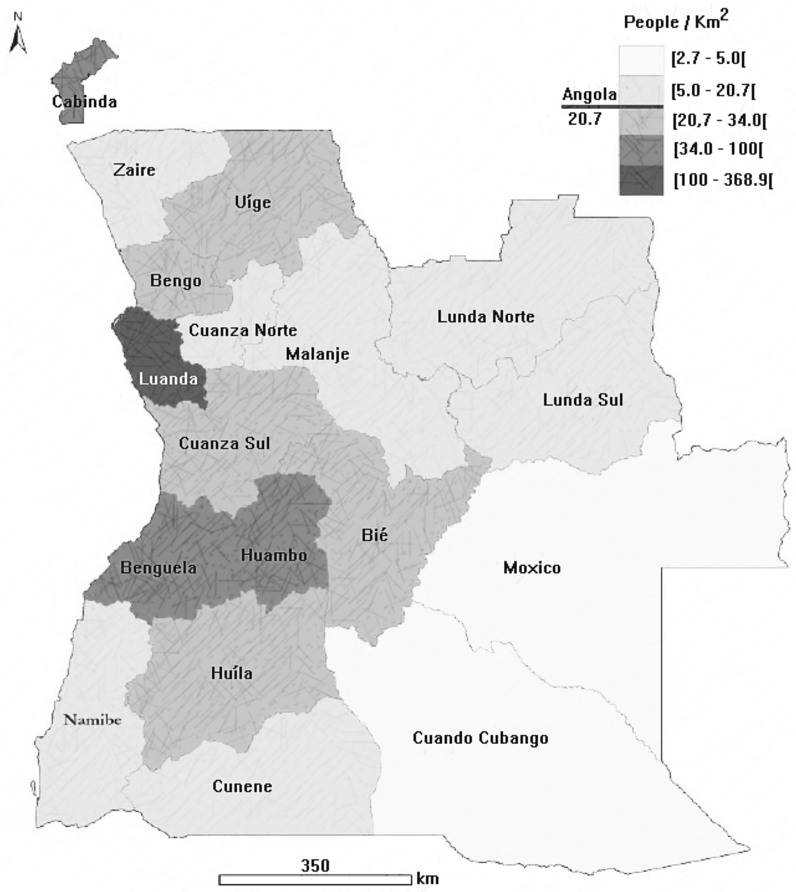


Figure 12 - Population density by province – Angola 2014

2.1.2 Access to electricity

The census also showed that electricity was accessible to less than one third of the population and that great asymmetries existed between urban and rural areas where only 2% of people had access to electricity of any kind (Table 17).

Table 17- Households with access to electricity by area of residence

Country and area of residence	number of households	% of population
Angola	1,770,729	31.9%
Urban	1,722,556	50.9%
Rural	48,173	2.20%

Adapted from (INE - Instituto Nacional de Estatística, 2016)

Moreover, only two provinces (Luanda and Namibe) had access rates over 50% of their respective populations.

The projected population for 2030 is of 41,777,194 people (INE - Instituto Nacional de Estatística, 2018) meaning potential demand for electricity services should also continue to grow and the energy sector should prepare to meet this demand in urban and rural areas. A national study (INE - Instituto Nacional de Estatística, 2018) showed that the population of rural areas that have almost no access to electricity is expected to grow from 9.6 million people in 2014 to over 17 million by 2030.

2.2 Angolan energy sector

Angola has vast oil reserves, being one of the largest producers in African continent (MINEA, 2018a). The country also has considerable reserves of natural gas and oil production has led to rapid economic growth. However, the existing electric sector infra structures did not accompany the demand growth. The suppressed demand has been offset by the generation (mostly private) from diesel, for the minority that can pay the price, which is not the case for the large majority outside urban areas.

2.2.1 Electric system

The Angolan national electrical grid is not unified. It is comprised of five main independent systems: Northern, Central, Southern and Eastern, plus Cabinda (MINEA, 2018a). Cabinda is considered an isolated mini-grid system. The largest electrical system in Angola is the Northern system in terms of infrastructures and demand. It serves the provinces of Luanda, Cuanza Norte, Malange, Uíge, Bengo and Zaire with Luanda having the largest demand (around 80% of global demand is from the province of Luanda which is supplied by the northern system). The Northern, Central and South systems are planned to be connected by the end of 2022 (Figure 13).

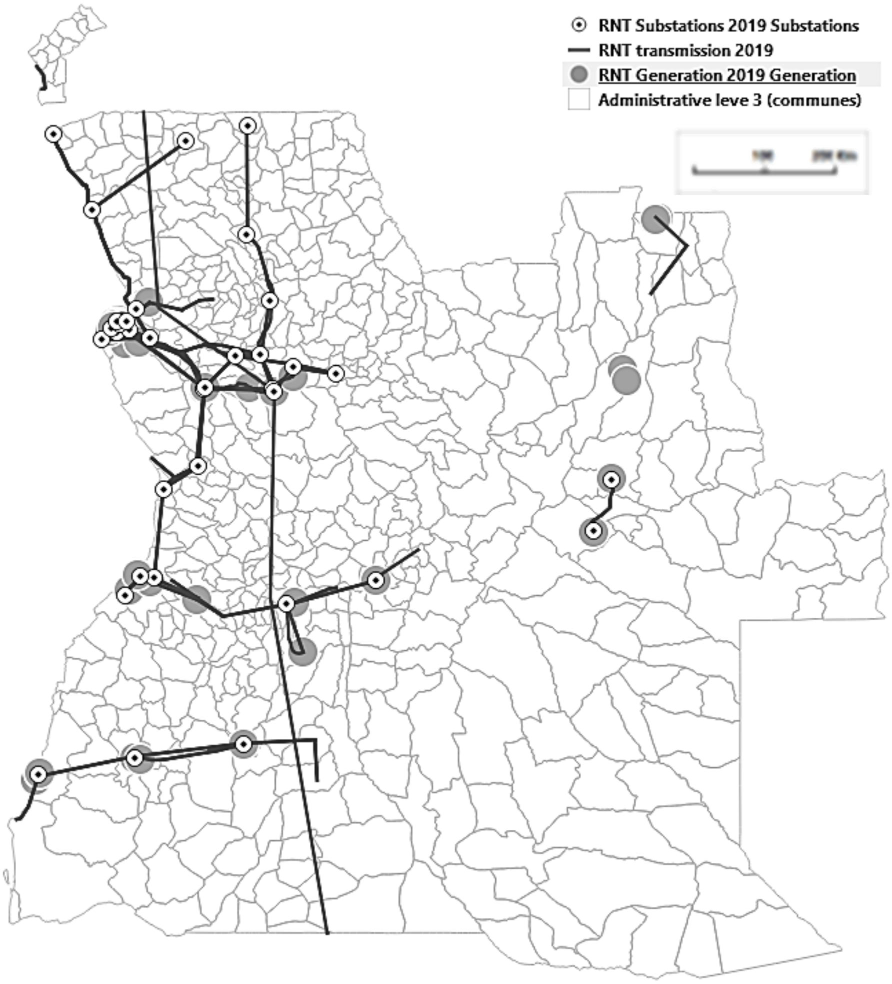


Figure 13 **–** Expected evolution of the national grid of Angola by 2022

Produced by the author in QGOS 3.10 using data supplied by Rede Nacional de Transporte (RNT)

Currently only the northern and central systems are connected and a large portion of communes will remain unconnected from the grid in 2022.

2.2.2 Generation

Angola's generation potential is managed and developed by a governmental company called Empresa Pública de Produção de Electricidade (PRODEL). The country's hydropower potential is vast and hydro energy is the country's focus in terms of generation. Angola started to explore hydropower during the 1950s and 1960s when several hydropower plants were built in the country. Angola's hydropower scheme is under rehabilitation and expansion. Currently the installed hydropower capacity reaches 4 GW, which represents roughly 70% of the total installed power capacity in Angola (MINEA, 2018a).

Unfortunately, despite the low generation costs and the benefits for future generations of the country, these projects are concentrated in the rivers with greater flows and in areas of higher slope, resulting in a high level of territorial concentration of power generation, in particular in the Kwanza River Basin and the Northern System (Figure 14). This means a significant share of this large installed hydropower capacity is unused. At this stage, there is a need for planning the expansion of the grid and investing in energy transport and distribution.

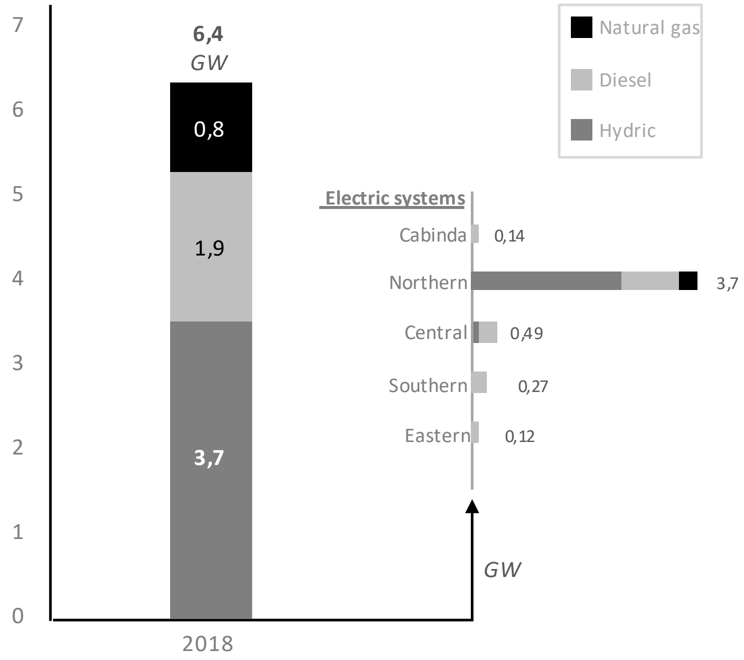


Figure 14 - Installed capacity for the electric system in Angola (2018)

Adapted from (MINEA, 2018a)

Exception made for the northern system in which hydro power generation is prevalent, diesel generation is present in a significant way. The northern system has a global capacity four times larger than the remaining systems combined.

2.2.3 Electricity Transport and Distribution

Rede Nacional de Transporte (RNT) is the public company responsible for energy transport and the development of the national strategy for grid extension. Distribution is managed by another company, called Empresa Nacional de Distribuição de Electricidade (ENDE), also in the public sphere, and the sector regulator is the Instituto Regulador dos Serviços de Electricidade e de Água (IRSEA).

In Angola, at the end of 2018, there were 4,012 kilometers of transport lines and 34 substations, across the four systems (Table 18).

Table 18 - Transmission line length by voltage

Voltage [kV]	400	220	150	132	110
Transmission line length [km]	1324	2194	192	57	245

Adapted from (MINEA, 2018b)

Concerning the public investment plan for the electric sector (MINEA, 2018b), it becomes apparent that the current national strategy for the sector involves a combination of public investment, private sector participation and regulation of tariffs. Due to the drop in oil prices, there is a need for contention that translates into a still considerable $6.974 million USD of public spending in the electric sector between 2018 and 2022. This strategy is, according to official documents (MINEA, 2018a), in line with the subscribed goals (ECCAS - Economic Community of Central African States, 2014). Its state of completion and feasibility, given current constraints, will be taken into consideration when establishing the baseline for this study. The current economic situation and the COVID-19 pandemic demand that every decision regarding investments are done in the most efficient way.

2.3 Investment prospects and budget

The SE4all investment prospectus for Angola (MINEA, 2018a) includes several projects linked with access to energy. Investment prospects (in global numbers) are not available to the public, but RNT kindly supplied an additional internal report on the national strategy for grid development (MINEA, 2018b). According to this report the global value for transport and distribution, between 2016 and 2025, was planned to be close to 11,000 million USD (Table 19). From these, 1,610 million USD are targeted at rural electrification, an average 161 million, per year.

Table 19 – Global Investment estimated for Energy 2016 - 2025 strategy

Typology	**investment** (USD millions)
Investment in Generation	16,944.0
Investment in distribution for urban areas	4,957.4
Investment in Rural Electrification	1,610.0
Extra high voltage power lines (EHV)	1,133.8
Interconnections (Congo Democratic Republic)	62.8
Transmission Substation EHV / HV	1,339.7
Power lines and substations to support rural electrification	1,225.4
Global investment (estimate)	**27,273.1**

Adapted from MINEA (2018a)

These ambitious goals have been delayed due to the recession that set in at the end of 2018. Angola is also under partial FMI intervention, which may imply reducing planned investment. With these macroeconomic conditions and the COVID-19 pandemic, expenditure has been, on average, below expectations. Official documents show that the budget execution has been around 50% of what was planned in 2019 (MINISTÉRIO DAS FINANÇAS, 2019).

The way in which all these elements and inputs were combined in the methodological approach that was employed in this study is detailed in Chapter **Error! Reference source not found.**.

Page intentionally left blank.

CHAPTER 3

Data collection and methodology

The challenge of developing a decision-making support model for energy planning in Angola was faced taking into account the literature on the topic. After careful consideration on the dimension and goals of the study, an approach was developed. This approach combines GIS tools to collect data inputs with heuristic methods, based on MST algorithms, to produce layouts of grid and off grid solutions. These layouts, while not technically accurate, are still useful for decision-making. This study can be classified, according to the literature, as an intermediate study (Ciller & Lumbreras, 2020), that uses what can be categorized as a comprehensive modeling approach (Corigliano et al., 2020).

The methodological approach for dealing with problem of access to electricity in rural communities of Angola has two main stages which are given in Figure 15:

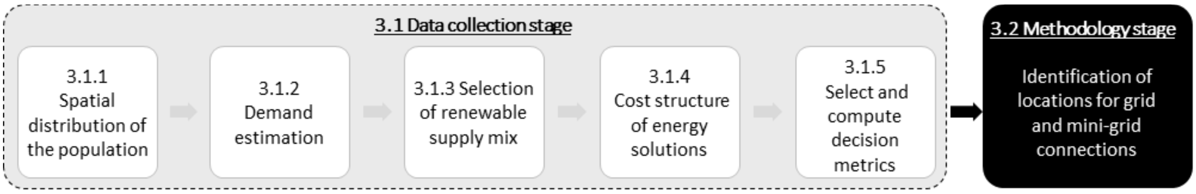


Figure 15 – Two stage methodology designed for this thesis.

i) The *data collection stage* encompasses: 1. finding GIS data with the spatial distribution of population in demand vertices, 2. determining the demand for each of these vertices, 3. definition of an adequate supply mix, according to data on local resources, 4. defining the cost structure for each technology in the supply mix, and 5. selection of suitable decision-making metrics;

ii) The *methodology* in which the best electrification option for each cluster, site or vertex is selected.

3.1 Data collection stage

This research employs data inputs commonly found in studies dealing with the rural electrification problem at a national level. For the case of Angola most of this data was not directly available and needed to be inferred though indirect methods. The way in which the data collection process was carried out is detailed in subchapters 3.1.1 through 3.1.5.

3.1.1 Data collection – spatial distribution of the population

Considering the purpose of this study, it was essential to map the population in different years, especially in 2030. Taking into account that Angola has subscribed a target coverage rate of 54% for the year 2030 (ECCAS - Economic Community of Central African States, 2014), this mapping should be forward looking. However, at the time of writing, the available spatial data was considered of insufficient resolution, namely regarding rural areas in Angola. The best data available was at a commune level (level 3 administrative division) and some communes are quite large. This fact coupled with the desire of maintaining a level of autonomy for the whole process led to the choice of a smart interpolation technic (Hay et al., 2005). Smart interpolation uses ancillary data such as the location of human-built structures and their attributes (size, urban or rural) for the distribution of the population in a non-uniform and non-random way. This approach requires finding suitable ancillary data. The data collecting process includes five sources of data (layers) from different sources. Following what has been established in literature for population geolocation, this study uses ancillary data that is a proxy for human settlements (night time lights) (Hay et al., 2005) and data that is directly representative of human settlements (OSM human built structures – points and polygons) (Tobler et al., 1997). For population density, census and administrative division information is used (Table 20).

Table 20 – Ancillary data for the mapping process

Data	Updated as of	Type	Format	Price	Resolution	Source	Included in this research
OpenStreetMap human built structures – *polygons*	December 2020	-	Vector	Free	-	OSM	Yes.
OpenStreetMap human built structures – *points*	December 2020	-	Vector	Free	-	OSM	Yes.
Administrative divisions – *levels 0, 1, 2 and 3*	version 3.4, April 2018	-	Vector	Free	-	GADM - Global administrative areas	Yes.
Angola Night lights 2016	2016	Land cover	Raster	Free	500 m	NASA Earth observatory	Yes.
Census data	2014	Census	-	Free	-	National Statistics Institute of Angola	Yes.

Produced by the author

3.1.1.1 Data quality considerations

In a broad sense, issues of scale, currency, semantics, model related issues and validation challenges are relevant for any population mapping approach (Leyk et al., 2019). This research addresses each one of these potential problems in the most efficient way possible:

i) Issues of Scale – as shown in Table 20 data of different types and resolution is used. To ensure the adequacy of combining these data layers, literature suggests that the distribution of population maps should rely on automated procedures, whenever possible, and rely on human skills when necessary. The general procedure that is applied was validated by an external GIS consultant.

ii) Issues of Currency and temporal agreement - for Census data, both the completeness of coverage and margin of error are crucial. In the Angolan case, the reported coverage rate was of 94.4% with an estimated error of 1% (INE - Instituto Nacional de Estatística, 2016). Administrative divisions in Angola have been in flux since the end of the civil war. Table 21 describes the administrative division revision file that is used. The administrative division vector files (levels 1, 2 and 3) that are used (the only ones available) have 18 provinces, 163 municipalities and 539 communes. To match the spatial administrative information with census population data it was necessary to identify which communes were merged and for these communes to sum their respective census population, in order to maintain agreement between census and spatial information for 2014. The majority of change had to do with the incorporation of urban communes in the zone of Luanda. It had little impact on research.

Table 21 – Changes in administrative divisions in Angola

	Provinces	**Municipalities**	**Communes**
Administrative division from the end of the civil war (2002) up to 2011	10	193	475
Administrative division at time of final Census results (approved in 2011)	18	162	559
Administrative division file for this study	18	163	539
Current administrative division (2021)	18	164	518

Adapted from http://www.mat.gov.ao/

iii) Issues of semantics are minimized using the place of residence or Nightlights (NTL) (de jure) concept. This is a commonly used approach in literature.

iv) Model related issues - the smart interpolation (Hay et al., 2005) approach that is used is a relatively common technic that is referenced in the literature. The relation between night-lights and settlement data with population (including POI data) has been established in the literature. This way the issues commonly associated with modeling were minimized.

v) Validation issues - According to literature it is at the discretion of the analyst to decide if the data product (gridded map or ancillary data) is adequate for the purposes of any given study. There is no systematic procedure to decide this. In spite of this, in this study, an internal validation (Grippa et al., 2019) is employed.

3.1.1.2 Mapping population distribution for rural areas of Angola

Some work was carried regarding each of the ancillary layers of data used in the mapping procedure:

i) *Human built structures layers:* Open Street maps (OSM) provide free online data on the location of settlements all over the world. For Angola, settlement layers (point and polygon layers) are used. A point layer containing 3,875 human built structures is extracted and imported into Quantum GIS (QGIS). Another layer, extracted from OSM, with polygons that represent areas smaller than 4 km^2 are treated as points (using polygon centroids[14]). This produces a vector layer with 99 points. These points are joined[15] with the point layer of 3,875 human built structure settlements totalizing 3,974 settlements locations.

ii) The *Night-lights* layer is used as proxy for urban extents: the night-lights original raster file concerns an area larger than the Angolan territory. Some of the areas with light are in the ocean (namely on the coast of Cabinda) and were removed[16] since they are irrelevant to our study. A clipped raster file with night-lights areas on land was obtained. The raster layer was converted to a polygonal layer (Figure 16). These polygons were subsequently used as proxies for urban extents. This produced a polygon layer with 37 urban extents. These areas represent large urban areas, close to major cities, with high population density.

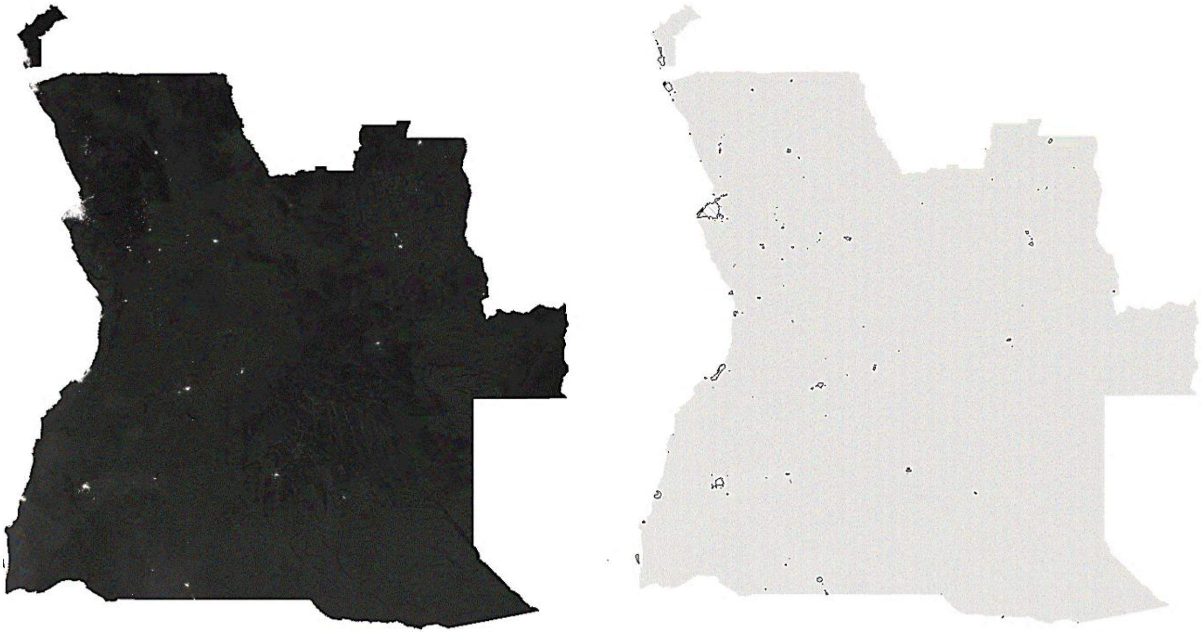


Figure 16 – Angola night-lights – raster to vector

Produced by the author on QGIS 3.10

---

[14] *This polygon layer was loaded in QGIS 3.10, then centroids were added using Vector > Geometry tools > Centroids.*

[15] *The two layers were joined using the command: Vector > Geo-processing tools > Union. This produced a new vector – point layer with a total of 3,974 settlement locations.*

[16] *i. Raster > Extraction > Clip Raster by mask layer; ii. Next the Saga Add On was used > Raster calculus > Raster calculator to isolate the areas with light as some lighted areas were in the ocean (oilrigs). These areas will be the urban extents. The following formulas in the Raster calculator dialog: ifelse(a<70, 0, a) - to remove all the non-lighted areas from the clipped raster file. "Clipped_raster_just_lights" was produced. iii. To get the same value in all pixels of the lighted areas of new the raster file ifelse(a>=70, 255, a). Finally, the clipped raster file was polygonised. Raster > Conversion > polygonise.*

iii) Distribution of the population through the urban extent layer: The settlement point layer and the urban extent layer were joined. The population of these 37 large urban areas was added from Census data. This totaled 14,257,874 people in 37 urban areas (INE - Instituto Nacional de Estatística, 2016). These urban areas will always be a target for grid connection in the remaining years. They are not the focus of this study, as their connection plan is already largely drawn by central government. They represent approximately 54% of the population that is in line with subscribed SE4all (MINEA, 2018b) and (ECCAS - Economic Community of Central African States, 2014) goals.

All points falling inside these urban extents were removed (266 settlements). In total the final point layer only included points outside urban areas, representing 3,708 rural settlements.

3.1.1.3 Population mapping outside large urban areas of Angola in 2014

According to the 2014 census, beyond urban extents, 11,531,151 people lived in rural areas. These people are distributed across the 539 communes of Angola. Among these 101 communes, which according to the 2014 census represent 2,157,121 people and has no human-built structure marked in the OSM database. A layer was created for each one of these communes, rather than ignoring this limitation or using, for example, commune centroids. For each of these communes, settlement locations were manually added, analyzing Google Earth[17] images. After this procedure, 849 settlements were added. The final point layer had 4,557 points, the previous 3,708 plus the 849 that were added, representing human built structures. The assumption was that these points account for the large majority of the 11,531,151 people in rural areas.

In order to address the issue of population distribution for these 4,557 points, OSM settlement classifications were used (city, town, village and hamlet). In the OSM classification a human built structure is classified as city, town, village or hamlet according to the approximate number of people it contains (Table 22)[18]. This process was audited by an external specialist[19], in line with (Herfort et al., 2019). The population distribution procedure looks to place more population in where the settlements are larger according to OSM classification.

The number of settlements inside each commune was counted. Given that the number of rural settlements as well as the total rural population, per commune, was known it was only necessary to distribute that number of people across the settlements inside each commune, according to the OSM

---

[17] *Google images are updated, yearly, for several provinces of Angola. This might be an adequate way of keeping input settlement data current.*
This can be accessed via google earth pro.

[18] *Available at openstreetmap.org.*

[19] *PhD in Geophysical Sciences and Geoinformation.*

classification.

Table 22 – OSM –place (human-built structure) classifications

Code	class	Description	OSM tag
1001	City	As defined by national/state/provincial government. Often over 100,000 people.	Place=city
1002	Town	As defined by national/state/provincial government. Between 10,000 and 100,000 people.	Place=town
1003	Village	As defined by national/state/provincial government. Generally, smaller than a town, below 10,000 people.	Place=village
1004	Hamlet	As defined by national/state/provincial government. Generally smaller than a village, just a few houses.	Place=hamlet

Adapted from openstreetmap.org by the author

In Figure 17 the general procedure population distribution is shown.

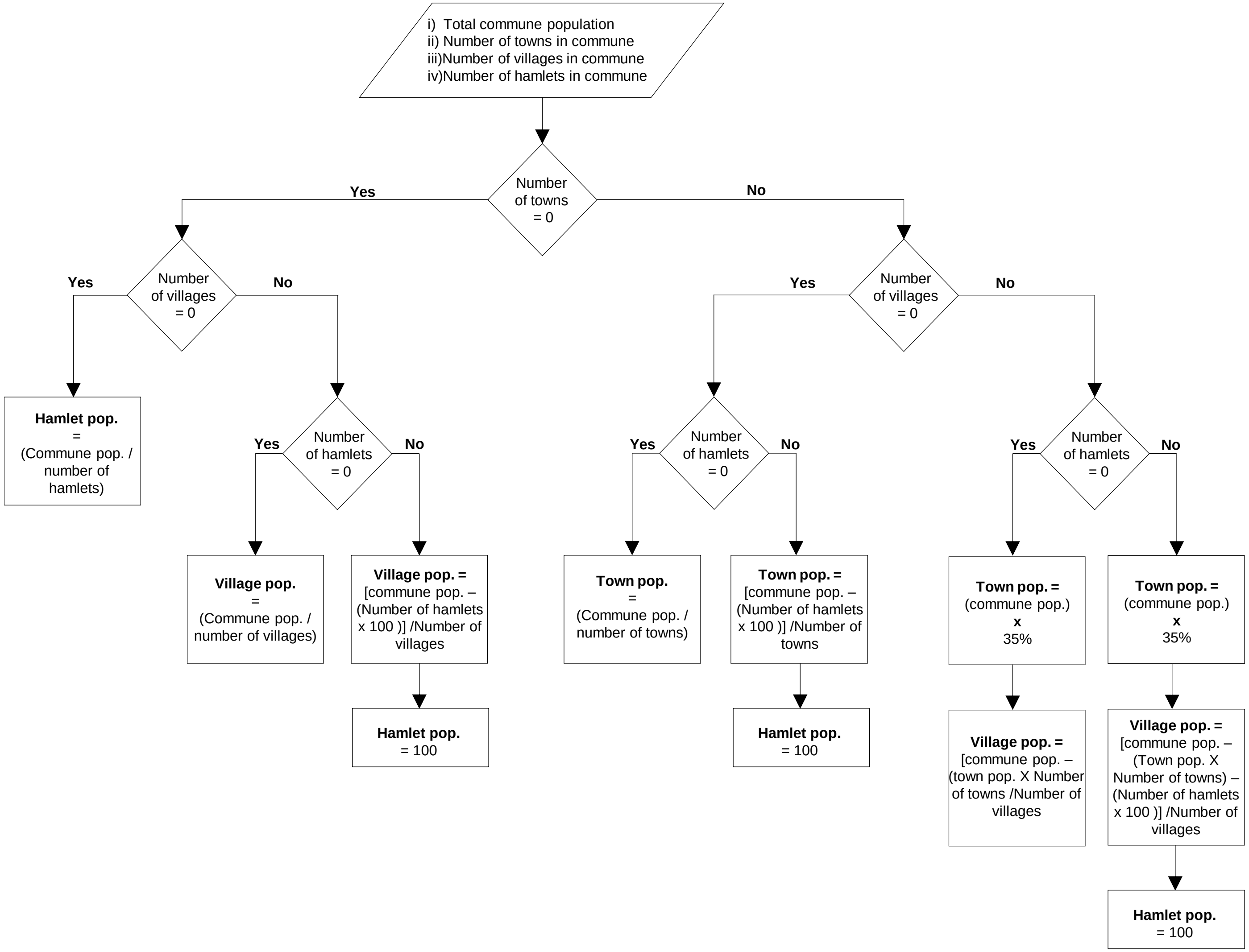


Figure 17 – Population distribution algorithm for each commune

Produced by the author

This process looks for the relative population sizes of each one of the 4,557 settlements. The 11,531,151 people were distributed, by each settlement, using the total population of the commune (according to census), and OSM classifications (a maximum of 100,000, 10,000 and 100 people by town, village and hamlet, respectively). None of the settlements were classified as "city", which was somewhat expected given that the data concerns rural communities. In some cases, a commune might have 10,000 people and no towns or hamlets. In these cases, exceptions are introduced: if the number of towns is zero the lion's share of the population is distributed by the villages, ignoring OSM limits. If towns and villages are present, the towns receive 35% of the commune's population[20]. If both the number of towns and villages is zero the population is distributed by the hamlets, ignoring OSM rules. This way the relative population sizes inside the commune are preserved. This procedure was applied to each commune using a spreadsheet. For each commune the output was saved as a comma-separated-value (CSV) file and imported into QGIS[21].

For the case of the commune of Ringoma, for example, the census had a count of 8091 people. From OSM data, 15 villages and 3 hamlets were identified (Table 23).

Table 23 – Example of the population distribution procedure for the commune of Ringoma – population data from 2014

	Number of settlements in commune				**Total pop.**	**Avg. Population by type of settlement**			
	city	*town*	*village*	*hamlet*		*city*	*town*	*village*	*hamlet*
Commune of Ringoma	0	0	15	3	8,091	0	0	519	100

Produced by the author

According to the procedure, to each of the three hamlets was attributed 100 people and to each of the 15 villages 519.4 people (rounded to 519 for practical reasons). In Figure 18 it is possible to assess the advantages of this approach over the use of commune centroids. For example, most villages are far from the center of the commune, and are of similar size. Using a centroid would mean losing all this detail.

---

[20] *This was the national average concentration of people in larger population aggregates (like towns and cities) in rural municipalities of Angola (INE - Instituto Nacional de Estatística, 2016).*

[21] *Using the table join feature in QGIS (https://www.qgistutorials.com/en/docs/3/performing_table_joins.html)*

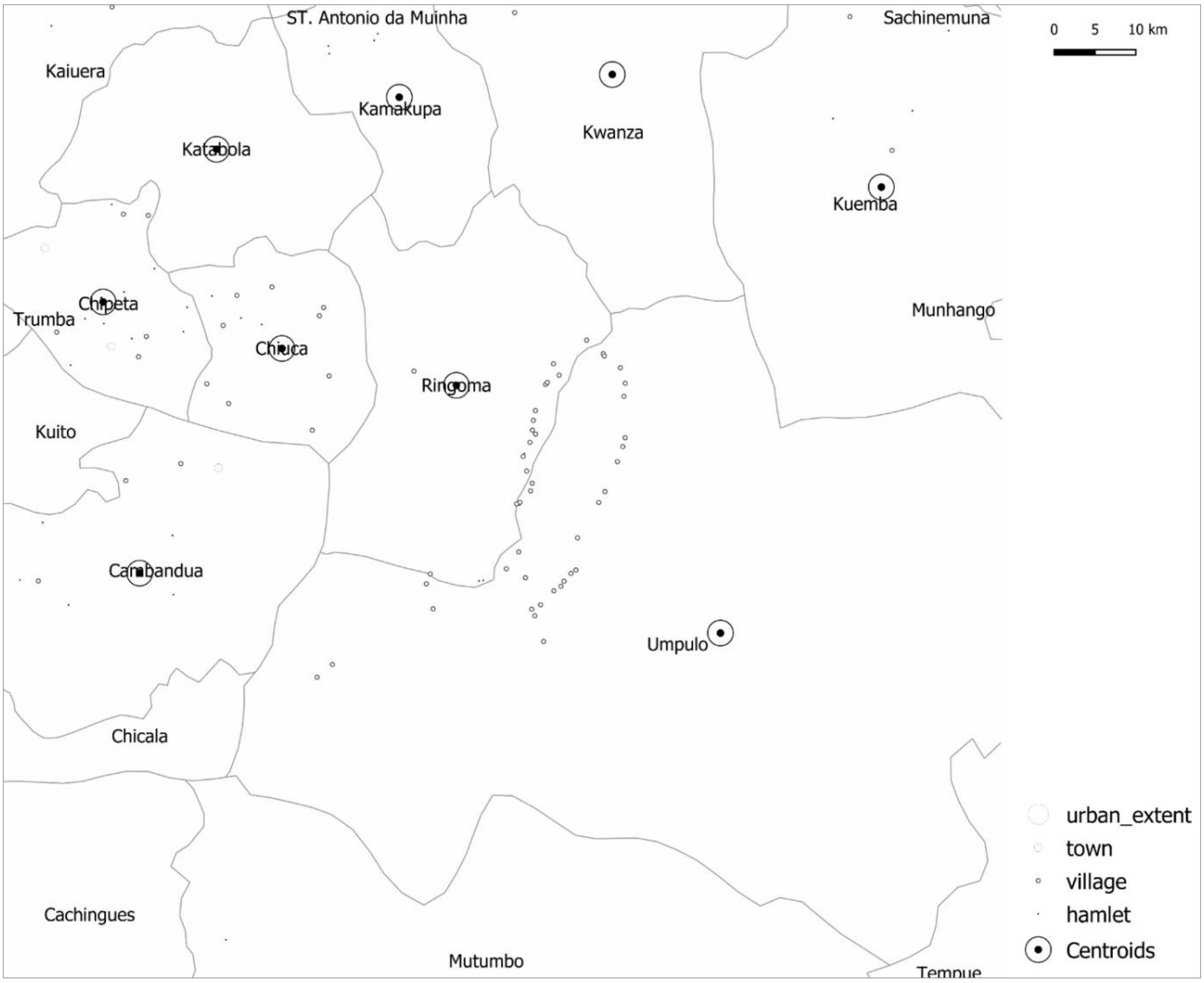


Figure 18 – Angolan population distribution map for 2014

Produced by the author in QGIS 3.10

This procedure was carried out, for all communes, in the same way.

3.1.1.4 *Internal validation*

Limitations on existing data are a reality in SSA. Validation of the results is often impossible due to the lack of resolution of the input data that is available (Leyk et al., 2019). This can influence validation procedures. To validate the population distribution procedure a method identical to the one proposed by Grippa et al. (2019) was used. These methods consisted of three main stages:

Stage one – A sample of 3 coarser administrative unitswas created. Two with similar area and a smaller one to measure the impact of area size in the results. These areas encompass 18 communes and over 350,000 people. Each coarser administrative aggregates neighboring communes and contains between 65,000 and 162,000 people.

Stage two – The total rural population (from census) for the coarser areas was distributed according to the procedure described in Subchapter 3.1.1.3 and consequently the population for the communes contained in the coarser areas was determined.

Stage three –The predicted rural population numbers for the communes belonging to the coarser areas were compared to the real reference data from the 2014 census. The metric used to test the performance of the smart interpolation procedure was the normalized metric root mean square error (%RMSE), given in Equation (6). It can be interpreted as the average error (distance) between a predicted value and the observed value, from census (Grippa et al., 2019). This measure is very sensitive to moderate errors in a small number of large values and much less sensitive to serious errors in small values (Batista e Silva et al., 2013), which is adequate for this study.

$$\%RMSE = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(pred_i - ref_i)^2}}{\sum_{i=1}^{n} ref_i} * 100\% \tag{6}$$

The results for the validation process are given in Table 24, Table 25 and Table 26.

Table 24 – Validation process - Sample area 1

Commune	**town**	**village**	**hamlet**	**zone 1 - Pop remaining (reference - census)**	**Pop. by com. (predicted)**	**%RMSE**
Bembe	0	7	1	135,117	16,225	46.6%
Kiende	0	3	1		7,011	
Kindege	0	9	4		21,132	
Kinvuenga	0	3	6		7,511	
Lemboa	0	4	3		9,514	
Luvo	0	0	8		800	
Mabaia	0	0	25		2,500	
Madinba	0	2	3		4,907	
Nkuso	0	0	7		700	
Quiximba	0	4	3		9,514	
Songo	1	3	11		55,302	
				Total	**135,117**	

Produced by the author

Table 25– Validation process - Sample area 2

Commune	**town**	**village**	**hamlet**	**zone 2 - Pop remaining (reference - census)**	**Pop. By com. (predicted)**	**%RMSE**
Mungo	1	12	13	162,433	75,451	41.8%
Lunge	1	15	14		86,982	
Chiumbu	-	-	-		0	
				Total	**162,433**	

Produced by the author

Table 26– Validation process - Sample area 3

Commune	**Town**	**village**	**hamlet**	**zone 3 - Population (reference value - census)**	**Pop. By com. (predicted value)**	**%RMSE**
Cambandua	1	3	9	66,574	12,895	18.7%
Chipeta	2	6	9		24,890	
Chiuca	0	9	3		12,985	
Kaiuera	0	11	3		15,804	
				Total	**66,574**	

Produced by the author

For the three coarser areas, the results compared favorably with studies at a national level, like Grippa et al. (2019) which attained a global %RMSE of 44% across several individual countries. For the case of these sample areas error decreases as sample area size decreases (sample area 1 has over one hectare whereas sample area 3 has one third of that). Area seems to be more relevant than population density in terms of %RMSE. This is in line with Palacios-López et al. (2021) that found that RMSE is likely to report higher values influenced solely by a larger area and sample size.

Given that in this research, the process of population disaggregation was carried out in smaller areas (communes), it is reasonable to expect smaller errors for the population disaggregation procedure. This provides enough confidence to use this approach for producing approximations of population distribution numbers, at a resolution that was previously unavailable, for Angola. To attain better quality of spatial population distribution would require access to higher resolution satellite imagery that is outside the scope and budget of this research.

3.1.1.5 Estimating the population from 2022 to 2030

Population numbers and growth rates, between 2022 and 2030, were already available for rural areas, from a previous study (INE - Instituto Nacional de Estatística, 2018), but only at a municipality level (level 2 administrative division). This meant these growth rates had to be generalized for all communes in each municipality.

There is an underlying assumption of reduced migrations from rural areas to urban areas (INE - Instituto Nacional de Estatística, 2018). For the period of 2022 – 2030 numbers point to a growth of 53% of urban populations against a 77% growth in rural areas. Additionally, the COVID-19 pandemic has accentuated this trend. Movement between provinces has practically halted since 2020, due to lockdowns.

To map the spatial distribution of the rural population, from 2022 to 2030, the same procedure already described in Subchapter 3.1.1.3 was applied. The total population estimates for the period between 2022 and 2030 is given in Table 27.

Table 27 – Population estimates for Angola (2022 – 2030)

Year	Population estimate
2022	33,086,278
2023	34,094,077
2024	35,121,734
2025	36,170,961
2026	37,243,484
2027	38,340,130
2028	39,461,732
2029	40,608,969
2030	41,777,194

Adapted from INE - Instituto Nacional de Estatística (2018)

These estimates are made under the assumption the current settlements will grow with little fragmentation. The reasons why this is reasonable are varied:

- The official data supports this (INE - Instituto Nacional de Estatística, 2018);
- Angola has had low mobility of its population due to lack of roads and functioning public

transportation. Most people do not have cars able to undertake long journeys outside urban areas. This has worsened immensely after 2020, due to the lockdowns that are still ongoing in 2021.

- The timeframe of analysis is not very long (under 10 years);

  - In the case of investment in mini-grids, unlike grid extension, there is periodic reinvestment, meaning it is possible to course correct if a new settlement appears and the current ones do not expand as expected. While this option was not explored in this study, it is viable since a multi-period approach was taken.

The estimated population numbers serve as reference for demand and supply estimation.

3.1.2 Data collection – Demand estimation for rural communities of Angola (2022 – 2030)

After knowing where people are likely to be it is possible to produce an estimate of their demand for electricity. Electricity is only useful if it allows the energy services, which people need, to run appropriately. Different energy services (such as lighting, television, air circulation, refrigeration, space heating, etc.) require different levels of electricity supply in terms of quantity, supply duration, quality, and affordability. In addition, as access grows over time, demand patterns also tend to change.

For Angola, an econometric approach is not feasible due to data scarcity and undersupplied demand (Bhattacharyya et al., 2010). Thus, a bottom-up approach is used allowing the determination of yearly demand level, per location, at each time step of the research's period of analysis (from 2022 to 2030). This approach incorporates ideas from several studies (Deichmann et al., 2011; Mentis et al., 2017; Moner-Girona et al., 2019; Parshall et al., 2009) while not being identical to any of them. The main factors that will drive this approach are:

i) Rural population distribution by settlement (from Subchapter 3.1.1);
ii) Relative level of poverty by location (INE - Instituto Nacional de Estatística, 2019);
iii) A set of tiers or patterns of demand (IRENA, 2018);
iv) Data on the existence of schools and health units[22] near each settlement.

According to the results from the previous subchapter, an estimated 17,092,796 people will inhabit rural areas, in 2030, across 4,557 settlements. For the purposes of demand estimation, these settlements have to be categorized according to demand tiers. The demand in each settlement included four possible components:

- Household demand;

---

[22] *Extracted from https://www.who.int/.*

- Productive demand – uses of electricity for small business and agricultural purposes (computed as a function of domestic demand). Large scale projects and industrial needs were considered outside the scope of this research;
- For health demand a yearly fixed value from literature (United Nations Development Program, 2015) was used for each clinic present in a location;
- For education demand a yearly fixed value is equally used for each school present in a location. The source for this value was also the (United Nations Development Programme, 2015a);

According some authors, like Morrissey (2019), some of these features has been sometimes lacking in previous studies.

For household demand tiers, the source was the Global Tracking Framework from the United Nations (United Nations Development Programme, 2015a). These tiers have underlying assumptions regarding the types of energy services and appliances people will be able to use. These assumptions are expressed in Table 28.

Table 28 – Tiers of domestic electricity demand for the case of Angola

Tiers	**Tier 1**	**Tier 2**	**Tier 3**	**Tier 4**	**Tier 5**
Lighting	Task light	Multipoint general lighting	-	-	-
Communication and entertainment	phone charging, radio	Television, computer, printer	-	-	-
Cooling and heating	-	Fan	Air cooler	-	Air conditioner, space heater
Refrigeration	-	-	Refrigerator, freezer	-	-
Mechanical loads	-	-	Food processor, water pump	Washing machine	Vacuum cleaner
Product heating	-	-	-	Iron, hair dryer	Water heater
Cooking	-	-	Rice cooker	Toaster, microwave	Electric cooker

Adapted from (United Nations Development Programme, 2015a)

In a study by INE - Instituto Nacional de Estatística (2019) a multidimensional poverty index (MPI) is produced, for each municipality of Angola, using the approach proposed by Alkire and Foster

(2011) (see appendix A). To each settlement was assigned the MPI of the corresponding municipality.

To classify each settlement according to aforementioned tiers the information in Table 28 was used.

Table 29 – Domestic electricity tiers, by MPI quintile (Q).

Domestic demand tiers by quintile (Q1 – lowest level of poverty to Q5 – highest level of poverty)	**Q1** Tier 4	**Q2-Q3** Tier 3	**Q4-Q5** Tier 2
Peak capacity (Watt)	800	200	70
Minimum average number of hours per day	16	12	12
Demand (Wh) per household per day	3,400	1,000	275
Demand (kWh) per household per year	1,241	365	100

Source: Tiers adapted from (United Nations Development Programme, 2015a) and poverty index from INE - Instituto Nacional de Estatística (2019).

These demand tiers are set per household, not by the number of people. To compute the number of households in each location the population of each settlement was divided by the average household number in the commune it belongs to (INE - Instituto Nacional de Estatística, 2016). To attain the yearly household demand for each settlement the number of households was multiplied by the yearly demand (kWh) per household, according to its corresponding demand tier.

For health demand, a list of health facilities government and non-government sources from 50 countries in SSA was used (available from the World Health Organization[23]). It provides a comprehensive spatial inventory of 98,745 public health facilities. A georeferenced file was extracted, containing 1,575 health facilities in Angola. From these, large public hospitals and private facilities located in urban areas were excluded, since they are facilities are outside the scope of this study. The remaining smaller public health facilities, that serve rural areas, were taken into account to estimate demand. Since settlements and health facilities are represented by points that do not have identical geolocation, there was a need to perform some clustering. Using QGIS[24] the settlement closest to each of the health centers was determined, forming a demand cluster. Equation (7) is used to determine an Euclidean distance between health facilities and the settlements. The clustering process assigns each health center with the closest settlement, under a 1 km radius. Health centers outside this radius were discarded. In the end, 617 health facilities remained, each one associated with one settlement.

---

[23] https://www.who.int/malaria/areas/surveillance/public-sector-health-facilities-ss-africa/en/.

[24] *In QGIS 3.10 select Vector>Analysis tool > Distance matrix (linear Euclidean distance, k=1).*

$$D(x,y) = \sqrt{\sum_{i=1}^{m}(x_i - y_i)^2} \qquad (7)$$

Where,

x_i – represents a point belonging to the health facilities list.

y_i – the nearest settlement, from the settlements list.

D(x,y) – distance (in meters) between a health facility and its nearest settlement.

Regarding demand for education, it was not possible to attain reliable data regarding the location of schools. Thus, it is assumed that near each one of the 617 health centers, primary and secondary schools are, or can be, present.

For schools and health facilities, the yearly demand reference values suggested by Blechinger et al. (2019) ae used. Moreover, the yearly productive demand is computed as a function of domestic use under a daily 4-hour restriction (Table 30).

Table 30 – Productive, health and education demand

Demand	**Q1**	**Q2-Q3**	**Q4-Q5**	**Demand kWh / year**	
Peak capacity	200	200	200	Health facility	2,400
kWh per household per year	310	122	33	Primary school	1,200
kWh for production (kW) per	52	20	5.5	Secondary	2,400

Source: Produced by the author

In rural areas, the introduction of new energy technologies impacts the social-economic structures of communities (Riva et al., 2018). Some models ignore the impact of access to electricity on demand growth, while others look at historic appliance use data. For Angola, this data is not available, with sufficient resolution. Instead, a constant rate of demand growth was used. The Ministry of Energy of Angola estimated a constant demand growth of 2.61%, for the period from 2022 to 2025 (MINEA, 2018a). This rate was extrapolated for the full period of analysis (from 2022 to 2030). Finally, for each rural settlement, domestic, productive, health and education demand were estimated using the procedure detailed in Figure 19.

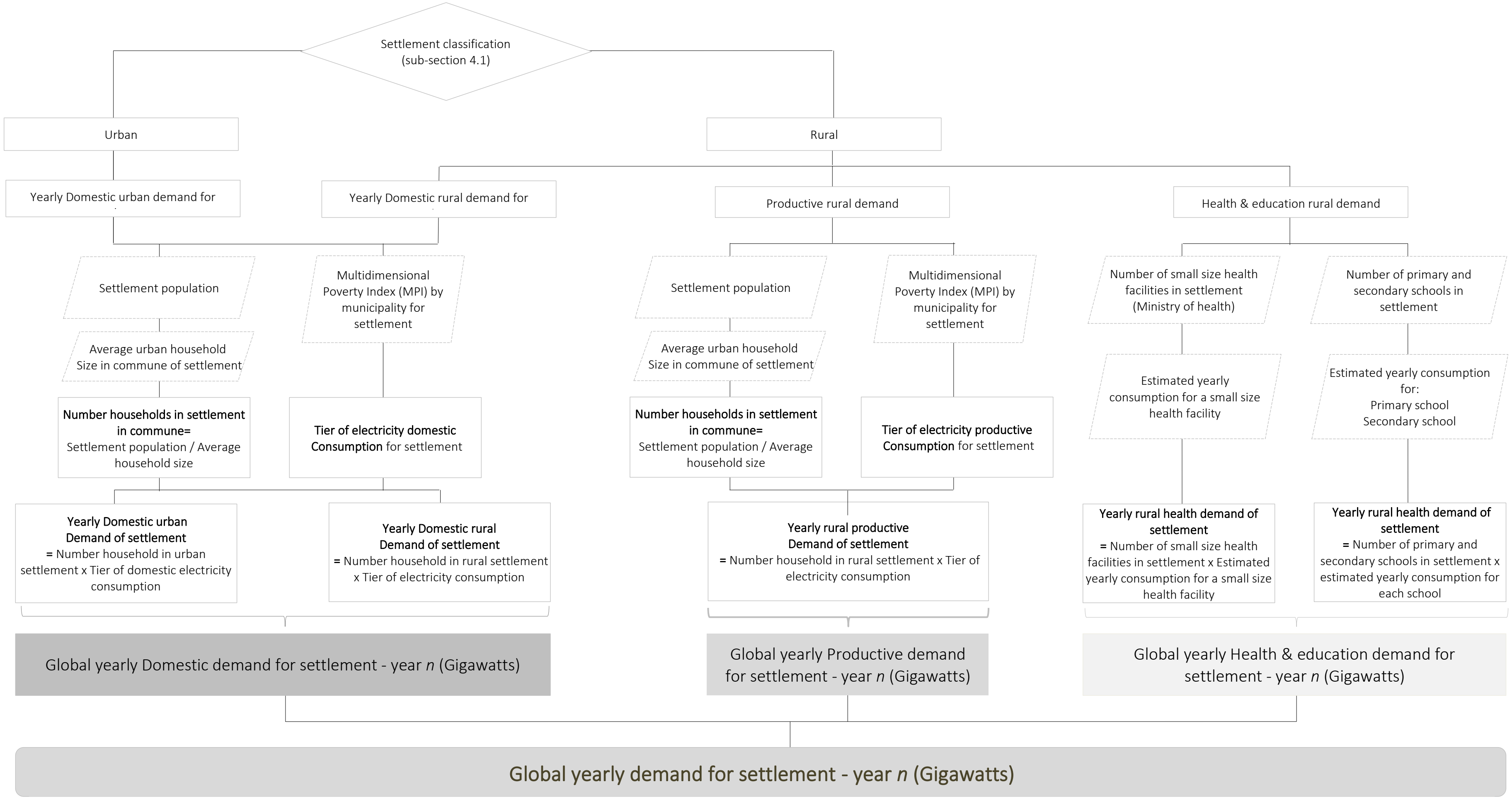


Figure 19 **–** Algorithm for rural settlement yearly demand

Produced by the author

This procedure was carried out in a spreadsheet, for each one of the 4,557 settlements, in the database.

The global rural yearly demand estimates for households, productive uses, health and education, for the period of analysis are given in Table 31.

Table 31 – Global electricity demand estimates for rural areas of Angola (2022-2030)

	Giga-watts hour per year				
Year	**Domestic - Rural**	**Productive**	**Health facilities - rural**	**Schools - rural**	**Total Gigawatts hour per year**
2022	1,311.0	287.2	1.8	2.7	1,602.7
2023	1,345.2	294.7	1.9	2.8	1,644.6
2024	1,380.3	302.4	1.9	2.9	1,687.5
2025	1,416.3	310.3	2.0	2.9	1,731.5
2026	1,453.3	318.4	2.0	3.0	1,776.7
2027	1,491.2	326.7	2.1	3.1	1,823.1
2028	1,530.2	335.2	2.1	3.2	1,870.7
2029	1,570.1	344.0	2.2	3.3	1,919.5
2030	1,611.1	352.9	2.2	3.4	1,969.6

Produced by the author

Globally speaking, by 2030, the lion's share of demand will come from household applications (81.8%) and productive uses (17.9%), while health and education are residual, despite their significant impact in the quality of life of these rural communities.

3.1.3 Data collection - Supply mix selection for the case of Angola

In the near future, grid extension will always be the most significant component of the supply mix of Angola, according to official documents (MINEA, 2018a, 2018b). Data for the approximate current state of the network, by the end of 2019, was kindly supplied by RNT[25] since it is not freely and easily available, at the time of writing. In terms of regional integration, there are connections to the north

---

[25] *Data was supplied by RNT. Our meetings took place between May and June of 2018, at RNT headquarters in Camama – Angola.*

to the Democratic Republic of Congo, and to the South to Namibia. Additionally, there is an agreement with Zambia to supply electricity to bordering communes (to the East), but it is still in the planning stages. The current situation for the grid is given in Figure 20[26].

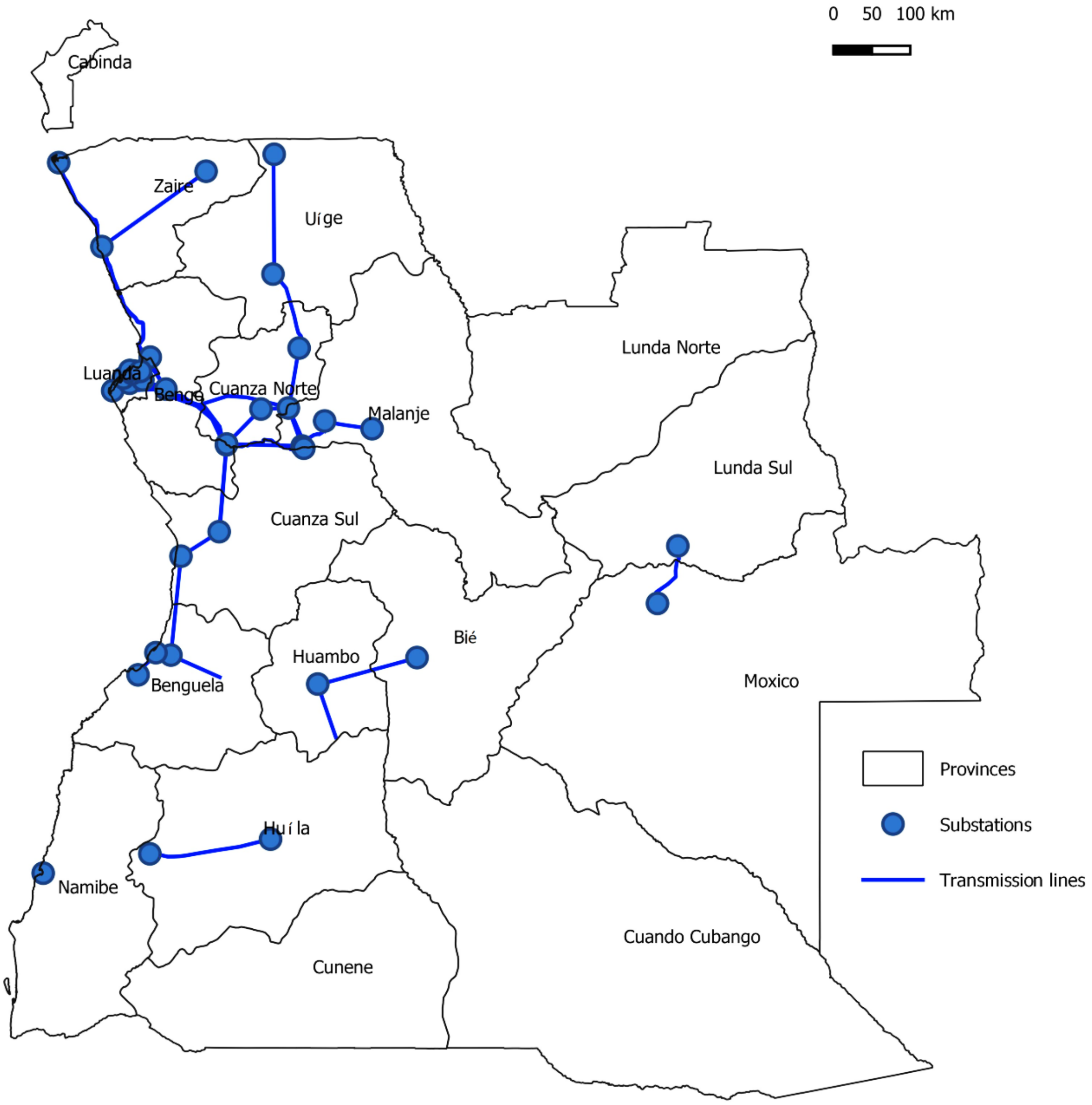


Figure 20 – National grid of Angola in 2019

Produced by the author in QGIS 3.10 using data supplied by RNT

From Figure 20 it is clear that most of the rural areas of Angola have no access to electricity and other forms of modern energy. Biomass (especially wood) still constitutes one of the energy sources most commonly used in the rural areas for heating and cooking. The country has high photovoltaic

---

[26] *This includes HV/MV substations.*

potential and is currently one of the countries in the world with the largest penetration of renewable energy in the electric sector due to the high potential of its basins powering its hydroelectric plants. This has improved the generation capacity significantly, but transport and distribution are still extremely problematic.

This research was developed under the hypothesis that these large plants will not be a viable solution, for access to electricity, for a significant number of rural settlements, in the coming years.

Quantitative data was gathered to assess the potential and interest of adding several energy solutions to the supply mix that will serve rural areas.

3.1.3.1 Hydro potential

The literature shows the large potential for smaller hydro mini-grids in SSA (Korkovelos et al., 2018; Szabó et al., 2013) but these solutions are not part of the national Angolan strategy. The focus is on large scale investments (MINEA, 2018a).

Although most of the generation, for the grid, in the country, is from hydropower (MINEA, 2018a), data on the location of potential sites for hydro mini-grids is scarce, at a national level. Data with the bodies of water and rivers is available (Korkovelos et al., 2018) and shown in Figure 19.

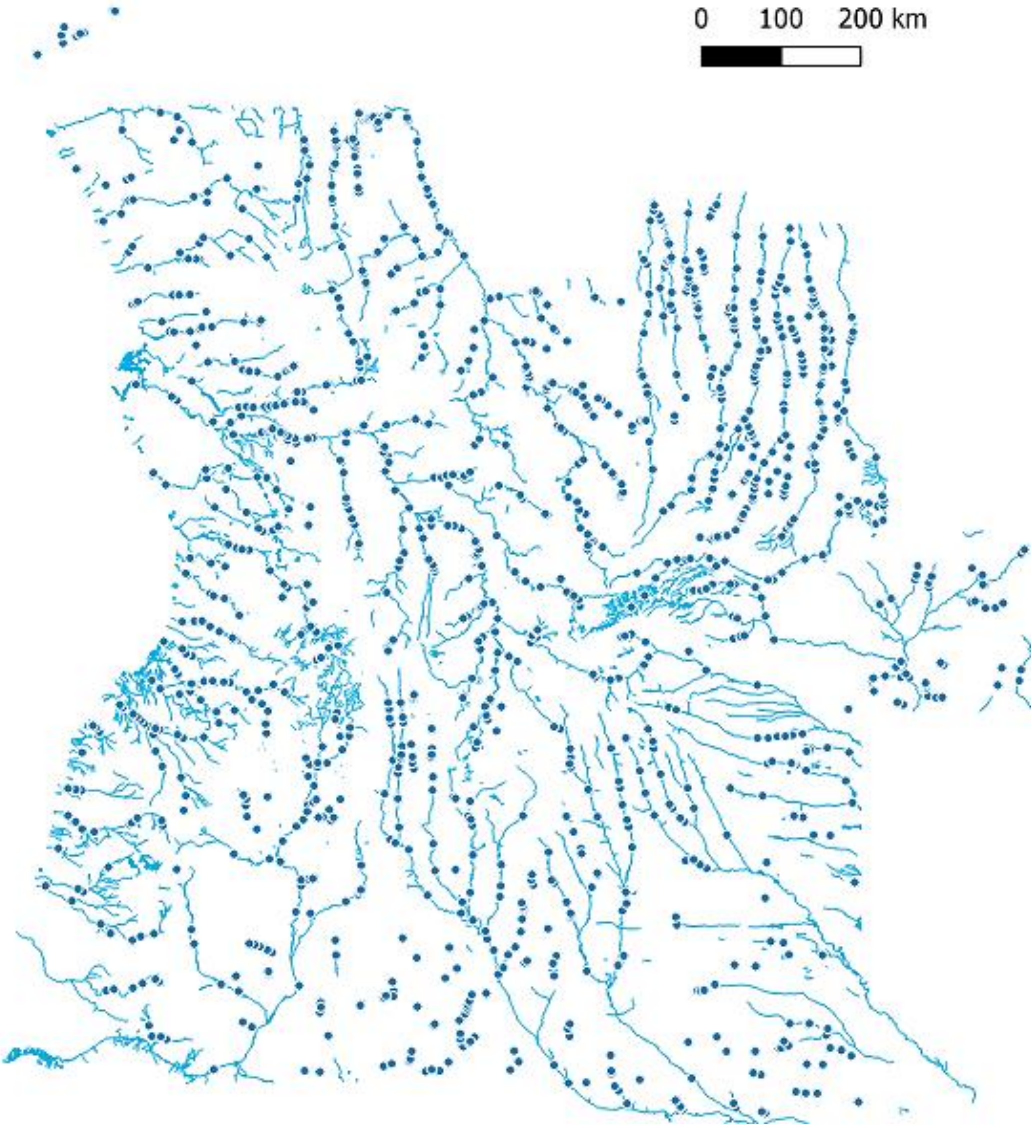


Figure 21 – Waterways, bodies of water and hydro potential in Angola.
Produced by the author in QGOS 3.10 with data from HdX (Humanitarian data exchange – UNDP) and Korkovelos et al. (2018)

Although hydro potential is present in the country as shown in Figure 21, several authors (Korkovelos et al., 2018; Morrissey, 2019) have stated that hydropower would play a limited role in the electrification of communities outside large urban centers, due to data scarcity. For the case of Angola unfortunately detailed and reliable data was unavailable at a national level. As such, hydropower was not included in the supply mix.

3.1.3.2 Photovoltaic potential

Solar energy constitutes the most widely spread and more uniformly distributed renewable resource of the country (MINEA, 2018a). It also has the fastest installation time, which is under a year in most cases. It is a viable solution for rural areas, but complete alternative solutions to grid connection, including batteries, tend to be expensive.

Raster data with yearly average potential photovoltaic production (kWh) was attained from Global Solar Atlas 2.0[27] (Figure 22).

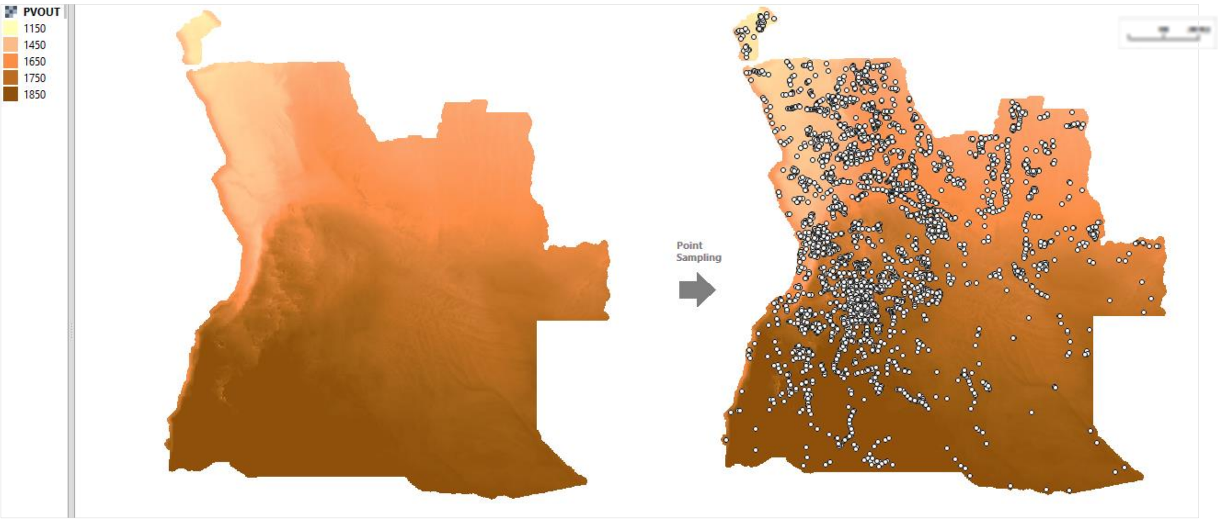


Figure 22 – GIS raster data with yearly average potential PV production (kWh) in Angola

For the purposes of this research, there is a need to know the potential in each settlement in the population maps. The data was extracted using the point sampling plugin in QGIS. With the knowledge of the potential yearly PV output, the feasibility of solar energy was assessed for each of the 4,557 rural settlements in Subchapter 3.1.1.3.

Due to the richness of the data that was available, this resource was included in the supply mix.

3.1.3.3 Diesel

While unpopular, for environmental reasons, diesel generation is still a cornerstone of power generation in SSA. It is used in production and in households, as main generation source, or as

---

[27] *Available from the World Bank at: www.esmap.org/RE_Mapping.*

backup for other sources of energy (in grid and mini-grid systems). The cost of buying and operating a generator is size-dependent, which in turn is determined by the expected demand. After sizing, prices can be attained from suppliers. Diesel power will be part of the supply mix, as a backup to solar energy.

3.1.3.4 Wind power

In the literature, wind energy is included in according to its availability. In Angola, there are limited applications for wind energy in isolated or rural areas. Most of the potential is in urban areas. The wind Atlas for Angola[28] has allowed the identification of enough potential for electricity generation in the southwestern region of the country (Figure 23).

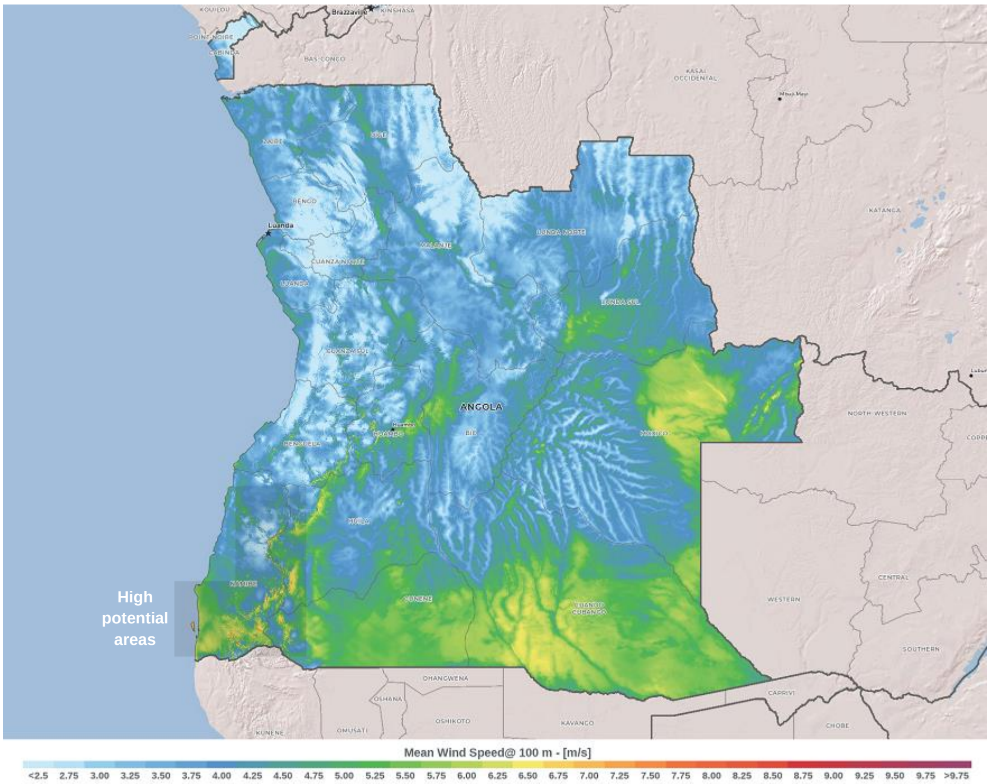


Figure 23 – GIS raster data with yearly average potential wind power in Angola.

These sites are close to the main network and sub-stations, which have enough capacity to absorb this energy without technical restrictions or significant investments. This means they will not

---

[28] https://globalwindatlas.info/area/Angola

supply isolated communities, but be directed toward increasing generation capacity for more urban areas (Republic of Angola Ministry of Energy and Water, 2017).

It is for this reason that wind power was not included in the research.

3.1.3.5 Supply and Demand

In any supply system, energy is lost (Blume, 2007). This loss, on the supply side, varies according to the technology that is considered. This means more energy needs to be generated, than the expected demand. Supply is expressed in Equation (8), as function of these two elements

$$Supply\ (kWh) = \frac{Demand\ (kWh)}{(1 - loss\%)} \tag{8}$$

For example, the national grid has a declared loss rate of 14% (MINEA, 2018a). For PV systems, in the literature, loss tends to be smaller, around 2%, also because the systems are much smaller. Diesel generators are not very efficient. In literature efficiency rates range from 28% to 33%, meaning that loss ranges from 67% to 72%.

For each technology included in the supply mix, loss was taken into account since it impacts the sizing of energy systems, which in turn drives the cost of these solutions over time.

3.1.4. Data collection - Cost structure of energy solutions for rural areas of Angola

The cost structure for energy solutions includes implementation and maintenance costs. While this is not an engineering study, enough detail should be used so the analysis is useful.

3.1.4.1 Grid extension for Angola – technological and economic assumptions

For the purposes of research, grid extension was initially defined as the process of bringing electricity from the nearest access grid point to households. The process was divided into three main components (Figure 24):

(i) Medium voltage (MV) lines;
(ii) Transformers from medium to low voltage (MV/LV transformers);
(iii) Low Voltage (LV) lines.

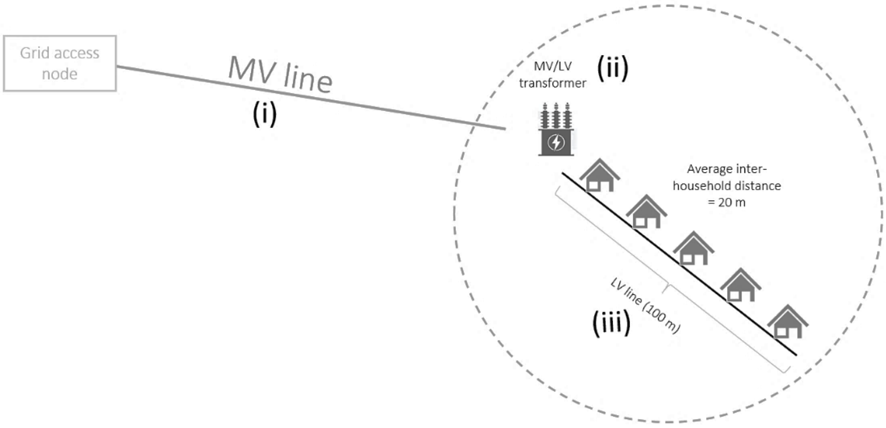


Figure 24 – Basic grid extension process.

Produced by the author using art from https://www.pinterest.com/

The cost of grid extension includes medium voltage line, low voltage line and transformers. Any point in the known grid was considered a viable interface with the grid. In order for a settlement to be connected to the grid, a line is extended from a nearby settlement or MV line. The maximum allowed MV line length considered was 165 km. That is the lengthiest MV line in Angola that could be identified from the data[29]. Data concerning grid-extension in Angola was gathered from different sources. The most important assumptions in the model are given in Table 32.

---

[29] *Distance between the Substations of Lubango and Matala (MINEA, 2018a).*

Table 32 - Case Study of Angola - Main inputs and assumptions for Grid extension

Assumption	**Input**	**Source**
WACC	11%	AO Eurobond interest rate (coupon %)
Exchange rate (Akz/USD) 26/04/2019	320.25	https://www.bna.ao/Servicos/pesquisa_cambios.aspx?idc=825&idl=1
High Voltage (HV)	60 >= kV > 35	(Diário da República II Série - N.º 54, 2019)
Medium Voltage (MV)	35 >= kV > 1	
Low Voltage (LV)	<= 1 kV	
Capital Cost for Grid extension (aerial lines)	**USD per km**	**Source**
MV line (15 kV and 33 kV)	28,026	(MINEA, 2018b)
LV line (<= 1 kV)	7,612	
Transformer (MV to LV) in USD	5,000	
Cost of replacement = Cost * (1-residual value %)	90%	Literature review (Table 10)
Transformer residual value	10%	
Line lifetime (years)	20	
Transformer lifetime (years)	10	
Recurring Costs	**Input**	**Source**
O&M – Grid	3% of Grid connection (local cost)	Literature review (Table 10)
O&M – Transformer	3% of Transformers cost	
Minimum distance between LV lines	100 m	(Diário da República II Série - N.º 54, 2019)
100 m of LV line allows for the connection of 5 household		
Technical losses	14%	(MINEA, 2018a)

Produced by the author

It is worth mentioning that beyond the initial capital investment, there are costs with reinvestment as well as those related with operational and maintenance activities. The cost of grid extension is given by Equations (9) through (14).

Grid connection Cost = (i) MV line cost + (ii) Transformer MV/LV + (iii) LV line cost +
+ (iv) Replacement cost + (v) Recurring cost (9)

where

(i) MV line cost = (distance[30] *to the nearest MV line or Grid connected settlement) x cost/km* (10)

(ii) Transformer MV/LV = (Number of households/30[31]*) x cost per transformer* (11)

(iii) LV line cost = (number of households/5) x (100 m / 1,000 m)x cost per km (12)

(iv) Replacement cost = Transformer MV/LV x (1-residual value %) (13)

(v) O&M_Grid + O&M_transformers, over a period of 20 years (14)

All calculations were carried out in a spreadsheet, for a period of 20 years, considering an adequate weighted average cost of capital (WACC). The period of 20 years comes from the expected lifetime for the line (listed in Table 32). This same period of time was applied to other technologies to enable objective economic comparison. This is standard practice in comparative feasibility analysis studies.

3.1.4.2 Photovoltaic in Angola – technological and economic assumptions

Regarding photovoltaic systems, assumptions came from both local suppliers and the literature (Table 33).

---

[30] *Distances were measured in a straight line.*

[31] *The transformer can cover 300 m radius (assumption of 30 households for 1 transformer) (MINEA, 2018b).*

Table 33 – Case study of Angola - Main inputs and assumptions on PV energy solutions

Capital Cost for PV systems (aerial lines)	**Input**	**source**
LV line (<= 1 kV)	7,612	(MINEA, 2018a, 2018b)
Transformer (MV to LV) in USD	5,000	
Cost of replacement = Cost * (1- residual value %)	5%	literature (Table 10)
Transformer residual value	10%	
PV system (panels + charger + inverter): USD/kWh	350	from supplier
Batteries: USD/kWh	315	
PV system residual value	5%	from literature (Table 11)
Recurring Costs	**Input**	
O&M – PV system (panels + charger + inverter)	3%	from literature (Table 11)
O&M – Transformer and LV Lines	3%	
Minimum distance between LV lines	100 m	(Diário da República II Série - N.º 54, 2019)
100 m of LV line allows for the connection of 5 household	5	
Household coverage per transformer	30	expert opinion
Brands / Models		Module - Open Renewables / 310-HM60 + control panel + structure. Batteries OPzS / 680 OPzS 2900 AhC10-2V.
WACC	11%	AO Eurobond interest rate (coupon %)
Technical loss	0%	Negligible
Tariff expected growth (epr)	0%	-
PV degradation factor	1%	(Jordan & Kurtz, 2013)
Line lifetime (years)	20	literature (Table 10)
Transformer lifetime (years)	10	
PV system lifetime (years)	20	
Batteries lifetime (years)	4	

Produced by the author

The cost of a PV system was modeled according to Equations (15) through (18):

$$\text{PV system Cost} = \text{(i) PV cost} + \text{(ii) Replacement costs} + \text{(iii) Recurring costs} \tag{15}$$

where

$$\text{(i)} \quad \text{PV cost} = \text{PV system (USD/Kw)/correction coeff.}^{32} + \text{Batteries (USD/Kw)} \tag{16}$$

$$\text{(ii)} \quad \text{Replacement cost}^{33} = \text{Batteries (USD/Kwh)} \times (1 - \text{PV system residual value \%}) \tag{17}$$

$$\text{(iii)} \quad \text{Recurring costs} = \text{O\&M\_PV} + \text{O\&M batteries} \tag{18}$$

The remaining cost components were described in Subchapter 3.1.4.1. They include LV lines, transformers and respective O&M and they are computed according to Equations (10) through (14). These elements are inherent to all solutions in this research.

3.1.4.3 Hybrid systems for Angola – technological and economic assumptions

Diesel generation is not environmentally sound, but it is an essential element as backup, for the implementation of isolated systems in SSA. Its use, in this research, has been reduced to a secondary role as a backup solution to photovoltaic energy (Figure 25).

---

[32] *This correction coefficient is equivalent to the normalized values of PV potential extracted in Subchapter 3.1.3.2. Higher values means, on average less solar panels and less cost.*

[33] *Over the 20-year period only the batteries are expected to be replaced.*

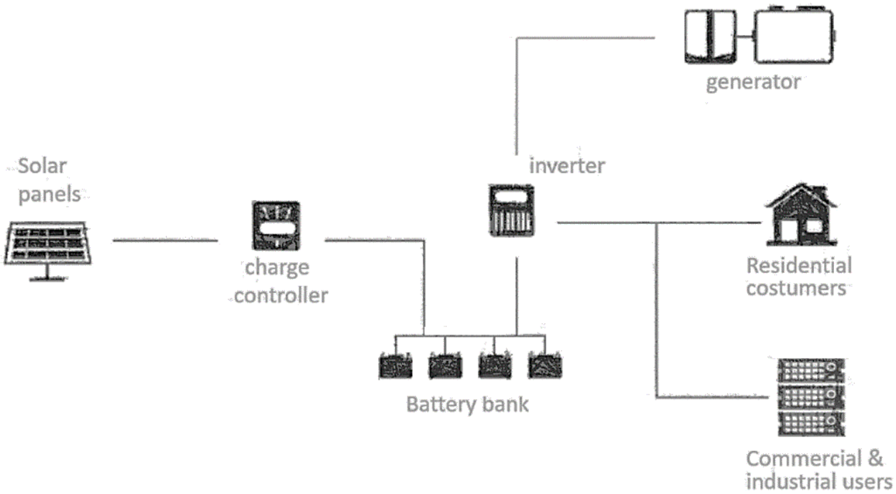


Figure 25 **–** Simplified scheme of a hybrid PV/Diesel installation

Adapted from Mamaghani et al. (2016) using art from https://www.pinterest.com/

With the inclusion of hybrid systems (PV and diesel) all the assumptions from PV systems apply and new ones related to diesel generation are introduced in Table 34.

Table 34 – Case study of Angola - Main inputs and assumptions for diesel energy

Capital Cost for Diesel Generator extension (aerial lines)	**Input**	**source**
Generator residual value	0.10	from supplier
Diesel Generator cost (equipment, transport and installation) : USD/kW	315	from supplier
O&M – Diesel Generator	5%	from literature (Moner-Girona et al., 2019; Szabó et al., 2013)
Generator conversion efficiency (l/kWh)	0.20	from supplier specification
Transport truck diesel consumption (liters/hour)	12	(Moner-Girona et al., 2019; Szabó et al., 2013)
Volume of diesel transported (liters)	300	
Generator Brand / Model		Himoinsa / 1 HFW-125 T5 with 240 liter tank
WACC	11%	AO Eurobond interest rate (coupon %)
Tariff expected growth (epr)	0%	(MINEA, 2018a, 2018b)
Diesel Generator lifetime (years)	10	(Moner-Girona et al., 2019; Szabó et al.,

Produced by the author

The cost of diesel generation was modeled according to Equations 19 through 25.

Diesel generator cost = ***(i)*** *Generator cost +* ***(ii)*** *Replacement costs +* ***(iii)*** *Recurring costs* (19)

where

(i) Generator cost = Diesel Generator cost (USD/ kW) x Power (kW)[34] (20)

(ii) Replacement cost = Generator cost x (1-Generator residual value) (21)

(iii) Recurring costs = O&M_diesel_generator + Complete cost of fuel[35] (22)

(iv) Comp. cost of fuel = (Fuel$_{direct}$ *+ Fuel*$_{transport}$*) x (supply/kWh) x (Generator conversion efficiency)* (23)

(v) Fuel$_{direct}$ *= pump price*[36] *(USD/liter)* (24)

(vi) Fuel$_{transport}$ *= 2 x Fuel*$_{direct}$ *x Truck diesel consump. (liter/hour) x Time factor (hours) / Vol. of diesel transp.* (25)

For fuel transport, distances were measured as straight lines between two points and corrected using a travel time factor. This travel time factor was extracted using a travel time map (Weiss et al., 2018). More time means a higher correction of travel distance and, consequently, a higher cost.

Oil prices are notoriously difficult to predict. The price of oil in the international market was used as a proxy of real pump price of diesel. Commercial logistics margins were added according to literature (International Monetary Fund, 2015). The real pump price per liter for January 2020 was 0.54$/liter[37].

All calculations were carried out in a spreadsheet, for a period of 20 years, considering an adequate cost of capital.

---

[34] *Power is estimated, for each settlement, as given in sub-chapter 3.1.3.5.*

[35] *Partially adapted from Szabó et al. (2013) and Moner-Girona et al. (2019).*

[37] *Brent Barrel Price/capacity (in liters) x (1+ margins in %) = 55$ / 159 liters x (1.55) = 0.54 $/liter.*

3.1.5 Data collection - Computing NPC and LCOE for each technology in the supply mix

Having defined the cost structure, its elements and using some assumptions, metrics that will support decision-making were computed.

Two metrics were elected, as basis for decision-making:

i) The NPC, which is a measure of the global cost of a technological solution. It was measured over a period of 20 years. This allowed for the comparison of all cost elements described in Subchapter 3.1.4, not only initial investment. Operating under the assumption that investments will be supported with public financing (in USD), the Angolan Eurobond interest rate (coupon %) was used as discount factor. NPC was computed using the formulation in Equation (2), under the assumption that tariffs would remain stable as they have been for years through subsidizing.

ii) The LCOE is a measure of the cost in USD per unit of energy produced (USD/kWh). It was computed using the formulation in Equation (3). The same assumptions that were used for NPC were applied here. Total supply was determined using Equation (8), with the appropriate loss rate, for each technology. In this study, issues of intermittency were minimized, as the hybrid mini-grids that are being proposed include batteries and/or a diesel generator for backup. This measure allows for accurate comparisons between grid and mini-grid solutions.

Very few studies in literature do not include these metrics. Exceptions are studies where the economic component is irrelevant, which is not the case in Angola since economics are the most relevant factor for nationwide decision-making.

In order to have a deeper knowledge on the cost factors that drive grid extension, photovoltaic energy and diesel generation, a sensitivity analysis of these two metrics was also carried out.

Preliminary calculations were independently performed for each alternative in the supply mix using overnight-built modeling (Morrissey, 2019).

Cost analysis, including sensitivity analysis of key cost factors was performed for different technologies:

- Grid extension with the assumptions listed in Table 32;
- PV mini grid (with batteries) and diesel under the assumptions given in Table 11;
- Diesel generation to be implemented in all households of rural communities in Angola according to the assumptions from Table 34.

Among these technologies, grid extension was the most complex in terms of cost estimation.

3.1.5.1 Cost analysis of grid extension, for rural communities of Angola

On a countrywide study, it is not possible to use detailed network designs. A planning study should not replace detailed engineering analysis of each location, but should produce a layout that, while not technically accurate, is a reasonable approximation for cost of grid expansion, which can be used to support decision-making (Blechinger et al., 2019; Ciller & Lumbreras, 2020; Deichmann et al., 2011; Mentis et al., 2016; Parshall et al., 2009). To ensure that the modeling solution fulfills its purpose, graph theory was employed to produce an initial output with a full grid growth scenario. Vector data containing 4,557 georeferenced rural settlements was used as input in the QGIS MST plug-in implementation of Kruskal’s MST algorithm described in Çalışkan and Anbaroğlu (2020), to produce the spanning tree. The resulting vector layer provided the minimum spanning tree for grid using inter-settlement distance as edge weight. The approach minimizes inter-settlement distance, in line with literature, as distance is a defining cost factor for grid extension. The output of this procedure is given in Figure 26.

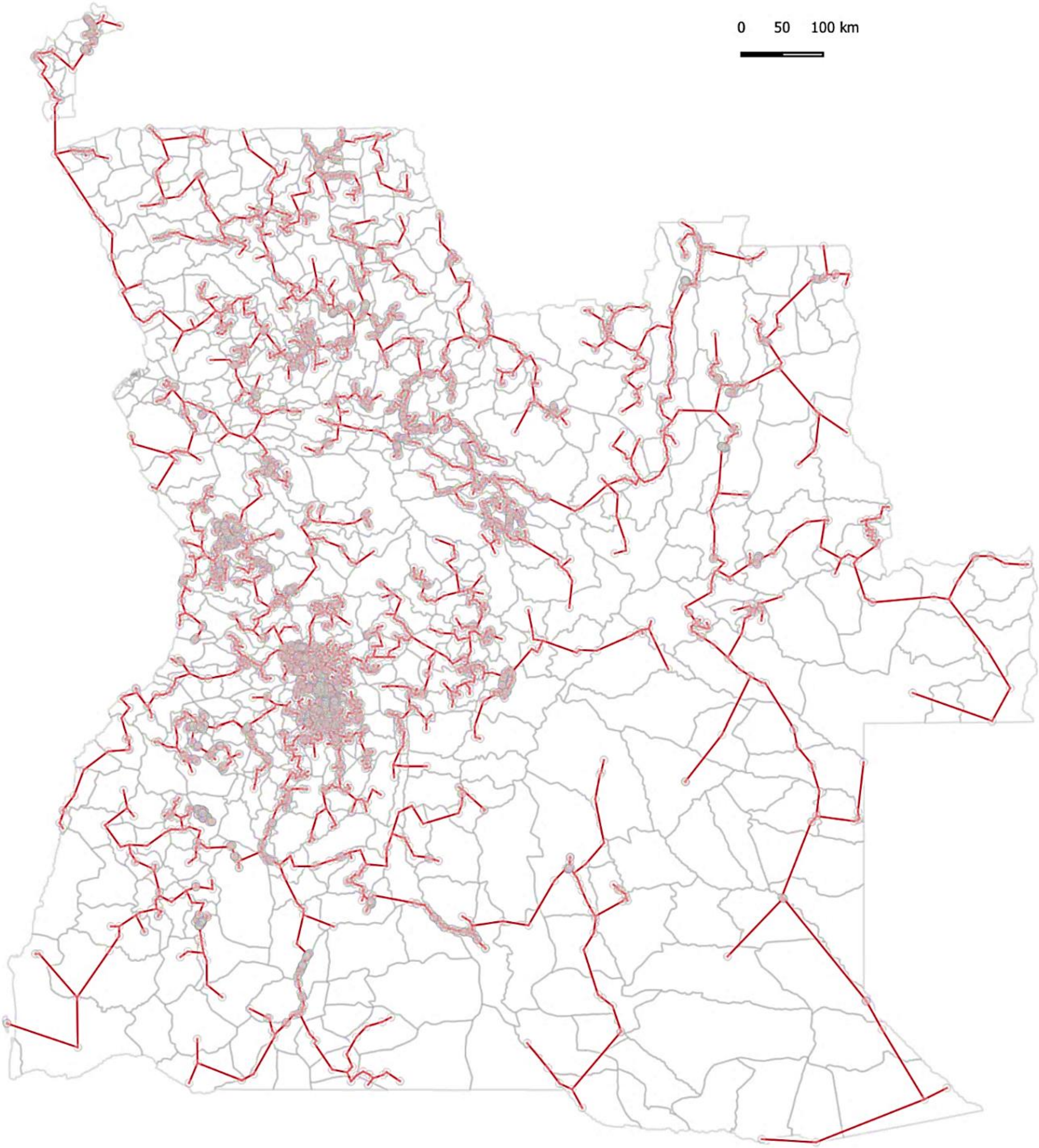


Figure 26 – Complete Grid extension for access to electricity using Kruskal's algorithm to solve the minimum spanning tree problem.

Produced by the author in QGIS 3.10

This MST provided useful data in terms of total MV line length in kilometers and a first look at what a complete grid might look like. The full output matrix with settlements and distances was exported to a spreadsheet. This was combined with assumptions (Table 32) and Equations (9) through (14), from Subchapter 3.1.4.1, to compute NPC and LCOE, per new vertex connected, for all 4,557 rural settlements. The global results are given in Table 35. Some of the communes do not contain any of these rural settlements. This is due to urban settlements with high population densities (in the center of the country) or small population sizes (in southern areas of the country).

Table 35 – Global cost if rural Angola is connected using grid extension

Global NPC (USD)	2,245,156,759
Average LCOE (USD/kWh)	0.8305

Produced by the author

A simple sensitivity analysis was performed for some key variables for NPC and LCOE (Figure 27). To this end tornado charts were produced using a spreadsheet[38]. These charts are a modified version of bar charts and are also one of the classic tools of sensitivity analysis used by decision makers to have a quick overview of the risks involved. They show how changes in the inputs (cost components) can potentially impact on the output (cost in USD). In this study the impact was measured by recalculating total costs considering changes from -20% to 20% for each key input.

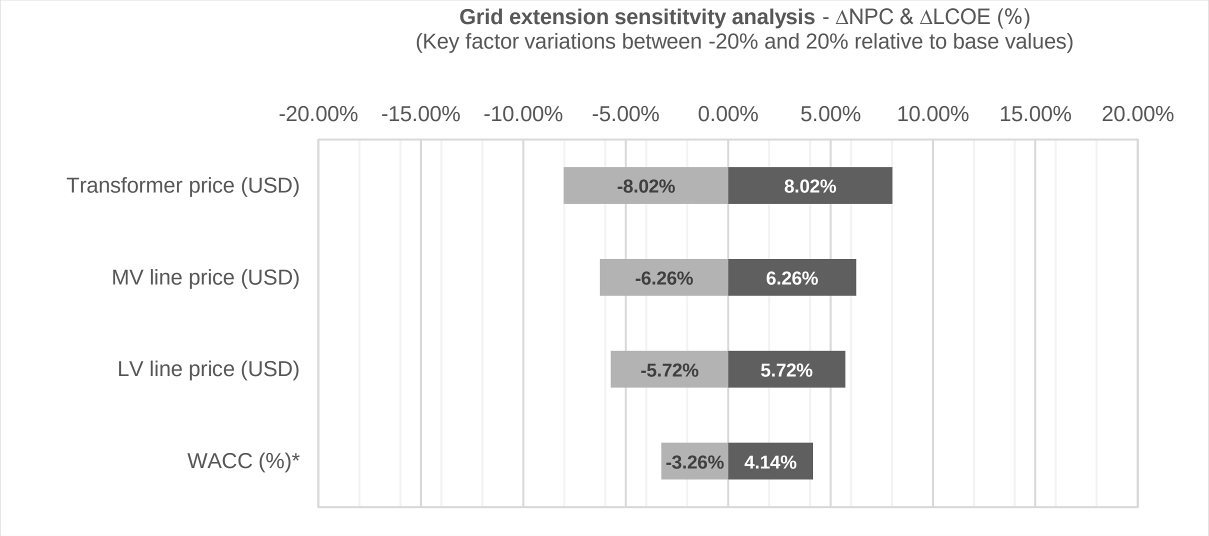


Figure 27 – Grid extension: Sensitivity analysis for NPC & LCOE vs. main cost factors[39].

Produced by the author

The analysis confirms not only distance (MV line) as one of the main factors in grid extension, but also that transformer cost significantly influences NPC and LCOE results. This is a global overview

---

[38] *The way in which these charts can be produced is given in https://www.wallstreetmojo.com/tornado-chart-in-excel/.*

[39] *For WACC the sensitivity analysis is relative to NPC.*

and does not exclude the need to individually treat the results of each location, in order to better support decision-making as is presented in Subchapter 3.2. The remaining energy options (PV and diesel) are also analyzed.

3.1.5.2 The cost of PV, for rural communities of Angola

In terms of sensitivity analysis PV and diesel were analyzed separately, so the factors that drive cost, in each technology are clearer.

Cost analysis for photovoltaic systems (including batteries) is more straightforward. Under the assumptions in Table 33 (PV assumptions) all costs were considered, over a span of 20 years. This meant including in the calculation Equations (15) through (18) for PV. NPC and LCOE were computed in a spreadsheet for each one of the 4,557 settlements. Global cost is given in Table 36.

Table 36 – Case study of Angola – Global PV results

Technology	NPC (USD)	LCOE (USD/kWh)
Hybrid (PV + Batteries)	6,532,489,374	0.393

Produced by the author

Although a global hybrid solution for all rural communities is economically unfeasible, for some rural settlements, hybrid solutions may be a viable option. A sensitivity analysis was carried out for PV systems and the same procedure and tornado charts used for Grid extension were applied to this technology. The results are given in Figure 28.

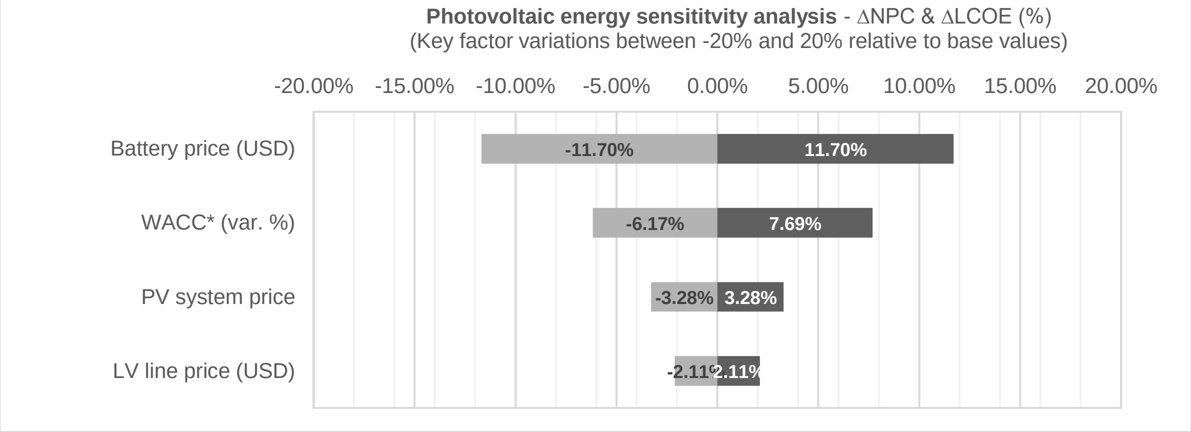


Figure 28 – PV energy: Sensitivity analysis for NPC & LCOE vs. main cost factors[40].

Produced by the author

---

[40] *For WACC the sensitivity analysis is relative to NPC.*

It came as little surprise that, for PV solutions, the price of batteries represented one of the largest costs in absolute terms but also the most significant source of potential volatility of NPC and LCOE.

Solar home systems (SHS) were also analyzed. SHS are quite limited in terms of their applications and are a weak replacement for grid connection. Regardless, the global NPC and LCOE of a solution to electrify rural communities in Angola using SHSs were computed under the assumptions from Table 12. The results are given in Table 37.

Table 37 – Case study of Angola – Global SHS results

Technology	**NPC (USD)**	**LCOE (USD/kWh)**
Diesel	1,550,716,018	0.344

Produced by the author

This technology, while obviously cheaper, is the only one that will be proposed where the issue of intermittency is truly present and represents the bare minimum in terms of access to electricity. SHS's are a sub-set of PV energy. Their main cost factor is price since the solution comes pre-packaged and is not customizable, unlike the remaining alternatives. Sensitivity analysis was not performed for this solution.

It was considered that, for the purposes of the research, this alternative would not be included in the potential supply mix.

The last component of the energy mix to be analyzed is diesel generation.

3.1.5.3 The cost of diesel, for rural communities of Angola

Cost calculation for diesel generation combines the assumptions from Table 13 and Equations (19) through (25).

The total cost of implementing a diesel solution that encompasses all rural communities, are given in Table 38.

Table 38 – Case study of Angola – Global diesel results

Technology	NPC (USD)	LCOE (USD/kWh)
Diesel	5.024.562.776	0.319

Produced by the author

Diesel generation remains a cornerstone of access to electricity in SSA, but it is not an economically feasible solution for all communities since the cost can be very high. A sensitivity analysis was also carried out for diesel power (Figure 29). The procedure is in line with the one used for grid extension and PV.

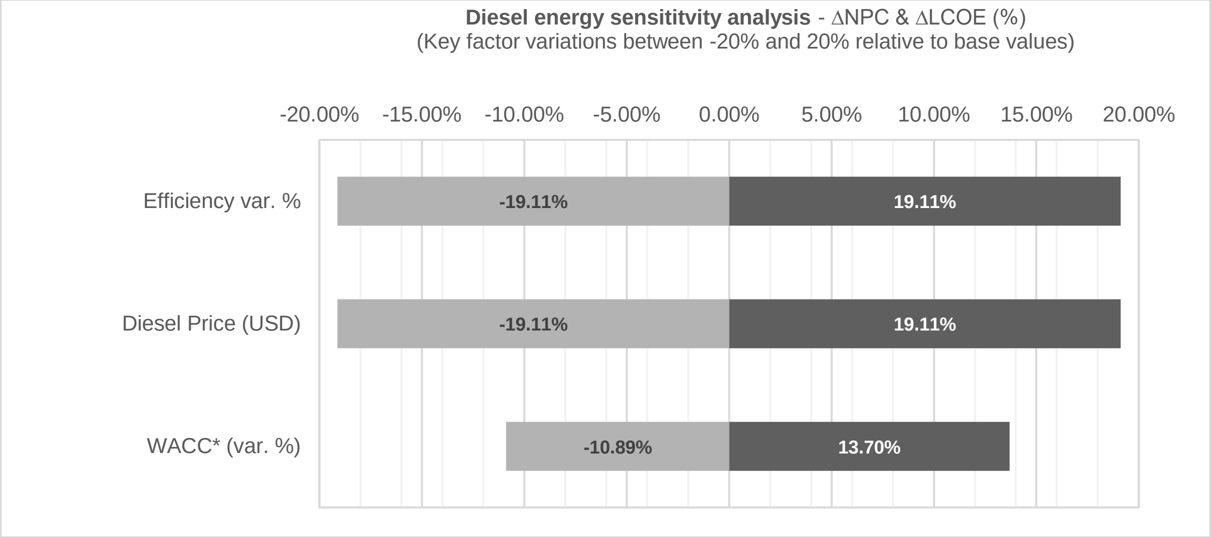


Figure 29 – Diesel: Sensitivity analysis for NPC & LCOE vs. main cost components.

Produced by the author

Although the cost of full rural electrification using diesel or PV and diesel hybrid solutions is prohibitively high when compared to grid extension, this does not mean these solutions do not make sense in many settlements. For example, electrifying large cities using PV is currently unfeasible but connecting several smaller settlements is often desirable economically. Beyond economic considerations it is useful to consider that off-grid solutions can target isolated communities instantly and grid growth favors proximity to the current infra structure deepening regional asymmetries.

Comparing these results globally, hybrid grids and diesel are much more expensive and its cost is subject to greater volatility but as mentioned, this preliminary analysis does not compare alternatives on a location-by-location basis and has the usual limitations of overnight built models.

On the last stage of this study's methodology a multi-period analysis under different scenarios was carried out for all rural settlements that were previously identified, considering combinations of grid, PV and diesel power. The outcomes of potential alternatives should be clearly presented, to support decision-making as efficiently as possible. This procedure is detailed in Subchapter 3.2.

3.2 Selection of rural communities for mini-grid connection

The goal of the last stage of this research's methodology is the selection of rural communities of Angola where all relevant factors coalesce for access to electricity using hybrid-mini grids and grid extension. This study's decision-making support tool allows the production of outputs, on a yearly basis. The assumption for this multi-period approach is that the yearly budget will be in line with previous national budgets and that the analysis will have 2030 as an important milestone as this year is a reference for Sub-Saharan Africa countries in terms of the goals subscribed in United Nations electrification programs. In this time period the available budget may not be enough to ensure full grid coverage at a nationwide level.

3.2.1 Scenarios for rural electrification – supply mix and assumptions

There is significant uncertainty associated with decisions on the subject of electrification at a nationwide level. The territory is vast with some settlements being close to the grid while others are distant, and the people have different ways of life that constrain their demand for energy. It is often difficult to decide what would the "best" outcome that holistically considers all these elements. It is in this type of problem that the value of scenario planning is realized by a range of features rather than the ability to predict a particular result. The use of scenarios can make leaders comfortable with the ambiguity of decisions and provide a deeper foundation of knowledge.

In order to propose a solution for this study's central problem and provide as much information as possible for decision-making, scenario analysis was employed. Three scenarios were considered, representing general directions that may be followed in electrification of settlements in rural areas.

- Scenario 1 - A grid exclusive scenario where only Grid extension was considered with assumptions from Table 32 and Equations (9) through (14). This scenario is a useful baseline to assess the cost and degree in which grid extension, on its own, can address the problem of rural electrification in the timeframe from 2022 to 2030 although it leaves out other potential alternatives.
- Scenario 2 – A grid intensive scenario combining grid extension and hybrid mini-grids (PV

with batteries and diesel). This scenario mitigates one of the main limitations of the previous scenario since it considers an alternative to grid extension. This alternative is a complete substitution solution for grid extension but costs may increase when compared with simpler hybrid solutions. In this scenario, for the mini-grids, PV represents the majority of yearly supply while diesel is a backup. For grid extension the assumptions and equations remain the same as in the previous scenario. For PV energy the model assumptions were those from Table 33 while Equations (15) to (18) were used for calculations. Regarding diesel generation, assumptions from Table 34 were used in the model along with Equations (19) through (25). NPC and LCOE from both technologies (grid and hybrid) are computed for each location.

- Scenario 3 – A balanced scenario, that also combines grid extension and hybrid mini-grids but with a different configuration. In this scenario there is a trade-off in the completeness of the solution for some expected reduction in cost. This scenario includes PV systems (without batteries) and diesel. Apart from assumptions regarding batteries, all other assumptions and equations from the previous scenarios were equally applied here. Only the contribution of each technology in the hybrid solution changed. NPC and LCOE from both technologies (grid and hybrid) were computed for each location.

These proposed scenarios run from 2022 to 2030 offering different degrees of complexity. In scenario 1 (grid exclusive) only one technology is present so selection can be made according to a criterion (cost minimization, for example), constrained by yearly expenditure of 180,000,000 USD.

The remaining two scenarios (scenario 2 – grid intensive and scenario 3 – balanced) are more complex and realistic, since the choice between technologies is allowed.

3.2.2 Multi-period modelling for Scenario 1 (Grid exclusive scenario)

From the beginning, it became clear that producing yearly results would be an asset for decision makers. One of the main criticisms leveled at most studies is that the grid is presented in its final state. To avoid this potential pitfall a procedure was put in place that, given a yearly budgetary constraint, enabled visualizing the periodic evolution of the grid and hybrid solutions, within the period of analysis (2022-2030).

Since scenario 1 only considers grid extension, the goal is to visualize its yearly growth. The MST that was previously attained in Figure 26 is the basis for this procedure. This time Prim’s algorithm was used. Despite having chosen Prim’s algorithm, the spatial data that was used was originally produced in QGIS so it was necessary to extract the data. In order to improve the time efficiency of the algorithm, an adjacency list of the previously attained Minimum Spanning Tree was produced. An adjacency list contains the vertices that are connected to each vertex. Adjacency lists are a good fit for sparse graphs. Typically, a sparse (connected) graph has about as many edges as vertices, which is this study's case with 4,557 vertices and 4,556 edges. The extracted adjacency list served as an input in a Python implementation of Prim's algorithm. The code is available in Appendix B – Python implementation of Prim's algorithm.

The code was tested in a Jupiter Notebook running Python V.3 and successfully ran with the extracted adjacency list. With this process it was possible to grow the network, starting from any vertex, and always attain a connected tree. This process allowed the visualization of the network evolution, at each time step, in the period from 2022 to 2030.

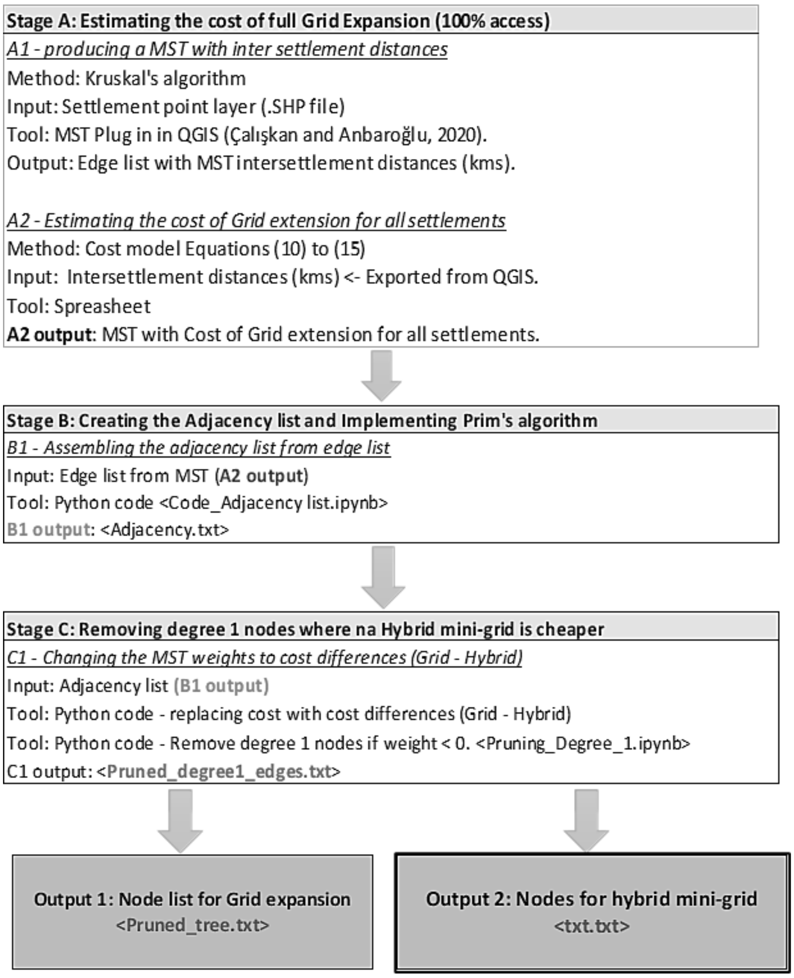
This procedure was used for all scenarios, starting with scenario 1 (grid exclusive). This scenario focused exclusively on grid extension and NPC minimization.

3.2.3 Multi-period modelling for scenario 2 (grid intensive scenario) and scenario 3 (balanced scenario)

Literature on energy planning is split into two main categories: studies that use linear programming and studies that use heuristic algorithms. The latter approach is almost uniformly employed in national level studies (Deichmann et al., 2011; Kemausuor et al., 2014; López-González et al., 2019; Mamaghani et al., 2016; Modi et al., 2013; Ohiare, 2015; Parshall et al., 2009). Given the dimension and complexity of planning for rural access to electricity in SSA a heuristic approach was used. The advantage of this type of approach is that heuristic algorithms produce feasible solutions fast. However, these are solutions that are not necessarily optimal solutions. While this is a problem for smaller, high-resolution issues (like supply chain problems) it is acceptable for a study that focuses on nationwide planning. Here the goal is to find trends, tradeoffs and possible avenues for development of the national grid that support decision-making. Studies that use linear programming employ detailed generation and transmission models, and typically have a smaller scope (Ciller & Lumbreras, 2020).

In scenario 2 (grid intensive) and scenario 3 (balanced), hybrid connections were introduced in the supply mix, with different configurations. Scenario 2 combines PV (with batteries) and diesel. In this scenario diesel represents only 20% of supply while scenario 3 uses PV (without batteries) but diesel represents 50% of supply, to compensate the absence of batteries.

For both scenarios there are communities that will always be cheaper to connect via hybrid solutions. An algorithm was introduced to find these communities. This procedure needed to be carried out while maintaining the connectivity of the MST that was produced using Prim's algorithm (Subchapter 3.2.2). The driving factor of the algorithm, at this stage, was the NPC metric. The general procedure that was carried out is given in Figure 30.



Stage A: Estimating the cost of full Grid Expansion (100% access)

A1 - producing a MST with inter settlement distances

Method: Kruskal's algorithm
Input: Settlement point layer (.SHP file)
Tool: MST Plug in in QGIS (Çalışkan and Anbaroğlu, 2020).
Output: Edge list with MST intersettlement distances (kms).

A2 - Estimating the cost of Grid extension for all settlements

Method: Cost model Equations (10) to (15)
Input: Intersettlement distances (kms) <- Exported from QGIS.
Tool: Spreasheet
A2 output: MST with Cost of Grid extension for all settlements.

Stage B: Creating the Adjacency list and Implementing Prim's algorithm

B1 - Assembling the adjacency list from edge list

Input: Edge list from MST (**A2 output**)
Tool: Python code <Code_Adjacency list.ipynb>
B1 output: <Adjacency.txt>

Stage C: Removing degree 1 nodes where na Hybrid mini-grid is cheaper

C1 - Changing the MST weights to cost differences (Grid - Hybrid)

Input: Adjacency list **(B1 output)**
Tool: Python code - replacing cost with cost differences (Grid - Hybrid)
Tool: Python code - Remove degree 1 nodes if weight < 0. <Pruning_Degree_1.ipynb>
C1 output: <**Pruned_degree1_edges.txt**>

Output 1: Node list for Grid expansion
<Pruned_tree.txt>

Output 2: Nodes for hybrid mini-grid
<txt.txt>

Figure 30 – Outline of the MST pruning procedure

Produced by the author

In this procedure, Stage A has already been accomplished when the MST was produced using Kruskal's algorithm. Stage B1 involves extracting the adjacency list from the MST. The code for this stage is given in Appendix C – Python code for < code_Adjacency list>.

The main challenge in removing vertices from an MST that represents an electric grid is to maintain connectivity. Removing an interior vertex disconnects subsequent vertices, with unforeseen consequences, for example in terms of cost. In the literature, mini-grid sites tend to be selected via pre-set criteria (minimum demand, maximum distance or maximum LCOE), which side steps this problem, by ignoring the impacts on the grid. This is a valid approach, but for this research an assessment was performed on how many sites could be effectively cheaper to connect via this research’s hybrid mini-grid configurations, using NPC as metric.

On Stage C, a procedure is used to remove from the MST all leaves with an NPC of installing a hybrid-mini grid inferior to the NPC of grid extension. By removing only the leaves there is an assurance that the grid remains connected. For the grid intensive scenario and balanced scenario, the edge weights of the MST are changed to the cost difference between hybrid mini-grid and grid extension. Thus, edges could have null, positive or negative weights. Those leaves with a negative weight are removed. These pruned vertices are suitable for hybrid mini-grid connection, from a global cost of investment perspective (lower NPC). The algorithm was coded in Python and is given in Appendix D – Python code for < Pruning_Degree_1 >). An illustration of the procedure is given in Figure 31.

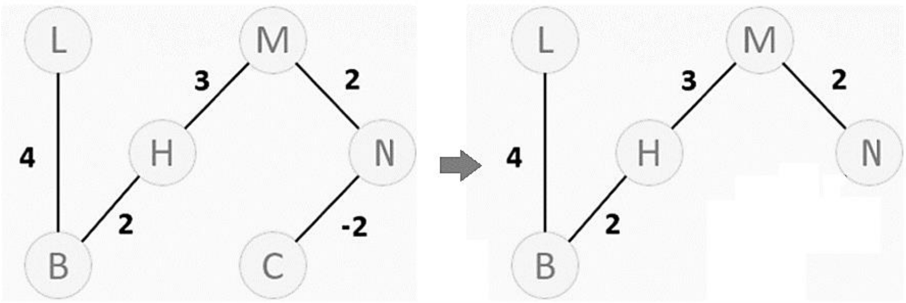


Figure 31 – Illustration of leaf removal procedure

Produced by the author

Prim’s algorithm is then used to select the hierarchy in which the connections should occur, for each technology, according to different independent criteria. Four potential criteria were considered producing scenario variations (Table 39).

Table 39 – Scenario analysis vs. criteria for the period from 2022 to 2030

Scenarios	Min{NPC}	Min{LCOE}	Max{MPI}	Max{Population size}
Scenario 2 (Grid intensive) & Scenario 3 (Balanced)	Connect settlements, by grid or hybrid, while minimizing NPC. Always select the settlement with lowest NPC, regardless if it is by grid or hybrid solutions.	Connect settlements, by grid or hybrid, while minimizing LCOE. The next settlement to be connected is always the one with the lowest LCOE. The focus is on the sustainability of the system	Connect settlements, by grid or hybrid, while maximizing MPI (Connect the poorest settlements first).	Connect settlements prioritizing those that have highest population first, regardless if it is by grid or hybrid solutions

Produced by the author

These goals or variations represent common concerns of decision makers and could potentially produce different ways of tackling the issue of rural access to electrify.

The python code for applying Prim's algorithm for each pair (scenario, criteria) is given in Appendix E – Python code – Prim's algorithm for settlement selection. This process employed the algorithm using NPC, LCOE, MPI and settlement population as edge weights, under a 180,000,000 USD yearly budgetary constraint (plus unspent budget from the previous year, if it is available), for the period from 2022 to 2030. The MST was readjusted to a new configuration that suited this study's approach. All pruned leaves from each scenario (hybrid-connected vertices from the procedure illustrated by Figure 31) were then reattached to the initial vertex of the MST. This allowed the MST to grow and remain connected, independently from hybrid connections. The procedure is illustrated in Figure 32 where vertices A, C and D represent leaves with negative weights that were pruned and reattached to the MST using an interface vertex.

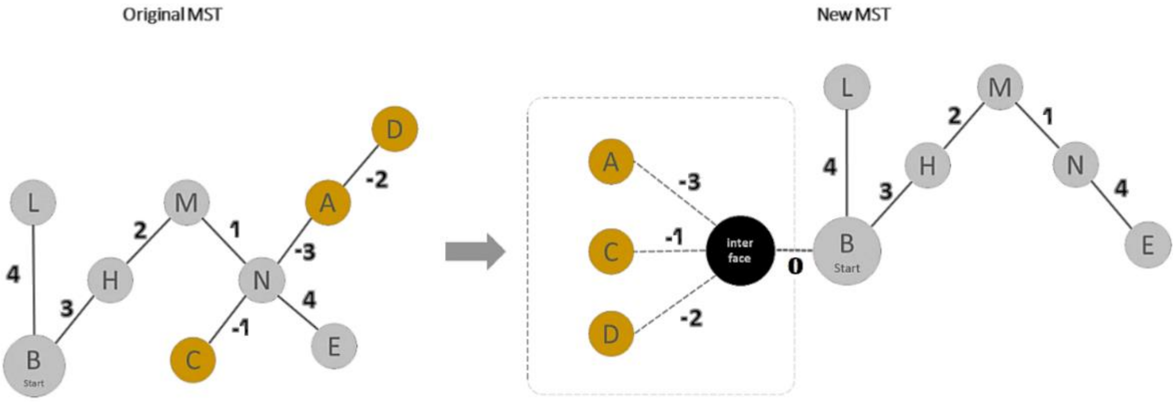


Figure 32 – Illustration of MST reorganization process

Produced by the author

The full results for these three scenarios are given in chapter **Error! Reference source not found.**.

CHAPTER 4

Results and discussion

The uncertainty associated with making decisions means that outcomes can be disrupted by a range of different factors, regardless of how exhaustive a model is. To mitigate this limitation three scenarios were proposed sharing the underlying assumption that 180,000,000 USD/year will be available for rural electrification since this is the average spending from previous national budgets (MINISTÉRIO DAS FINANÇAS, 2018, 2019, 2020, 2021). Another common element to these scenarios is the period of analysis between 2022 (next year) and 2030 which is a milestone year in SSA in terms of United Nations development goals that were agreed with Angola. The results presented in this chapter have been condensed in tables by year and graphical representations of these results were produced in QGIS to provide an overview of how these processes might evolve geographically.

4.1 Multi-period analysis and results for Scenario 1 (Grid exclusive scenario)

The first scenario (grid exclusive scenario) tackles the problem of electrification of 4,557 settlements in rural areas in Angola using solely grid extension. This scenario is a useful baseline for analysis and yearly results are given in Table 40.

Table 40 – Results from Grid exclusive scenario when minimizing NPC (2022-2030)

	Scenario 1 - Grid exclusive – yearly results				
Year	Yearly NPC	People	Avg. MPI	Nr. Settlements	Avg. LCOE (USD
2022	179,946,290	965,371	0.4816	678	0.7304
2023	179,992,198	870,705	0.5706	653	1.7698
2024	179,803,195	1,170,952	0.4994	354	1.0471
2025	180,235,757	1,399,286	0.4786	297	0.8714
2026	179,751,246	1,178,332	0.4059	527	0.2329
2027	179,981,594	1,448,579	0.3936	500	0.3379
2028	180,215,382	1,423,339	0.4286	342	0.7037
2029	179,331,544	1,756,638	0.5070	141	1.0361
2030	178,978,154	1,298,983	0.5899	360	1.9665
2022 - 2030	**1,618,235,539**	**11,512,185**	**0.4805**	**3,852**	**0.9409**

Produced by the author

In a way, this grid exclusive scenario is an extreme option in which rural communities are connected exclusively by the national grid. From Table 40, we can observe that only 67,7% people and 3852 out of 4557 settlements will be connected in 2030. Additionally, the first 3-year period (from 2022 to 2030) is the one with the highest increase on the percentage of people connected whereas the opposite occurs in the last 3-year period.

To visualize the grid growth and help decision-makers to better explore each solution and support their decisions, some maps were built. This is presented in incremental timesteps of 3 years, 2022-2024, 2025-2027 and 2028-2030 (Figure 33, Figure 34 and Figure 35, respectively).

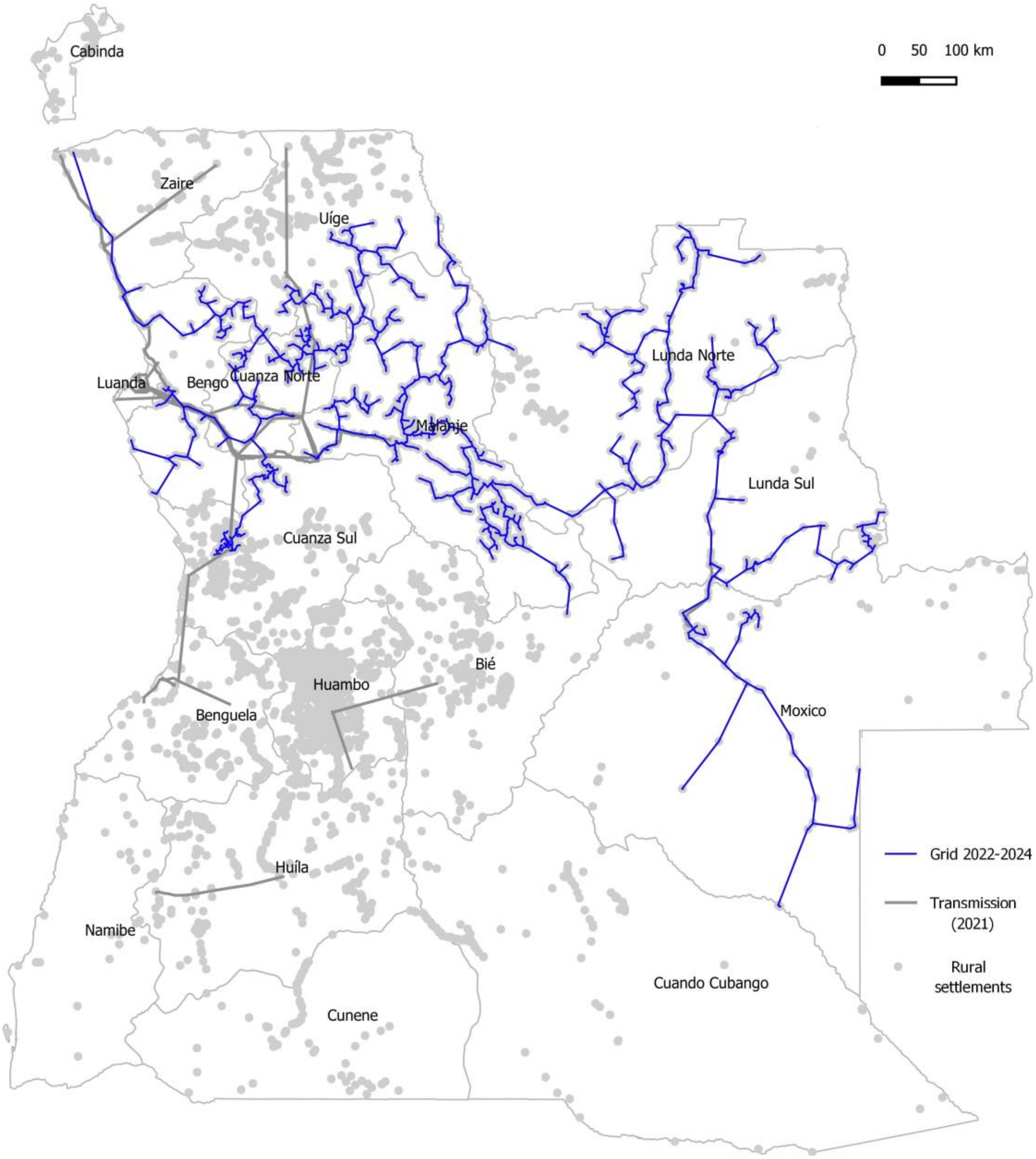


Figure 33 – Scenario 1 (grid exclusive) – grid growth (2022 – 2024)

Produced by the author

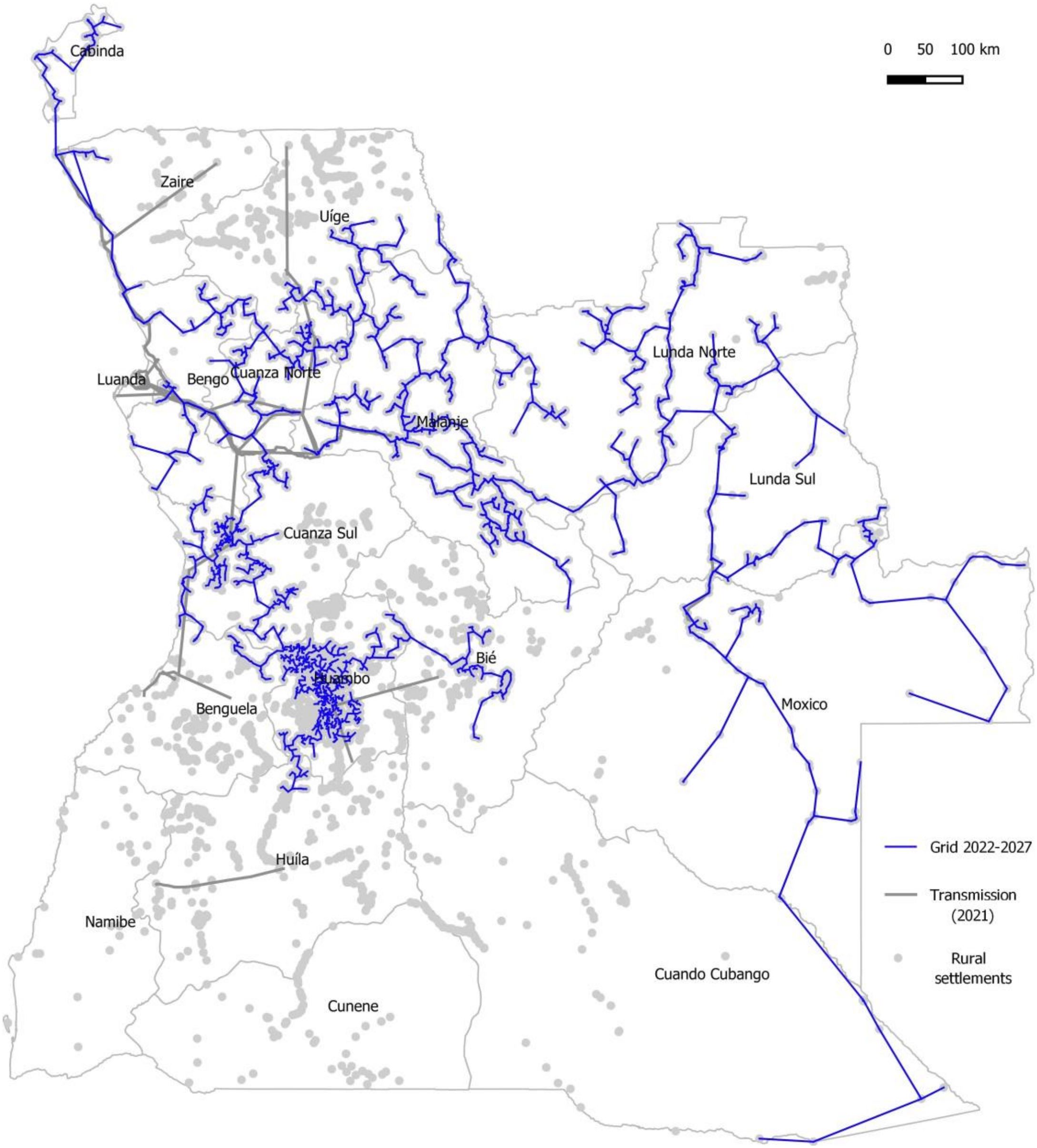


Figure 34 – Scenario 1 (grid exclusive) – grid growth (2022 – 2027)

Produced by the author

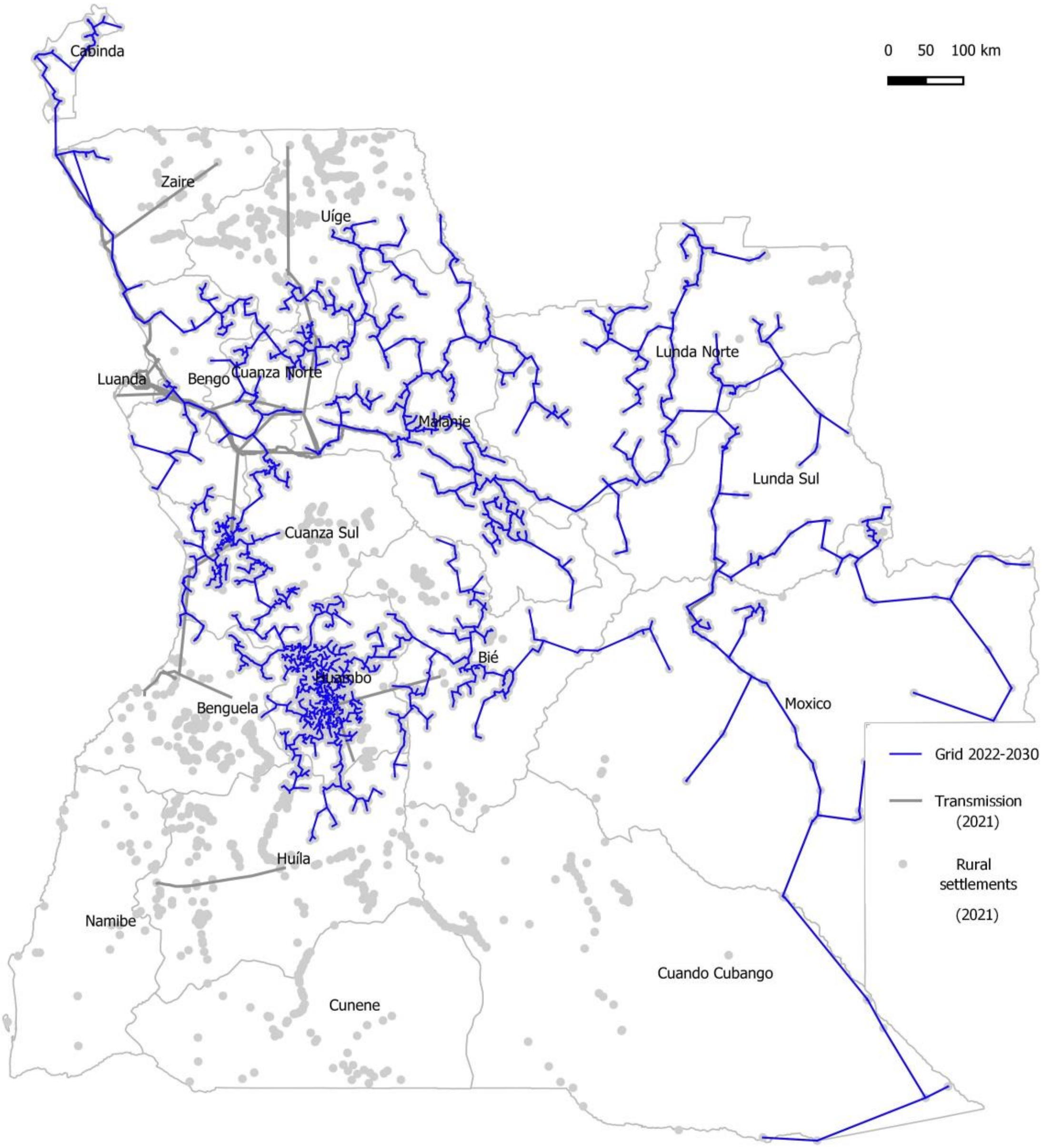


Figure 35 – Scenario 1 (grid exclusive) – grid growth (2022 – 2030)

Produced by the author

Results given in Table 40 and Figures 33, 34 and 35 show that:

- The average MPI (0.4805) remains almost the same all over the period;
- In terms of average MPI (0.4805) there is little change over the period, but poorer southern communities are not connected in this scenario, mainly in the provinces of Namibe, Huíla, Cunene, Cuando Cubango and parts of Moxico, which are precisely the five provinces with highest poverty index, with MPI values up to 0.7539 (INE - Instituto Nacional de Estatística, 2019). This scenario accentuates regional asymmetries with southern provinces having little access to electricity;

- Average LCOE is 0.9409 USD/kWh which is very high compared to other SSA countries (Moksnes et al., 2017). This happens because very small isolated communities are being connected by grid. LCOE is lower between 2026 and 2027 due to the connections in Huambo, which are highly concentrated;
- Grid growth in the first three years happens in the north of the country and only in the following years it will reach other parts of the country, in the east and south east along the border with Zambia and in the center in the provinces of Huambo, Bié and part of Huila.

If the procedure was allowed to continue until all people are connected total expenditure would be an estimated 2,245,156,759 million USD. This would take an additional 3.5 years (ending in 2034).

While this Scenario does not answer the central question of this research, it is a good baseline to test the remaining scenarios, in which hybrid mini-grids are introduced.

4.2 Multi-period analysis and results for Scenario 2 (Grid intensive scenario)

In this scenario, all the assumptions from the previous scenario are maintained and only the supply mix changes. Scenario 2 combines grid extension and hybrid mini-grids (PV with batteries and diesel generation). The mini-grids represents 80% of yearly supply while diesel accounts for 20% of total yearly supply. These mini-grids are complete hybrid substitution solutions for grid energy and include batteries to reduce diesel use. As such, they have higher cost than simpler solutions.

The results attained with NPC, LCOE, MPI and population, as decision criteria, are given in Table 41, Table 42, Table 43 and Table 44, respectively.

Table 41 – Results from Scenario 2 (Grid intensive) when minimizing NPC (2022-2030)

Year	Technology	Yearly NPC	People connected	Avg. MPI	Nr. Settlements	Avg. LCOE (USD/kWh)
	Grid	138,620,712	794,207	0.459	491	0.3001
	Hybrid	41,237,567	170,608	0.565	471	0.2832
2022	Yearly result	179,858,279	964,815	0.478	962	0.2971
	Grid	179,452,225	912,767	0.559	657	0.8045
	Hybrid	-	-	-	-	-
2023	Yearly result	179,452,225	912,767	0.559	657	0.8045
	Grid	174,047,399	1,157,769	0.488	315	0.4642
	Hybrid	6,366,936	22,454	0.503	3	0.2336
2024	Yearly result	180,414,335	1,180,223	0.488	318	0.4598
	Grid	166,780,484	1,341,666	0.472	262	0.3921
	Hybrid	12,110,454	46,196	0.505	5	0.2343
2025	Yearly result	178,890,938	1,387,862	0.473	267	0.3869
	Grid	177,075,349	1,160,540	0.404	520	0.1254
	Hybrid	4,080,883	23,679	0.603	1	0.3135
2026	Yearly result	181,156,232	1,184,219	0.408	521	0.1291
	Grid	179,808,560	1,453,103	0.393	494	0.1727
	Hybrid	-	-	-	-	-
2027	Yearly result	179,808,560	1,453,103	0.393	494	0.1727
	Grid	174,047,498	1,440,559	0.428	315	0.3672
	Hybrid	4,628,061	39,757	0.465	2	0.1910
2028	Yearly result	178,675,559	1,480,316	0.429	317	0.3624
	Grid	167,180,764	1,655,307	0.500	112	0.4908
	Hybrid	14,421,227	96,947	0.560	3	0.2379
2029	Yearly result	181,601,991	1,752,254	0.504	115	0.4768
	Grid	175,946,052	1,341,107	0.585	291	1.1251
	Hybrid	2,272,376	20,962	0.607	1	0.1507
2030	Yearly result	178,218,428	1,362,069	0.586	292	1.1101
2022 - 2030	Grid	1,532,959,04	11,257,02	0.475	3,457	0.4693
	Global hybrid	85,117,504	420,603	0.549	486	0.2511
	Global results	**1,618,076,54**	**11,677,62**	**0.477**	**3,943**	**0.4615**

Produced by the author

From Table 41, we can see that nearly 11.7 million of people (68% of the rural population) and 3,943 out of 4,557 existing settlements (86%) are connected. Moreover, hybrid solutions only contribute to connect 3,6% of people and 12.3% of the settlements. It is also worth remarking that hybrid solutions massively occur during the first year: 96.7% of the settlements and 40.5% of people. Concerning Grid, the expansion is more balanced all over the period. The average MPI is 0.4777 with slight variations resulting from the areas being connected: richer areas have lower

MPI's, and there is no strong difference between hybrid and grid solutions. On the contrary, the yearly average LCOE varies from 0.1291 in 2026 to 1,1101 in 2030, but with no apparent trend.

Table 42 – Results from Scenario 2 (Grid intensive) when minimizing LCOE (2022-2030)

Year	Technology	Yearly NPC	People connected	Avg. MPI	Nr. Settlements	Avg. LCOE (USD/kWh)
	Grid	160,135,204	658,070	0.4599	193	0.3691
	Hybrid	19,702,611	103,697	0.5871	228	0.4451
2022	Yearly result	179,837,815	761,767	0.5235	421	0.4071
	Grid	169,011,915	1,086,644	0.4369	488	0.3775
	Hybrid	8,989,590	67,627	0.4438	17	0.3086
2023	Yearly result	178,001,505	1,154,271	0.4403	505	0.3431
	Grid	173,816,978	1,214,577	0.4708	490	0.5290
	Hybrid	8,312,486	39,220	0.5129	32	0.3512
2024	Yearly result	182,129,464	1,253,797	0.4919	522	0.4401
	Grid	160,060,850	1,355,805	0.4667	348	0.4524
	Hybrid	10,292,501	72,888	0.4965	34	0.3385
2025	Yearly result	170,353,351	1,428,693	0.4816	382	0.3954
	Grid	186,976,896	1,598,954	0.4199	471	0.1619
	Hybrid	2,275,954	12,210	0.4726	5	0.2797
2026	Yearly result	189,252,850	1,611,164	0.4462	476	0.2208
	Grid	174,176,826	1,431,153	0.3658	496	0.1691
	Hybrid	3,783,450	24,008	0.4407	15	0.2768
2027	Yearly result	177,960,276	1,455,161	0.4032	511	0.2230
	Grid	169,428,733	1,558,018	0.4400	195	0.3349
	Hybrid	10,070,301	86,465	0.4655	15	0.2759
2028	Yearly result	179,499,034	1,644,483	0.4528	210	0.3054
	Grid	168,689,160	1,599,894	0.4898	219	0.5038
	Hybrid	11,163,148	99,998	0.4925	17	0.3117
2029	Yearly result	179,852,308	1,699,892	0.4912	236	0.4078
	Grid	172,255,336	1,383,917	0.5768	361	1.3095
	Hybrid	10,720,736	51,311	0.5905	27	0.4241
2030	Yearly result	182,976,072	1,435,228	0.5836	388	0.8669
2022 - 2030	Grid	1,534,551,89	11,887,03	0.4580	3,261	0.4669
	Hybrid	85,310,777	557,424	0.5083	390	0.3450
	Global results	**1,619,862,67**	**12,444,45**	**0.4603**	**3,651**	**0.3691**

Produced by the author

An analysis of the results given in Table 42 shows that nearly 12.4 million of people (73% of the rural population) and 3,651 settlements (79.7%) are connected. The percentage of people connected by hybrid solutions is 4.5% corresponding to 8.5% of the settlements, with 58.5% of them occurring during the first year. The average MPI is 0.4603 and the yearly average LCOE varies from 0.2208 in 2026 to 0.8669 in 2030, with a mean value of 0.3691.

Table 43 – Results from Scenario 2 (Grid intensive) when maximizing MPI (2022-2030)

Year	Technology	Yearly NPC	People connected	Avg. MPI	Nr. Settlements	Avg. LCOE (USD/kWh)
	Grid	173,023,508	774,752	0.405	375	0.3218
	Hybrid	6,732,864	35,525	0.437	28	0.2634
2022	Yearly result	179,756,372	810,277	0.421	403	0.2926
	Grid	167,696,654	1,296,457	0.454	310	0.3580
	Hybrid	11,861,085	98,056	0.523	43	0.3578
2023	Yearly result	179,557,739	1,394,513	0.488	353	0.3579
	Grid	172,472,330	1,483,277	0.448	329	0.3446
	Hybrid	8,008,854	41,469	0.524	57	0.3921
2024	Yearly result	180,481,184	1,524,746	0.486	386	0.3684
	Grid	178,853,966	1,411,381	0.349	655	0.1268
	Hybrid	911,802	1,569	0.367	5	0.2546
2025	Yearly result	179,765,768	1,412,950	0.358	660	0.1907
	Grid	169,082,323	2,094,581	0.468	312	0.3117
	Hybrid	10,643,133	104,511	0.451	24	0.2599
2026	Yearly result	179,725,456	2,199,092	0.460	336	0.2858
	Grid	173,284,088	1,006,509	0.445	180	0.2877
	Hybrid	7,369,001	42,210	0.504	12	0.2628
2027	Yearly result	180,653,089	1,048,719	0.475	192	0.2752
	Grid	175,909,500	1,455,100	0.542	188	0.9636
	Hybrid	3,954,387	15,080	0.560	16	0.4106
2028	Yearly result	179,863,887	1,470,180	0.551	204	0.6871
	Grid	170,219,959	1,362,534	0.498	411	0.6278
	Hybrid	9,546,681	58,737	0.548	54	0.3625
2029	Yearly result	179,766,640	1,421,271	0.523	465	0.4952
	Grid	158,606,388	1,228,552	0.518	408	0.8525
	Hybrid	21,813,524	144,303	0.600	115	0.4251
2030	Yearly result	180,419,912	1,372,855	0.559	523	0.6387
2022 - 2030	Grid	1,539,148,71	12,113,14	0.461	3,168	0.4660
	Hybrid	80,841,331	541,460	0.526	354	0.3470
	Global results	**1,619,990,04**	**12,654,60**	**0.464**	**3,522**	**0.4610**

Produced by the author

In Table 43 we can observe that nearly 12.7 million of people (74% of the rural population) and 3,522 settlements (77%) are connected. Only 4.3% of people are connected via hybrid solutions corresponding to 7.7% of the settlements, with 32.5% of them occurring during the last year. The average MPI is 0.4645 with the three highest values occurring during the last 3-year period. The yearly average LCOE varies from 0.1907 in 2025 to 0.6871 in 2028, and the average value is 0.4610.

Table 44 – Results from Scenario 2 (Grid intensive) when prioritizing more populated settlements (2022-2030)

Year	Technology	Yearly NPC	People connected	Avg. MPI	Nr. Settlements	Avg. LCOE (USD/kWh)
	Grid	160,112,421	633,551	0.516	304	0.8921
	Hybrid	19,534,535	102,901	0.587	223	0.3261
2022	Yearly result	179,646,956	736,452	0.552	527	0.6091
	Grid	171,576,036	1,505,319	0.451	314	0.1442
	Hybrid	8,548,216	68,880	0.479	19	0.2551
2023	Yearly result	180,124,252	1,574,199	0.465	333	0.1997
	Grid	178,935,691	1,487,070	0.392	529	0.1009
	Hybrid	381,690	168	0.511	1	0.2070
2024	Yearly result	179,317,381	1,487,238	0.451	530	0.1540
	Grid	174,009,936	1,464,857	0.371	402	0.1584
	Hybrid	6,255,830	51,790	0.412	15	0.2032
2025	Yearly result	180,265,766	1,516,647	0.392	417	0.1808
	Grid	170,344,279	1,564,007	0.490	236	0.4490
	Hybrid	10,211,093	66,005	0.464	20	0.2021
2026	Yearly result	180,555,372	1,630,012	0.477	256	0.3255
	Grid	166,580,898	1,544,048	0.577	178	0.8272
	Hybrid	11,889,938	93,928	0.555	18	0.2769
2027	Yearly result	178,470,836	1,637,976	0.566	196	0.5521
	Grid	171,385,464	1,701,906	0.585	185	1.0662
	Hybrid	9,500,673	73,799	0.625	23	0.3333
2028	Yearly result	180,886,137	1,775,705	0.605	208	0.6997
	Grid	165,345,360	1,165,949	0.448	289	0.4232
	Hybrid	10,806,464	51,652	0.470	18	0.2264
2029	Yearly result	176,151,824	1,217,601	0.459	307	0.3248
	Grid	177,331,183	1,283,164	0.460	417	0.2746
	Hybrid	7,148,882	37,876	0.510	28	0.2642
2030	Yearly result	184,480,065	1,321,040	0.485	445	0.2694
2022 - 2030	Grid	1,535,621,26	12,349,87	0.478	2,854	0.4700
	Hybrid	84,277,321	546,999	0.525	365	0.2693
	Global results	**1,619,898,58**	**12,896,87**	**0.480**	**3,219**	**0.3588**

Produced by the author

Regarding the results given in Table 44 nearly 12.9 million of people (76% of the rural population) and 3,219 settlements (70%) are connected. Grid connections accounts for the large majority of connections, with only 4.2% of people and 7.9% of the settlements being connected through hybrid solutions. Moreover, the number of settlements connected via hybrid connections occur mainly during the first year (62.5%) even representing only 23% of the people connected. The average MPI is 0.4801 being 2028 the year that has the highest yearly value (0.6052). The yearly average LCOE varies from 0.1540 in 2024 to 0.6997 in 2028.

For the final scenario (scenario 3 with the same variations) results are presented along with maps on the evolution of the system at three-year intervals. These results are given in Subchapter 4.3.

4.3 Multi-period Results and outputs from Scenario 3 (Balanced scenario)

Regarding the last scenario (scenario 3 – balanced), batteries are removed from the hybrid solution and the role of diesel is expanded representing 50% of demand supplied, with PV supplying the remaining 50%. Diesel participation takes into account that daytime in Angola ranges from 11.5 hours to 12.5 hours[41] and grid remains included in the mix.

The results for this scenario using NPC minimization are given in Table 45 as well as Figure 36, Figure 37 and Figure 38.

---

[41] https://www.worlddata.info/africa/angola/sunset.php

Table 45 – Results from Balanced scenario when minimizing NPC (2022-2030)

Year	Technology	Yearly NPC	People connected	Avg. MPI	Nr. Settlements	Avg. LCOE (USD/kWh)
	Grid	240,640	2,854	0.3060	2	0.0410
	Hybrid (PV+diesel)	179,561,790	1,942,684	0.5145	1,833	0.0745
2022	Yearly result	179,802,430	1,945,538	0.5141	1,835	0.0745
	Grid	70,079,363	366,525	0.4532	267	0.2437
	Hybrid (PV+diesel)	109,900,964	1,307,854	0.5179	175	0.0770
2023	Yearly result	179,980,327	1,674,379	0.5038	442	0.1135
	Grid	136,320,828	753,324	0.5353	443	0.5391
	Hybrid (PV+diesel)	42,911,809	509,374	0.5253	28	0.0746
2024	Yearly result	179,232,637	1,262,698	0.5312	471	0.3517
	Grid	150,660,201	1,048,573	0.4608	253	0.2104
	Hybrid (PV+diesel)	29,951,468	342,983	0.5630	13	0.0738
2025	Yearly result	180,611,669	1,391,556	0.4860	266	0.1767
	Grid	128,891,621	1,054,664	0.3855	162	0.2638
	Hybrid (PV+diesel)	49,140,323	614,040	0.5415	15	0.0745
2026	Yearly result	178,031,944	1,668,704	0.4429	177	0.1941
	Grid	182,178,256	1,461,329	0.3741	523	0.1034
	Hybrid (PV+diesel)	-	-	-	-	-
2027	Yearly result	182,178,256	1,461,329	0.3741	523	0.1034
	Grid	152,930,494	1,398,167	0.4314	193	0.0410
	Hybrid (PV+diesel)	23,756,250	244,987	0.4990	5	0.0745
2028	Yearly result	176,686,744	1,643,154	0.4415	198	0.0460
	Grid	155,271,073	1,339,963	0.5552	197	0.5933
	Hybrid (PV+diesel)	27,079,024	276,977	0.5094	5	0.0568
2029	Yearly result	182,350,097	1,616,940	0.5473	202	0.5014
	Grid	178,154,556	1,591,790	0.5126	219	0.5245
	Hybrid (PV+diesel)	-	-	-	-	-
2030	Yearly result	181,050,491	1,591,790	0.5126	219	0.5245
2022 - 2030 period	Grid	1,154,727,032	9,017,189	0.4624	2,259	0.3141
	Hybrid (PV+diesel)	462,301,628	5,238,899	0.5217	2,074	0.0742
	Global results	**1,619,924,595**	**14,256,088**	**0.4842**	**4,333**	**0.2260**

Produced by the author

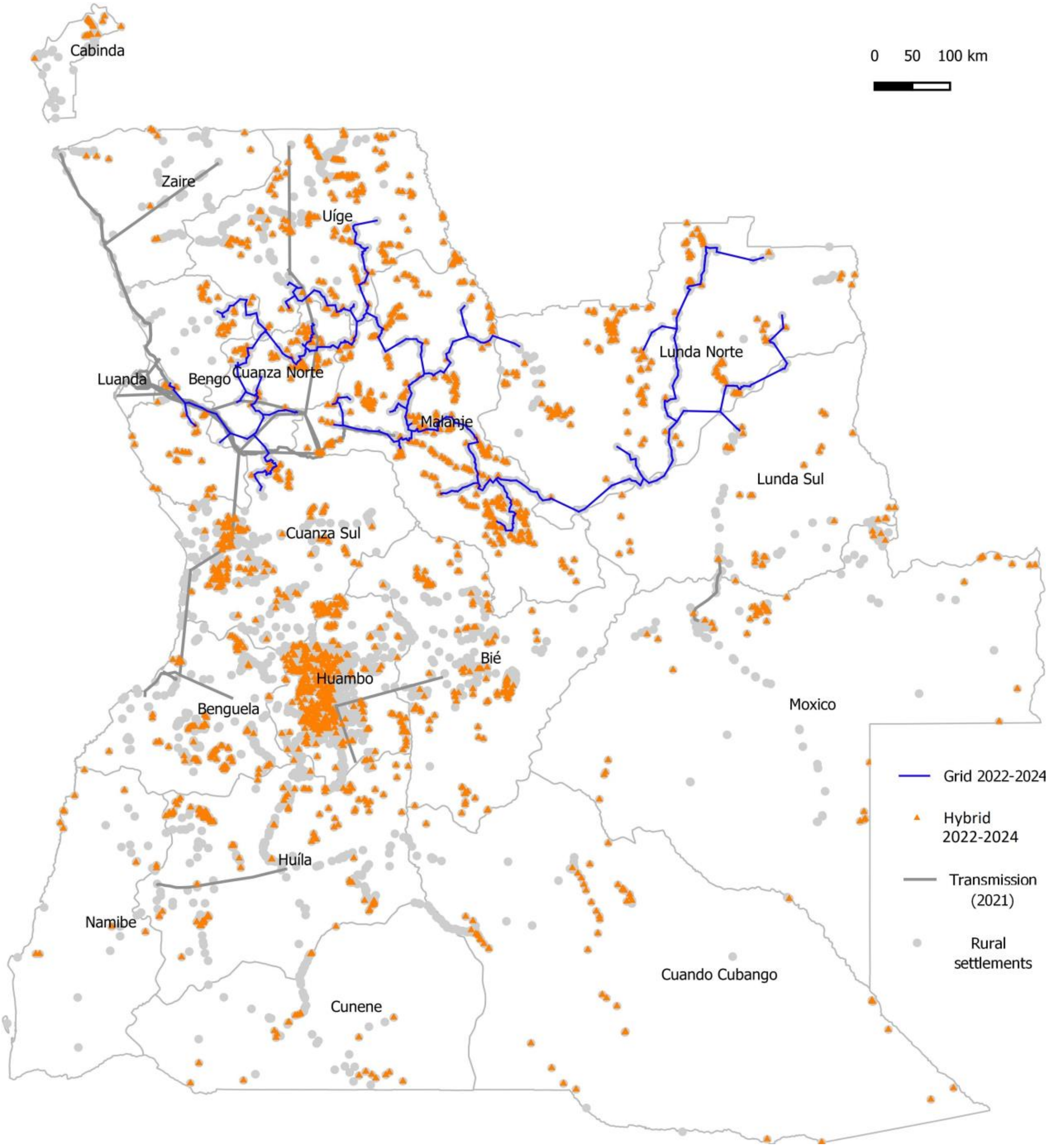


Figure 36 – Scenario 3 (Balanced) – grid and hybrid growth – minimizing NPC (2022 – 2024)

Produced by the author

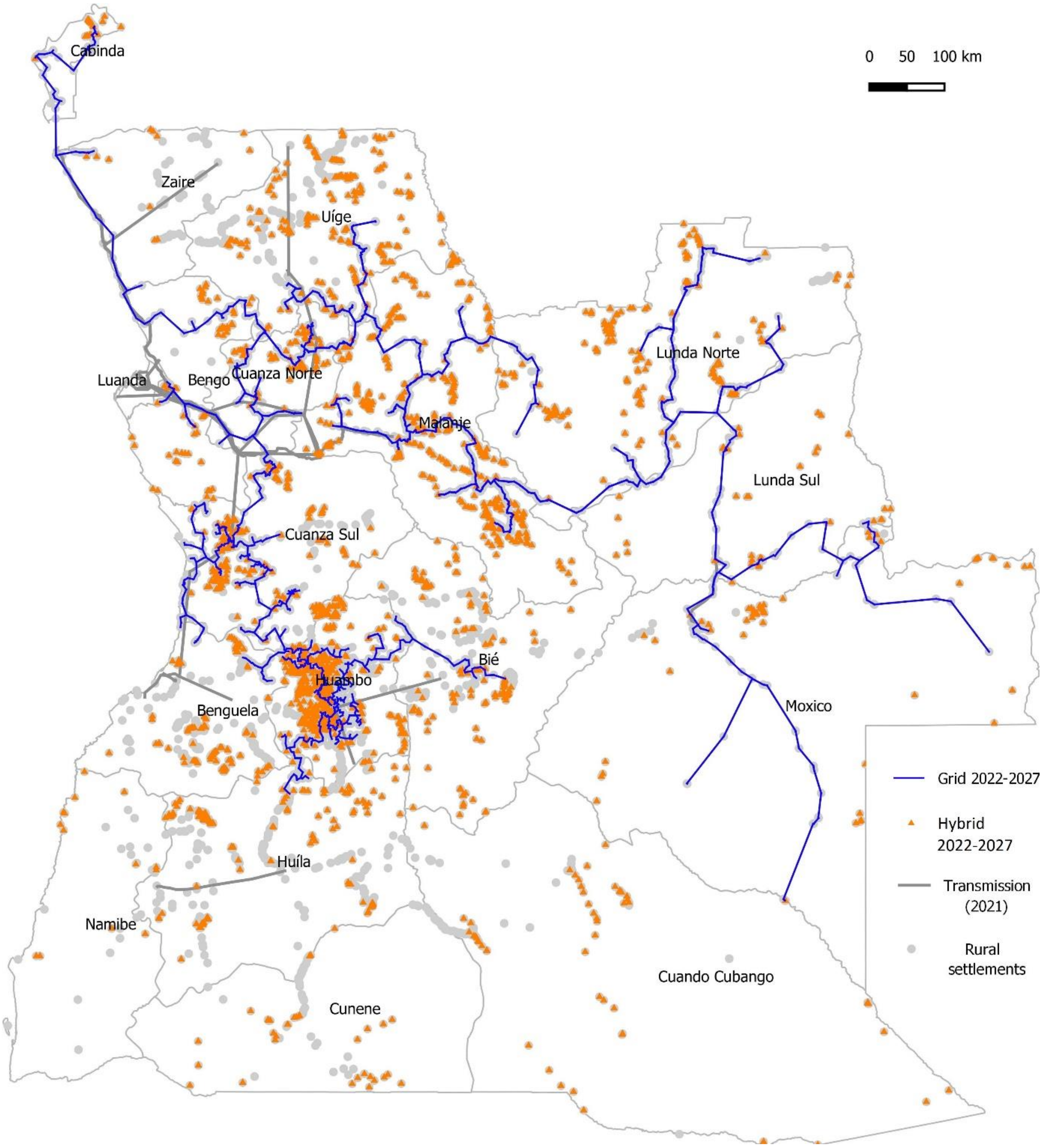


Figure 37 – Scenario 3 (Balanced) – grid and hybrid growth – minimizing NPC (2022 – 2027)

Produced by the author

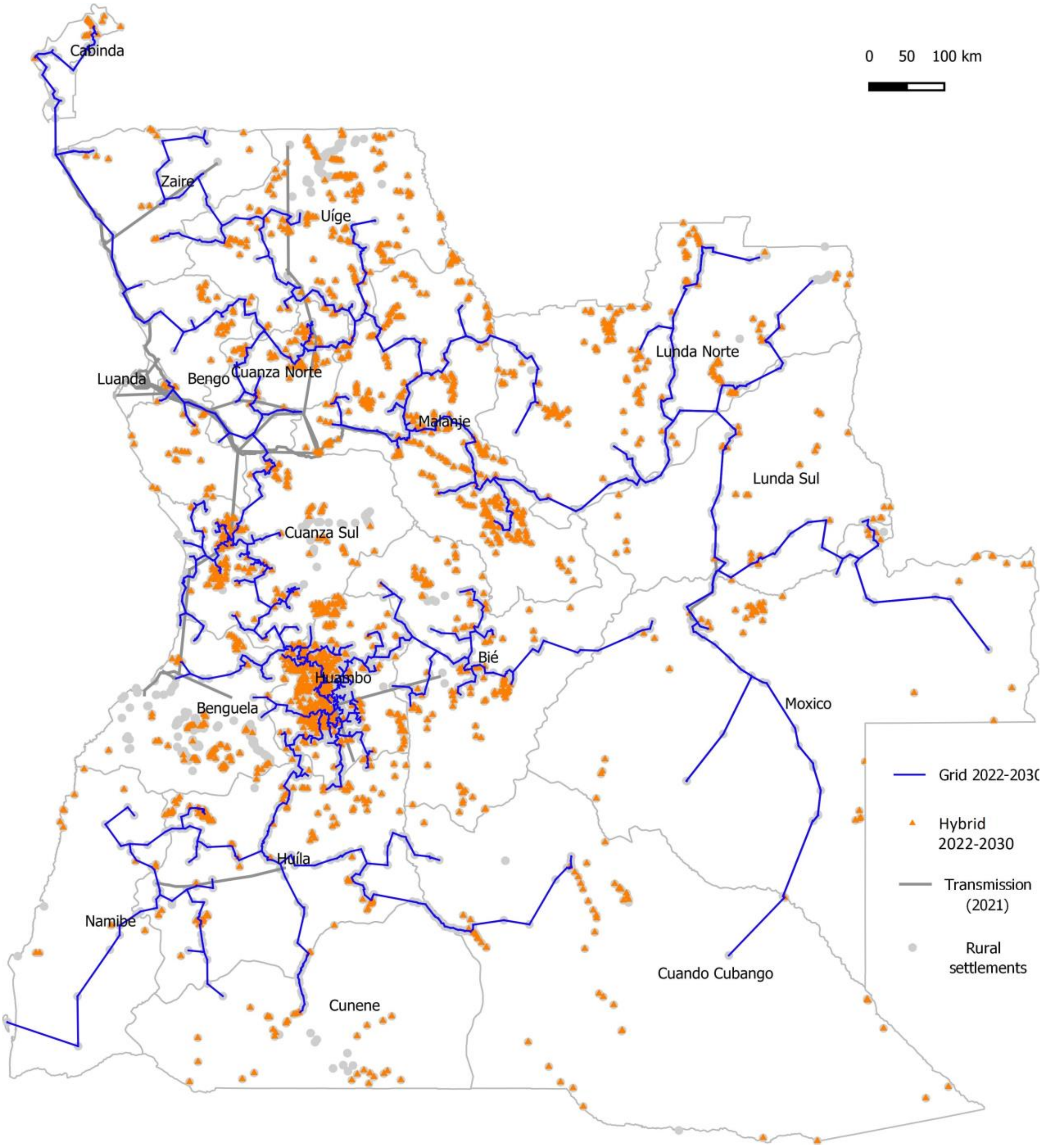


Figure 38 – Scenario 3 (Balanced) – grid and hybrid growth – minimizing NPC (2022 – 2030)

Produced by the author

Regarding the results given in Table 45, Figure 36, Figure 37 and Figure 38 nearly 14.3 million of people (84% of the rural population) and 4,333 settlements (95%) are connected, the large majority in the first year. Notwithstanding these results, grid connections still account for the large majority of connections (95.8%). Moreover, the number of settlements connected via hybrid connections occur mainly during the first year (88%). The average MPI is 0.4842 with 2029 being the year that has the highest yearly value (0.5473). The yearly average LCOE is 0.2260 ranging from 0.0460 in 2028 to 0.5245 in 2030. This implies more hybrid connections increase the overall sustainability of the system, lowering LCOE.

The results for this scenario using LCOE minimization are given in Table 46 as well as Figure 39, Figure 40 and Figure 41.

Table 46 – Results from Balanced scenario when minimizing LCOE (2022-2030)

Year	Technology	Yearly NPC	People connected	Avg. MPI	Nr. Settlements	Avg. LCOE (USD/kWh)
	Grid	1,732,687	6,832	0.2290	7	0.0287
	Hybrid (PV+diesel)	178,242,008	1,325,691	0.4149	159	0.0384
2022	Yearly result	179,974,695	1,332,523	0.4139	166	0.0383
	Grid	22,702,122	192,470	0.1792	12	0.0168
	Hybrid (PV+diesel)	157,295,359	1,578,542	0.4415	567	0.0624
2023	Yearly result	179,997,481	1,771,012	0.4130	579	0.0574
	Grid	22,788,398	152,176	0.3275	38	0.0444
	Hybrid (PV+diesel)	157,065,994	2,341,465	0.5210	434	0.1508
2024	Yearly result	179,854,393	2,493,641	0.5092	472	0.1443
	Grid	142,938,753	995,693	0.4196	304	0.1305
	Hybrid (PV+diesel)	37,172,917	530,118	0.5630	623	0.4223
2025	Yearly result	180,111,669	1,525,811	0.4694	927	0.2319
	Grid	176,433,698	1,299,912	0.4574	432	0.2083
	Hybrid (PV+diesel)	3,169,294	43,761	0.6017	206	1.0708
2026	Yearly result	179,602,992	1,343,673	0.4621	638	0.2364
	Grid	180,081,088	1,510,280	0.4070	370	0.0911
	Hybrid (PV+diesel)	-	-	-	-	-
2027	Yearly result	180,081,088	1,510,280	0.4070	370	0.0911
	Grid	179,436,522	1,575,553	0.3310	324	0.0793
	Hybrid (PV+diesel)	-	-	0.0000	-	-
2028	Yearly result	179,436,522	1,575,553	0.3310	324	0.0793
	Grid	180,418,693	1,628,299	0.4714	209	0.2097
	Hybrid (PV+diesel)	-	-	-	-	-
2029	Yearly result	180,418,693	1,628,299	0.4714	209	0.2097
	Grid	179,546,632	1,395,663	0.5916	397	0.6484
	Hybrid (PV+diesel)	818,733	11,256	0.6059	50	0.212
2030	Yearly result	180,365,365	1,406,919	0.5917	447	0.6450
2022 - 2030	Grid	1,086,078,592	8,756,878	0.4371	2,093	0.2353
	Hybrid (PV+diesel)	533,764,306	5,830,833	0.4799	2,039	0.2949
	Global results	**1,619,842,898**	**14,587,711**	**0.4542**	**4,132**	**0.2647**

Produced by the author

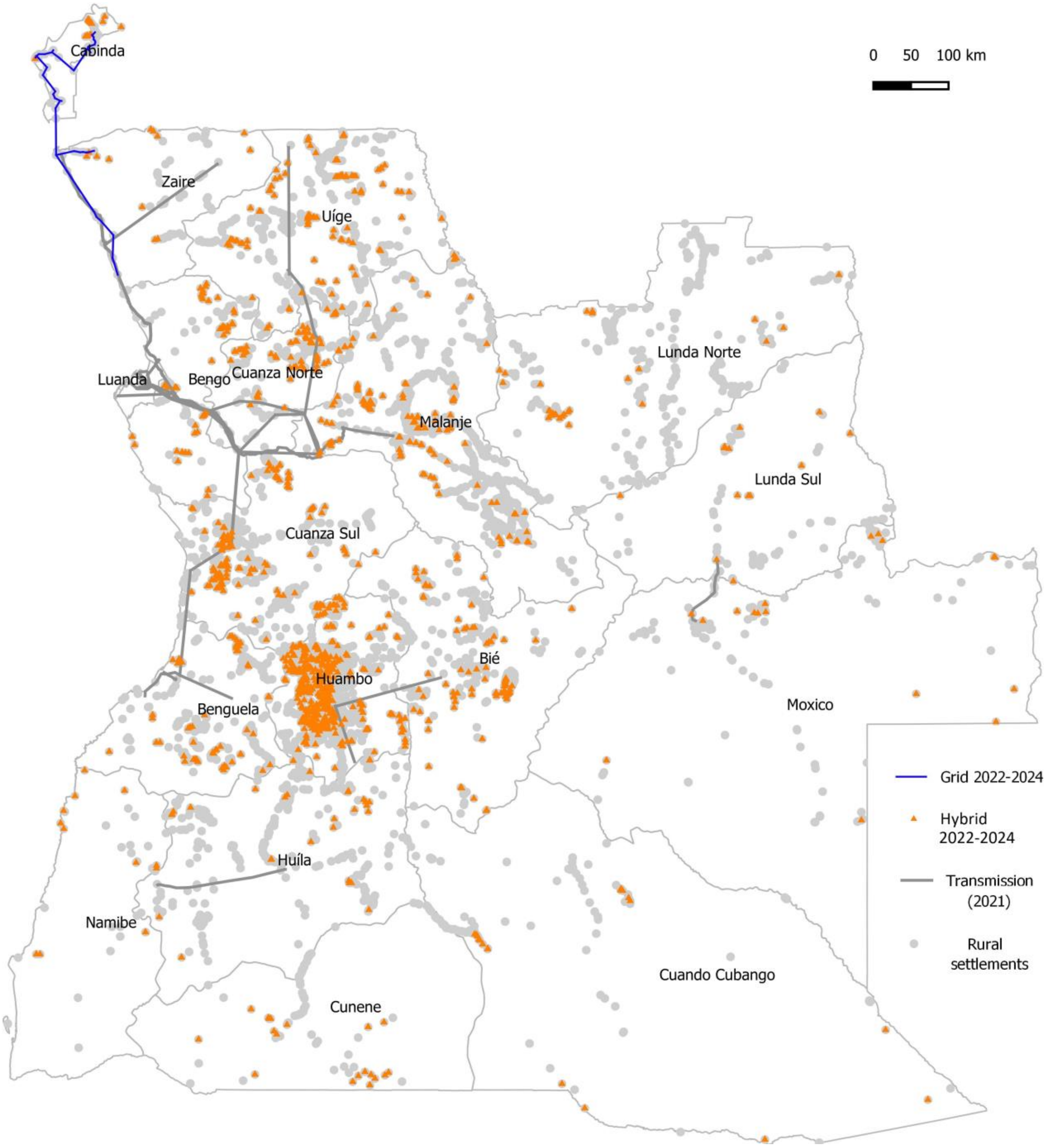


Figure 39 – Scenario 3 (Balanced scenario) – Expected grid and hybrid growth – minimizing LCOE (2022 – 2024)

Produced by the author

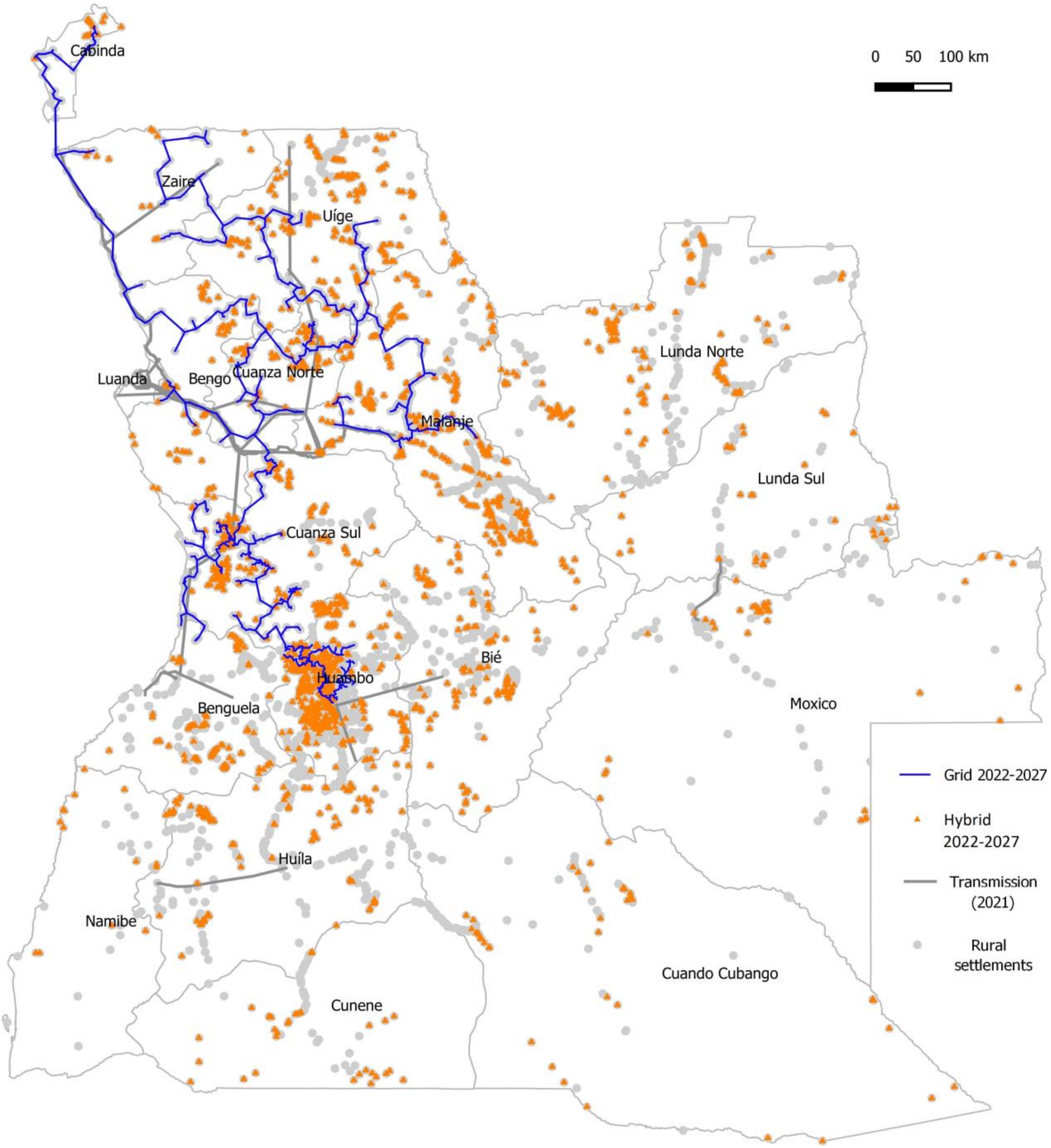


Figure 40 – Scenario 3 (Balanced scenario) – Expected grid and hybrid growth – minimizing LCOE (2022 – 2027)

Produced by the author

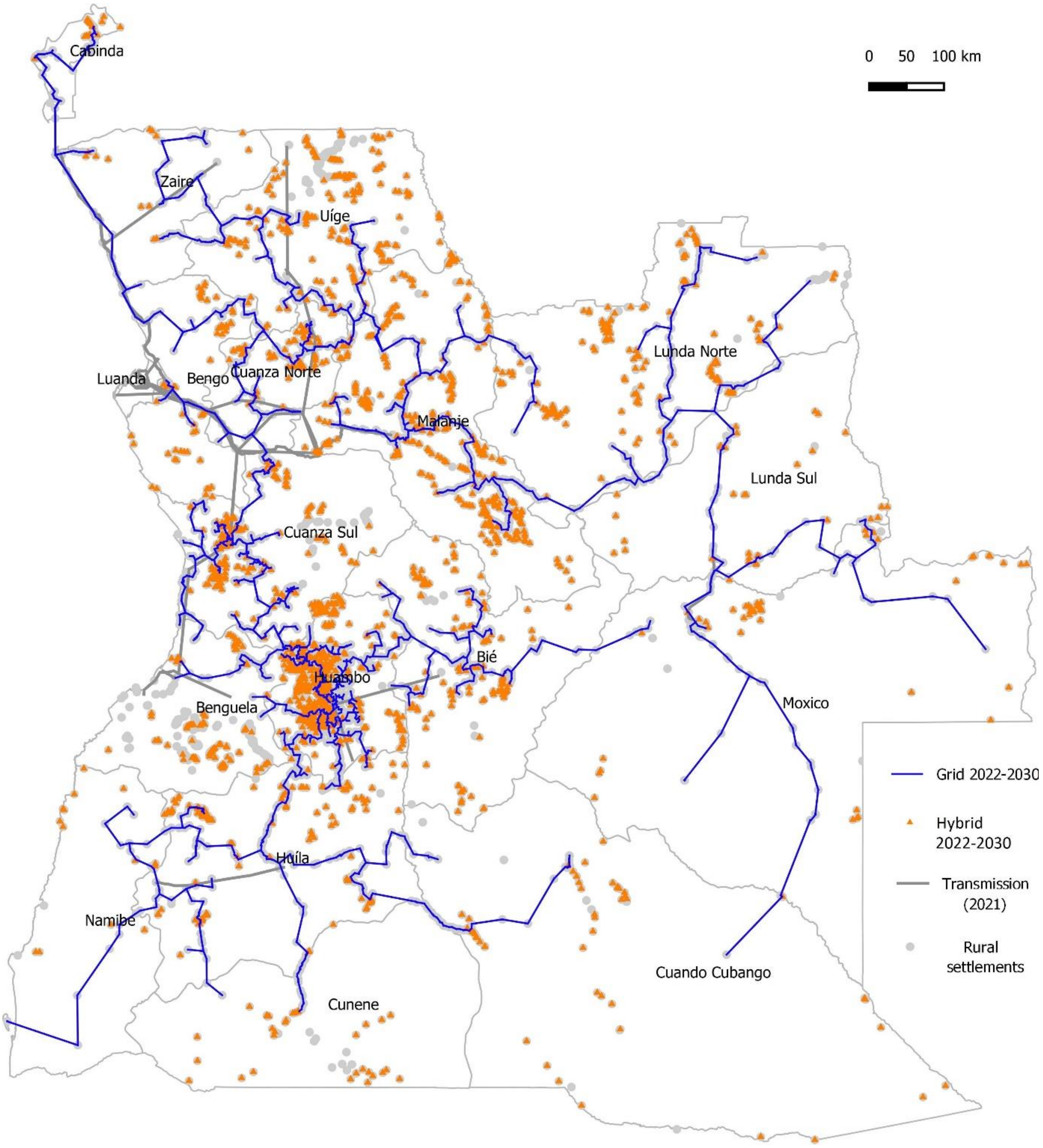


Figure 41 – Scenario 3 (Balanced scenario) – Expected grid and hybrid growth – minimizing LCOE (2022 – 2030)

Produced by the author

Some conclusions from Table 46 as well as Figure 39, Figure 40 and Figure 41 are that nearly 14.6 million of people (86% of the rural population) and 4,132 settlements (91%) are connected. While there is a similar number of connected settlements by the two technologies and the rate of connections is evenly spread over time, only 66% people are connected by grid extension. The average MPI is 0.4542 with 2027, a year where only grid connections happen, having the lowest yearly value (0.4070). The yearly average LCOE is 0.2647 ranging from 0.0383 in 2022 to 0.6450 in 2030.

The results for this scenario using MPI maximization are given in Table 47. In map format the results are given in Figure 42, Figure 43 and Figure 44.

Table 47 – Results from Balanced scenario when maximizing MPI (2022-2030)

Year	Technology	Yearly NPC	People connected	Avg. MPI	Nr. Settlements	Avg. LCOE (USD/kWh)
	Grid	240,640	976	0.2290	1	0.0281
	Hybrid (PV+diesel)	175,517,239	2,350,430	0.6233	665	0.7573
2022	Yearly Global results	175,757,879	2,351,406	0.6232	666	0.7570
	Grid	-	-	-	-	-
	Hybrid (PV+diesel)	183,834,800	2,012,557	0.6012	662	0.3912
2023	Yearly Global results	183,834,800	2,012,557	0.6012	662	0.3912
	Grid	16,696,688	92,674	0.4368	17	0.0591
	Hybrid (PV+diesel)	163,302,289	1,411,678	0.4891	717	0.1364
2024	Yearly Global results	179,998,977	1,504,352	0.4859	734	0.1316
	Grid	179,672,215	979,430	0.4875	556	0.4740
	Hybrid (PV+diesel)	610,811	2,543	0.2141	5	0.0471
2025	Yearly Global results	180,283,026	981,973	0.4868	561	0.4729
	Grid	179,812,926	1,175,816	0.5404	259	0.4458
	Hybrid (PV+diesel)	37,165	162	0.2100	1	0.0640
2026	Yearly Global results	179,850,091	1,175,978	0.5404	260	0.4457
	Grid	166,654,902	1,268,377	0.4995	350	0.1529
	Hybrid (PV+diesel)	10,777,069	59,860	0.4373	23	0.0584
2027	Yearly Global results	177,431,971	1,328,237	0.4967	373	0.1487
	Grid	182,646,135	1,625,680	0.6030	295	0.0863
	Hybrid (PV+diesel)	-	-	-	-	-
2028	Yearly Global results	182,646,135	1,625,680	0.6030	295	0.0863
	Grid	179,541,004	1,649,281	0.3858	305	0.2432
	Hybrid (PV+diesel)	-	-	-	-	-
2029	Yearly Global results	179,541,004	1,649,281	0.3858	305	0
	Grid	180,306,080	1,659,366	0.2720	237	0.7283
	Hybrid (PV+diesel)	-	-	-	-	-
2030	Yearly Global results	180,306,080	1,659,366	0.2720	237	0.7283
Period Global	Grid	1,085,570,589	8,451,600	0.4561	2,020	0.3494
	Hybrid (PV+diesel)	534,079,373	5,837,230	0.5811	2,073	0.4158
	Total	**1,619,649,962**	**14,288,830**	**0.5072**	**4,093**	**0.3830**

Produced by the author

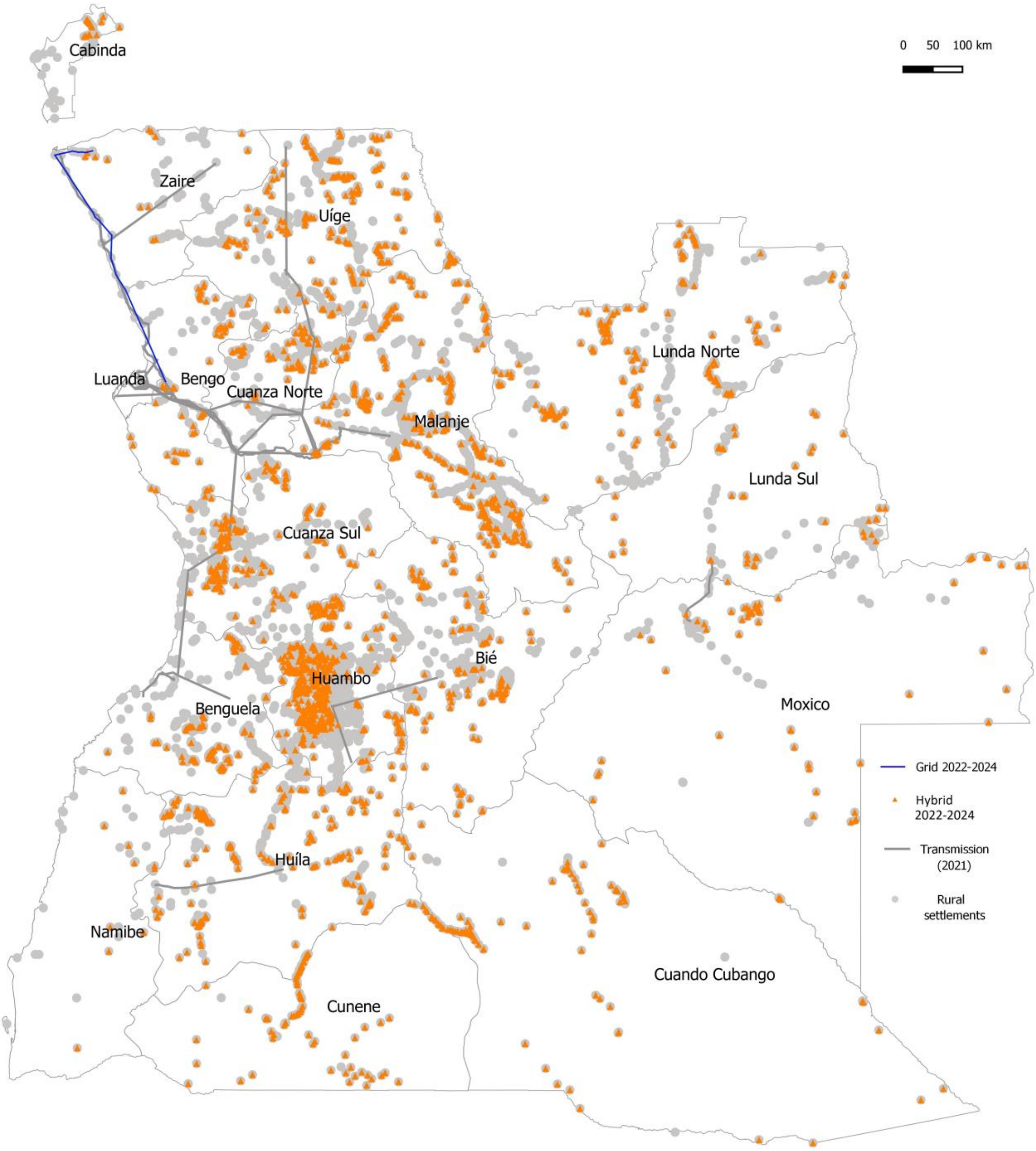


Figure 42 – Scenario 3 (Balanced scenario) – Expected grid and hybrid growth – maximizing MPI (2022 – 2024)

Produced by the author

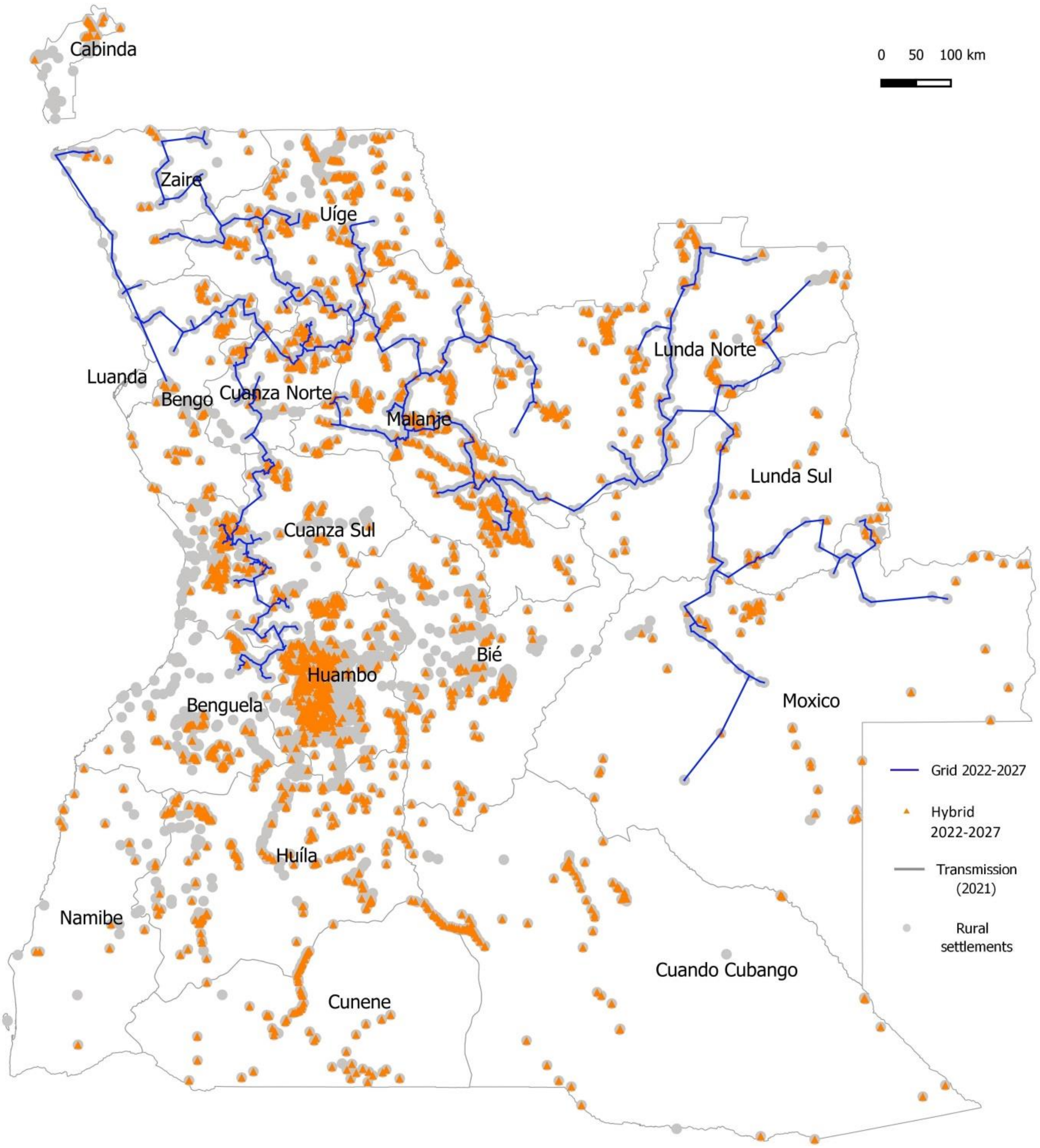


Figure 43 – Scenario 3 (Balanced scenario) – Expected grid and hybrid growth – maximizing MPI (2022 – 2027)

Produced by the author

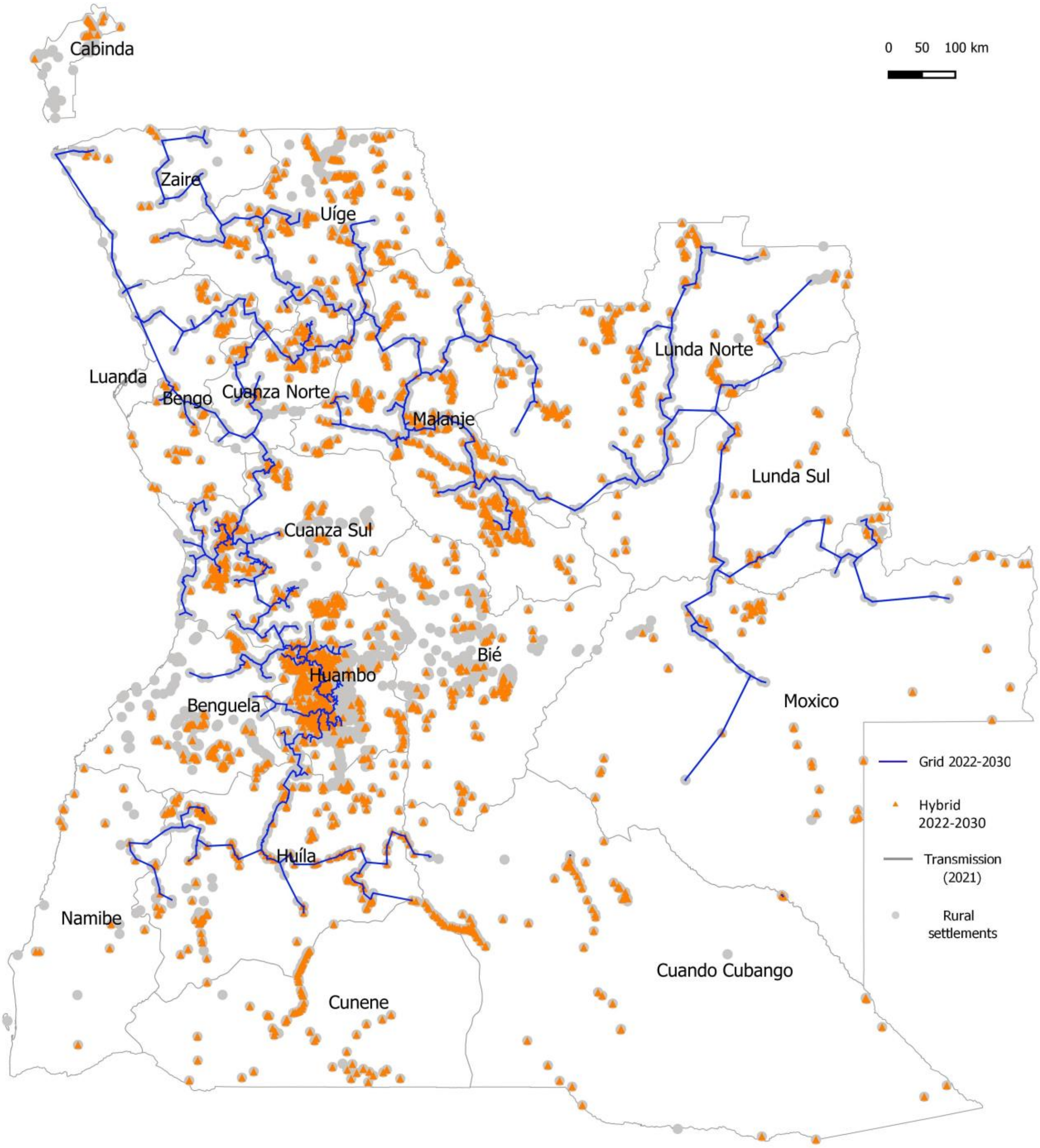


Figure 44 – Scenario 3 (Balanced scenario) – Expected grid and hybrid growth – maximizing MPI (2022 – 2030)

Produced by the author

From Table 47 as well as Figure 42, Figure 43 and Figure 44, it can be seen that nearly 14.3 million of people (84% of the rural population) and 4,093 settlements (90%) are connected. In the first three years the majority of connections happen with hybrid mini-grids. There is a similar global level of connected settlements although 66% people are connected by grid extension. The average

MPI is 0.5072 with a focus on poorer settlements (higher MPI) while average LCOE is 0.3830 USD/kW. Grid growth started in the northwest and spreads to the northeast and to center and southern areas along the coastline.

Finally, results for this scenario when prioritizing the connection of settlements with the highest population size are given in Table 48. Map outputs are given in Figure 45, Figure 46 and Figure 47.

Table 48 – Results from Balanced scenario when prioritizing more populated settlements (2022-2030)

Year	Technology	Yearly NPC	People connected	Avg. MPI	Nr. Settlements	Avg. LCOE (USD/kWh)
	Grid	2,113,310	9,686	0.2411	8	0.0310
	Hybrid (PV+diesel)	177,447,798	1,817,723	0.4824	1,817	0.4573
2022	Yearly result	179,561,108	1,827,409	0.4798	1,825	0.455
	Grid	26,510,777	155,786	0.3178	21	0.0888
	Hybrid (PV+diesel)	153,000,661	1,718,147	0.5033	212	0.1278
2023	Yearly result	179,511,438	1,873,933	0.48	233	0.1245
	Grid	180,393,910	901,111	0.4538	655	0.4521
	Hybrid (PV+diesel)	-	-	-	-	-
2024	Yearly result	180,393,910	901,111	0.4538	655	0.4521
	Grid	156,403,136	1,243,183	0.3507	257	0.1053
	Hybrid (PV+diesel)	22,279,163	190,144	0.4952	9	0.0771
2025	Yearly result	178,682,300	1,433,327	0.3649	266	0.1016
	Grid	133,983,924	878,065	0.436	183	0.4068
	Hybrid (PV+diesel)	47,084,324	584,346	0.5774	20	0.1059
2026	Yearly result	181,068,247	1,462,411	0.4833	203	0.2866
	Grid	163,300,256	1,238,741	0.2964	467	0.0965
	Hybrid (PV+diesel)	16,458,104	198,662	0.4884	5	0.0747
2027	Yearly result	179,758,360	1,437,403	0.3135	472	0.0935
	Grid	175,735,782	1,538,344	0.3926	247	0.2284
	Hybrid (PV+diesel)	3,698,351	44,950	0.496	1	0.0526
2028	Yearly result	179,434,132	1,583,294	0.3949	248	0.2234
	Grid	139,018,361	1,235,244	0.5021	145	0.6094
	Hybrid (PV+diesel)	42,517,348	550,682	0.5659	9	0.0978
2029	Yearly result	181,535,708	1,785,926	0.5202	154	0.4516
	Grid	178,694,615	1,689,536	0.5617	182	0.6953
	Hybrid (PV+diesel)	-	-	-	-	-
2030	Yearly result	178,694,615	1,689,536	0.5617	182	0.6953
2022 - 2030	Grid	1,156,154,071	8,889,696	0.4108	2,165	0.3308
	Hybrid (PV+diesel)	462,485,748	5,104,654	0.508	2,073	0.4155
	Global results	**1,618,639,819**	**13,994,350**	**0.4416**	**4,238**	**0.3722**

Produced by the author

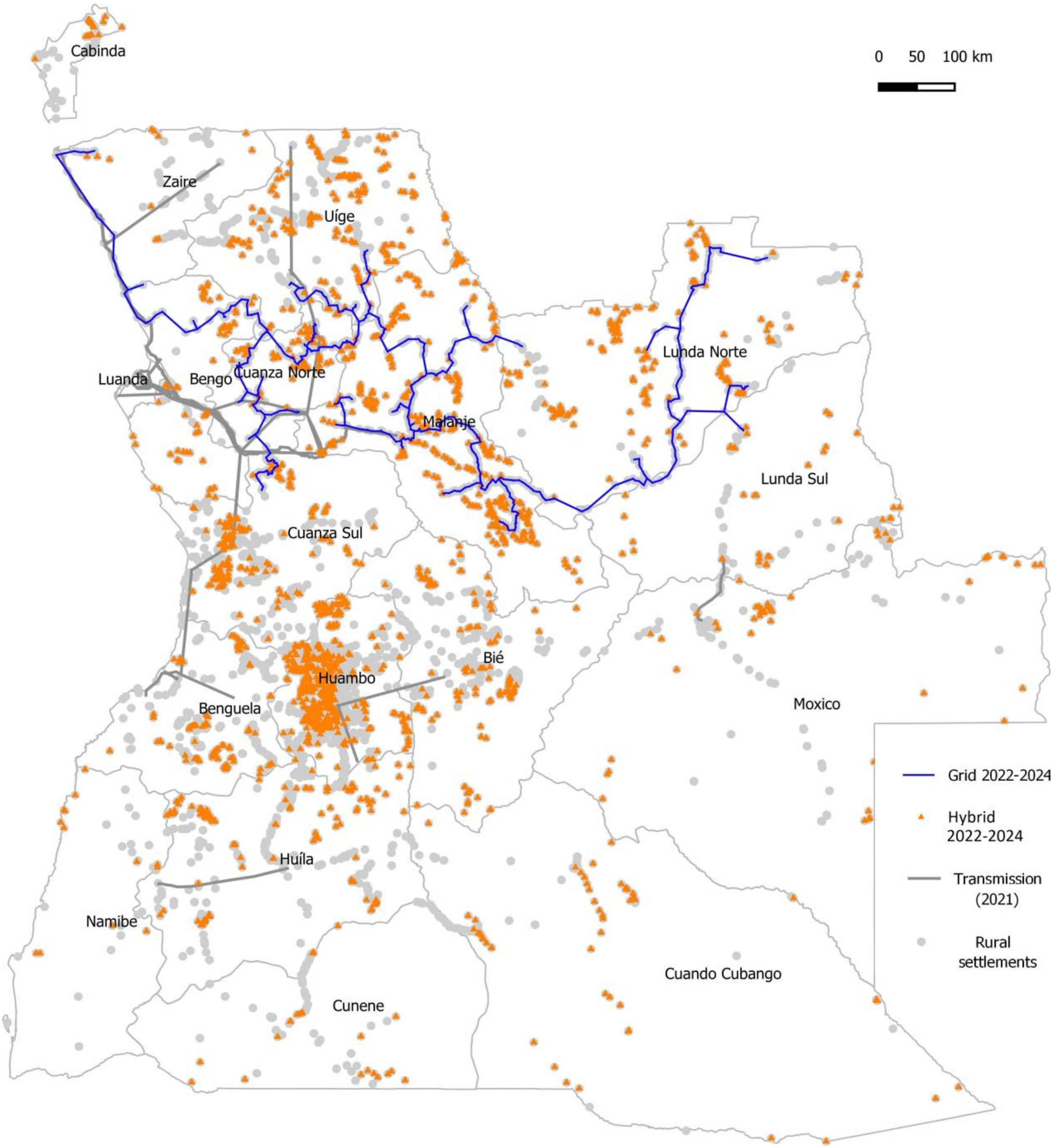


Figure 45 – Scenario 3 (Balanced scenario) – Expected grid and hybrid growth – prioritizing more populated settlements (2022 – 2024)

Produced by the author

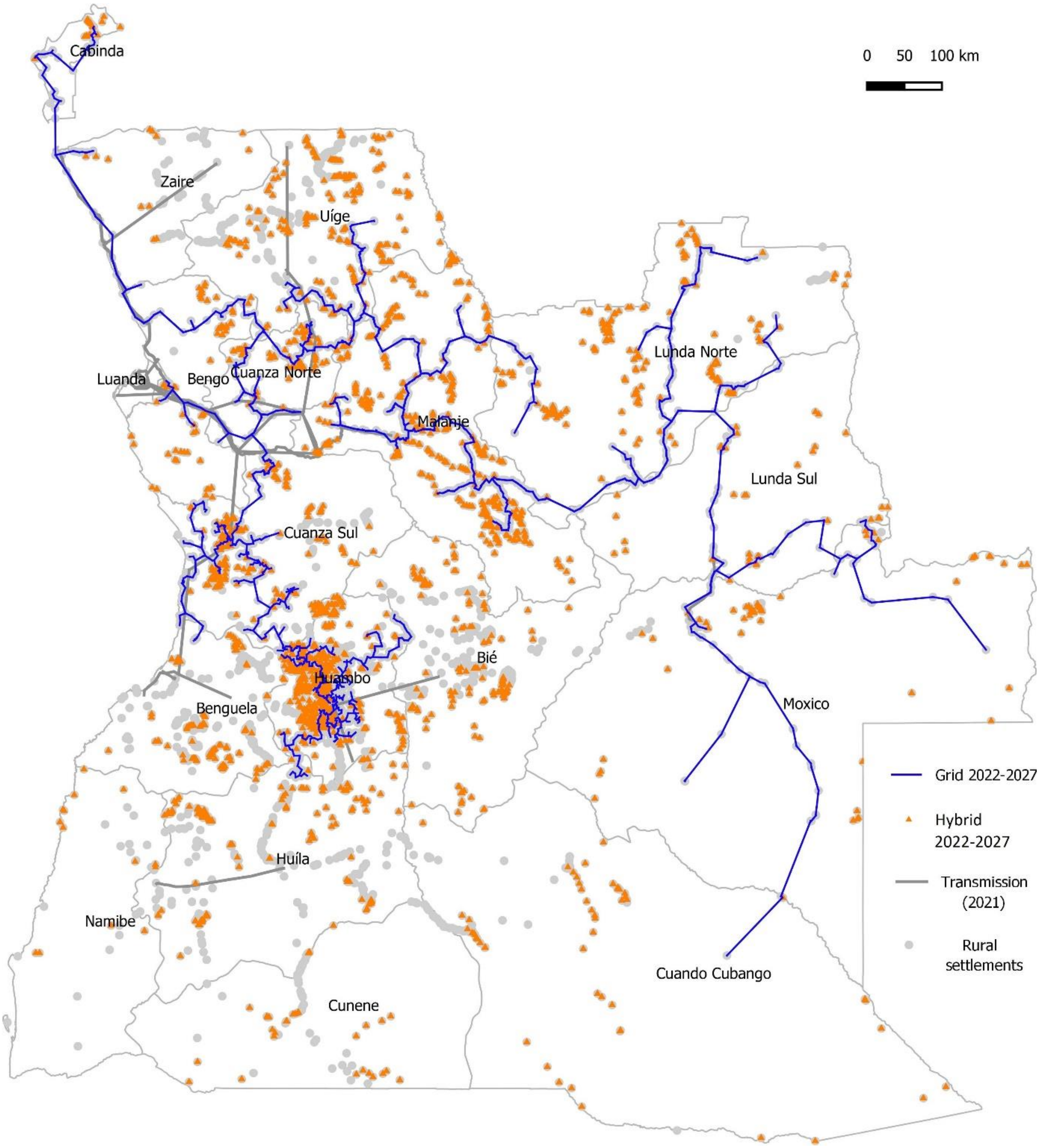


Figure 46 – Scenario 3 (Balanced scenario) – Expected grid and hybrid growth – prioritizing more populated settlements (2022 – 2027)

Produced by the author

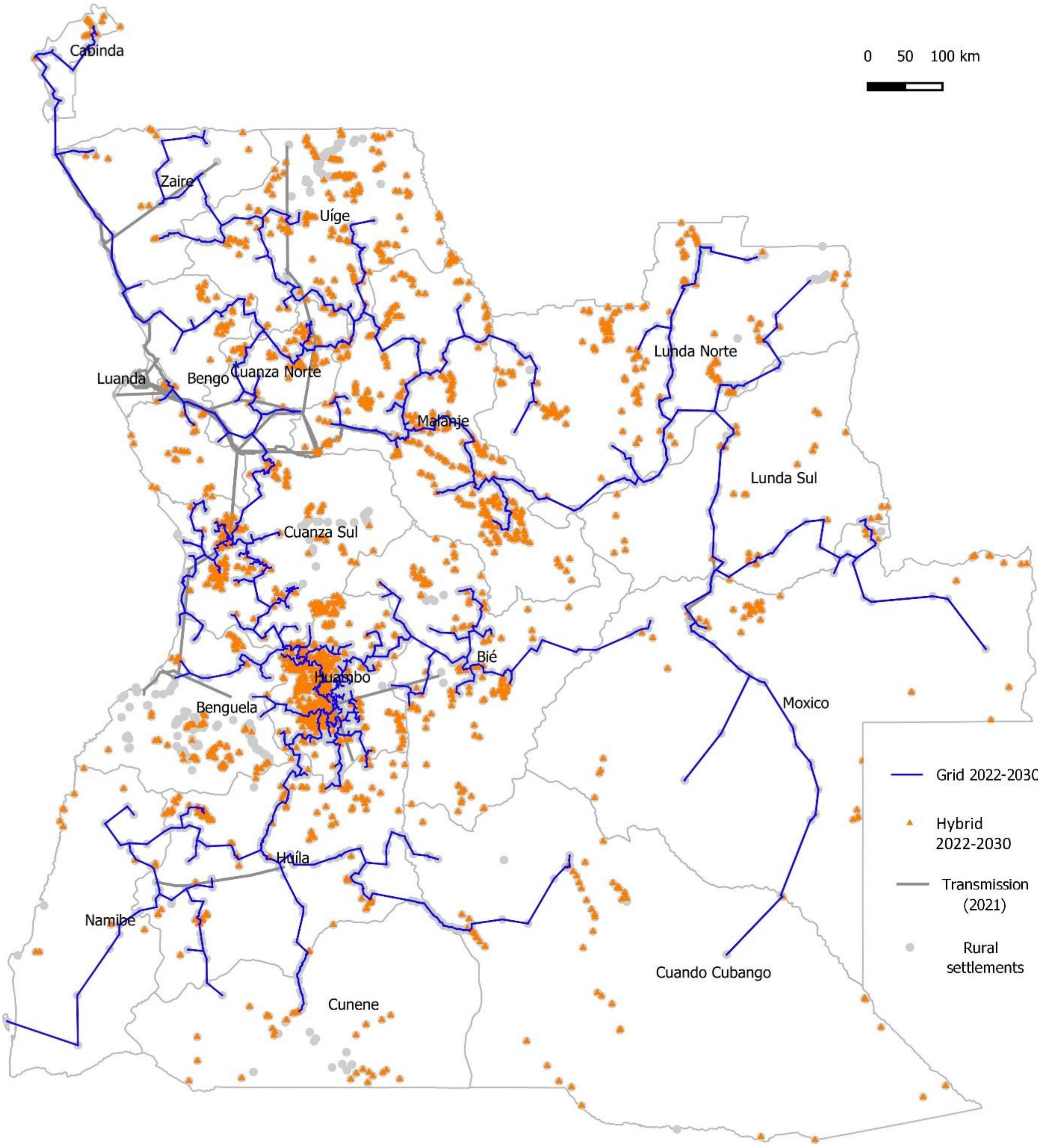


Figure 47 – Scenario 3 (Balanced scenario) – Expected grid and hybrid growth – prioritizing more populated settlements (2022 – 2030)
Produced by the author

Results from Table 48, Figure 45, Figure 46 and Figure 47 concern the scenario where population settlement size is prioritized. Some relevant notes on these results are that nearly 14 million are people connected, which represents 82% of the total rural population spread across 4,238 settlements while 51% of people are connected with grid solutions. In terms of average MPI the value is 0.4416 which could signal that, on average, some of the poorest communities (high MPI) are not located in larger settlements.

A comparison of all scenarios when using different criteria is given in Figure 48.

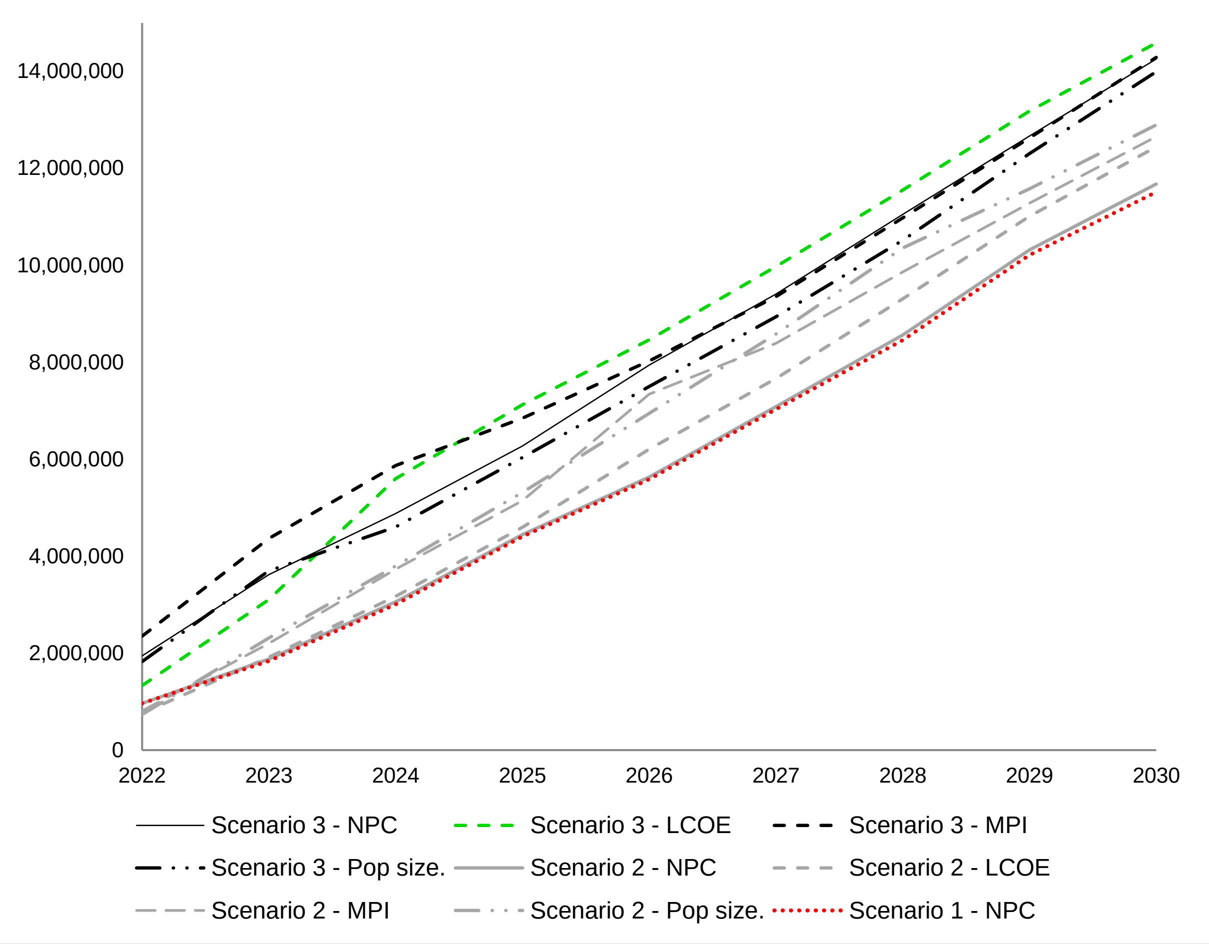


Figure 48 – Scenario comparison – people connected from 2022 to 2030.

Produced by the author

One clear conclusion from the results aggregated in Figure 48 is that, across all scenarios, the inclusion of hybrid mini-grids contributes to a higher estimated number of people connected. The main results for each scenario are given in Table 49.

Table 49 – Metric comparison for the three scenarios

	People connected	Settlements connected	Avg. LCOE	Avg. MPI
Scenario 1 min. NPC	11,512,185	3,852	0.9409	0.4805
Scenario 2 min. NPC	11,677,628	3,943	0.4615	0.4777
Scenario 2 min. LCOE	12,444,456	3,651	0.3691	0.4603
Scenario 2 max. MPI	12,654,603	3,522	0.4610	0.4645
Scenario 2 max. pop. Size	12,896,870	3,219	0.3588	0.4801
Scenario 3 min. LCOE	14,587,711	4,132	0.2647	0.4542
Scenario 3 min. NPC	14,256,088	4,333	0.2260	0.4842
Scenario 3 max. MPI	14,288,830	4,093	0.3830	0.5072
Scenario 3 max. pop. Size	13,994,350	4,238	0.3722	0.4416

Produced by the author

Scenario 1 is used mainly as a baseline and, according to Table 49, it is clear that generally, scenario 2 presented better results than scenario 1, notwithstanding a very slight drop in average MPI. Due to the cost of the hybrid solution in this scenario, and batteries in particular, hybrid technologies have a relatively low percentual contribution to the global supply mix, from 3.6% to 4.5%. It is also shown that connecting more settlements does not mean connecting more people. For scenario 2, the best results are attained when population size maximization is used as a criterion:

- The number of people connected, compared to the grid exclusive scenario (scenario 1) grew by 1,384,685 people to a total of 12.9 million people connected representing 76% of all rural population, the highest result across scenario 2;
- This is achieved while connecting 633 fewer settlements than in scenario 1, a reduction of 16%. This was expected since larger settlements are being connected first;
- The number of settlements connected is evenly spread in the period;
- Average LCOE is 0.3588 USD/kWh, the lowest in all of scenario 2;
- In terms of MPI the average value of 0.4801, close to the value from scenario 1;

Overall scenario 2, prioritizing larger settlements, is an improvement over scenario 1. The only downside, when compared to the initial scenario, is that fewer poor communities are being reached since the average MPI is 0.4801 which is slightly lower than the average MPI of 0.4805, from scenario 1.

For the balanced scenario (scenario 3) the main change is the removal of batteries which are the element with highest impact in the overall cost of photovoltaic technology, as given in the sensitivity analysis in Figure 28. This is compensated with the greater participation of diesel energy as a backup to PV energy. The results are a significant departure from previous scenarios in a number of ways. Scenario 3 has better results across the board, regardless of criterion of analysis (NPC, LCOE, MPI or population size). For example, when NPC is used as a minimization criterion, there is a substantial increase in system sustainability since the average LCOE is 0.2260 USD/kWh, the lowest attained in this research. On the other hand, using LCOE minimization as a criterion yields the best results in terms of people connected (14,587,711) representing 86% of the total rural population which represents an additional 3,075,526 people when compared to scenario 1 and, similarly to scenario 2, it does not yield the highest number of settlements reached. Average LCOE is 0.2647 USD/kWh which is slightly higher than in the previous scenarios. Hybrid connection rose to 33% of total connections, the highest across all scenarios. Other useful results are attained when MPI is used as a maximization criterion which yield the second highest number of people connected (14,288,830) and the highest MPI (0.5072) meaning poorer communities are being reached. This signals a trade-of between connecting the largest number of people and minimizing regional asymmetries but, regardless of the criterion that is chosen, these results show that scenario 3 is an improvement over the previous two. Therefore, it is preferable to reduce the current national focus on grid extension and introduce hybrid mini-grids to increase the number of people with access to electricity until 2030.

Page intentionally left blank.

CHAPTER 5

Conclusions

In this chapter some of the most relevant findings are summarized. The chapter is comprised of five elements: the conclusions, the theoretical contribution, the implications of the results, some limitations and recommendations for future avenues of research.

This study addresses the problem of low access to electricity, in rural areas of Angola, and the challenges faced by the decision-makers due to lack of reliable information. It is put forward that, for developing countries, mini-grids play an important role to accelerate electrification in remote areas. However, this is not being done in Angola, since the focus has been on the grid expansion. This apparent contradiction contributed to the formulation of the research hypothesis: "For some rural communities of Angola, hybrid mini-grids are the fastest and most economically feasible solutions for electrification, for the period between 2022 and 2030".

A literature review is carried out showing that there are few studies on this subject, especially in the non-Anglo Saxon and non-Francophone countries like Angola and, in order to make a contribution to filling this gap, a two-stage, multiperiod methodology is proposed to assess if the introduction of hybrid mini-grids in Angola's supply mix can speed up the access to electricity for rural communities. The approach is comprised of a data collection stage, with five components, and a methodology stage as shown in Figure 15.

Regarding the first stage, the location and size of rural communities were not available at the time of writing and, knowing where people live is a crucial information for planning and decision-making for access to electricity. This issue was tackled in Subchapter 3.1.1 in which a smart interpolation procedure is proposed (Figure 17) to the distribute the country's rural population for 2030 across 4,557 known settlements. Achieving these results implies combining ancillary data from several sources (e.g., demographic and GIS data) which served as inputs for the procedure. In Subchapter 3.1.2 a bottom-up approach is proposed to determine the yearly demand level, per rural settlement, at each time step of the planning process, from 2022 to 2030. This approach took into the account factors like the categorization of population as urban or rural, the population distribution by settlement, the relative level of poverty by location and a set of demand patterns. Applying the procedure shown in Figure 19 the total electricity demand for 2030, for each rural settlement, is estimated. The larger portion of this demand concerns household applications (81.8%), productive uses represented 17.9% while health and education needs are residual, despite being very important for populations.

Sub-chapter 3.1.3 focused on analyzing the available energy sources in the region and deciding which ones made sense for the rural electrification problem in Angola. Hydro power has great potential in Angola but is excluded due to lack of data, which is a recurring theme in literature, concerning this type of energy. Wind energy was equally excluded due to not having sufficient potential in the country's rural areas. The final supply mix included photovoltaic energy and diesel energy, in different configurations, as potential replacements for grid connection.

Decision makers, in most latitudes, take into account economic factors and this is also the case in SSA. With this reality in mind Sub-chapter 3.1.4 focused on the cost structure of energy solutions. In order to facilitate the decision-making process, it is important to know the approximate cost of implementing different grid and off-grid energy solutions. To do so it is necessary to define the cost of individual energy solutions, so that they can be combined, and their costs computed, for different demand sizes and applications. This was done combining inputs from the literature with local supplier information for grid extensions, PV solutions and diesel energy.

Subchapter 3.1.5 detailed the process of choosing adequate metrics to support decision making for the problem of electrifying rural communities. NPC is widely used in feasibility analysis and LCOE is a staple in the literature for comparing the cost of different energy solutions. Cost computations were carried out for grid extension, PV-diesel hybrid mini-grids with and without batteries. While hybrid mini-grids are isolated and independent from each other, grid extension is more complex, in terms of cost estimation, since each new connection depends on the previous one. To approximate the real cost of grid extension, Kruskal's MST algorithm was used to simulate grid growth, with the weights of each edge representing the cost of connecting a new settlement. Globally speaking, it was shown that connecting all of rural Angola with hybrid mini-grids is unfeasible, but these initial results did not take into account that some settlements might be better served with hybrid solutions while, for others cost might be prohibitively high. This means a settlement-by-settlement analysis has to be carried out for all 4.557 rural settlements. This is done in the methodology stage in Sub-chapter 3.2.

Given the dimension and complexity of planning for rural access to electricity in the second stage, a heuristic approach is used. The advantage of this type of approach is that heuristic algorithms produce feasible solutions fast. Three scenarios, that can support the job of decision-makers are proposed: a grid exclusive scenario (Scenario 1) representing the current national strategy focused on grid expansion, a grid intensive scenario, where hybrid solutions are allowed (Scenario 2) and a balanced scenario (Scenario 3) in which the participation of hybrid solutions increases. Scenario 2 combines PV energy, with batteries and diesel. In this scenario diesel represents only 20% of supply while scenario 3 also uses PV energy, but diesel represents 50% of supply, to compensate the absence of batteries.

For both scenarios, there are communities that will always be cheaper to connect via hybrid solutions. In Subchapter 3.2.2 an algorithm based on Prim's algorithm is introduced to select rural settlements to be connected by grid extension or hybrid solutions, according to four criteria (NPC, LCOE, MPI and population size). Python and GIS tools are used to apply the procedure, given in Figure 30, to both scenarios 2 and 3.

Globally, the third scenario presented the best results. This coincided with the introduction of a less costly hybrid solution, without batteries. In this scenario, minimizing LCOE allows the connection of the highest number of people in rural areas (14,587,711), coupled with an average LCOE of 0.2647 USD/kWh. Comparatively, Scenario 3, while minimizing NPC, connected only 14,256,088 people but presented the highest number of connected rural settlements (4,333) and the lowest average LCOE across all Scenarios (0.2260 USD/kWh), an important improvement in system sustainability. Nonetheless, if we consider MPI maximization, an average MPI value of 0.5072 is attained, the highest across all scenarios, which seems to signal that this scenario would target the largest number of impoverished communities.

In absolute terms, the scenario that best addresses the central problem of this research and confirms the hypothesis that, for some rural communities of Angola, hybrid mini-grids are the fastest and most economically feasible solutions for electrification, for the period between 2022 and 2030, is scenario 3 choosing LCOE minimization as a criterion since it allows the connection of an additional 3,075,526 million people when compared to a mostly grid exclusive scenario. Moreover, many of these people are connected from 2022 to 2024 and reside in remote rural areas that will not be connected by a grid exclusive scenario as is given in Figure 35 where southern settlements are mostly unconnected by 2030. In fact, after analyzing the results, it also becomes clear that grid expansion starts from north of the country to other areas, mostly near the coastline and eastern border and will always reach areas of the south at the end of the period or not at all. Hybrid systems are the solution to mitigate this problem and minimize regional asymmetries in Angola by granting access to electricity to remote rural areas.

Nonetheless, it is apparent that there is trade-off when choosing between the most people connected (scenario 3 with the LCOE minimization criterion) and prioritizing the poorest communities (scenario 3 with the MPI maximization criterion). Regardless, both are preferable to scenario 1 and the purpose of these results is to present these possibilities to the decision maker, which is achieved.

5.1 Theoretical contribution

There is a need for more energy modeling work to be done in developing countries (Morrissey, 2019) since only a few are prevalent in literature for energy planning (Trotter et al., 2017). This research is the first of its kind, for Angola, and hopes to incentivize further research for this country and others, in the region.

A decision-making support system is designed to help manage the expansion of the electric system in Angola using elements from the literature, but modeling and combining them in a unique way. An integration of these elements, free GIS tools, python code and spreadsheet data manipulation are used with good results, to produce a unique methodology.

The smart interpolation (Leyk et al., 2019) approach that is employed yielded interesting results when compared to similar studies and can be used for research in other sectors:

- Water sector services include drinking water and wastewater;
- Roadways;
- Education;
- Health services;
- Infrastructure projects in the private sector.

While assumptions for demand, supply and cost structure are derived from suppliers and literature, the scenarios that are put forward are ambitious since the hybrid solutions represent real alternatives to grid extension, allowing for an instantaneous and disruptive impact in the way of life of rural populations.

The modeling approach is carried out for the entire country, at a settlement level, which is unusual in an approach of this complexity. Other studies tend to focus on smaller sub-sections of countries (Blechinger et al., 2019; Moner-Girona et al., 2019) or have lower resolution (Trotter et al., 2019).

This research tackles the issue of uncertainty using sensitivity analysis of key cost components and three scenarios, considering criteria like NPC, LCOE, MPI and settlement population size. The lack of these types of uncertainty mitigation measures is one of the most relevant shortcomings of similar studies (Morrissey, 2019). Another common criticism is the use of overnight modeling since very few studies use time steps (Morrissey, 2019).This is addressed, since grid and hybrid connections are monitored, on a yearly basis, from 2022 to 2030.

This study also avoids using the assumption of a corridor around the grid inside which, all settlements are considered grid connected (Bertheau et al., 2017; Kemausuor et al., 2014; Moner-

Girona et al., 2019) and all settlements go through the same selection process between hybrid and grid extension.

A pruning technique is introduced, using graph theory, to side step the problem that removing one vertex from the grid, affects all subsequent vertices. Additionally, the methodology is transparent, scalable, custom-made and can be used in other countries or regions to support decision-making whereas, some premade tools used in other studies have models that are insufficiently known and have a high learning curve.

In the literature, the degree in which mini-grids participate in the supply mix varies significantly, from country to country. This is sometimes problematic since, there is an expectation that the results for SSA countries should not be so varied. This is a somewhat subjective view, since countries have differences. An objective contribution is made, by introducing different configurations of the same hybrid technologies (PV and diesel), with differing results. These results show that it is possible that some of the differences stem from the use of different hybrid solutions that represent, more or less robust, alternatives to grid extension. This has been shown to impact the results. For example, this research uses energy alternatives that provide a higher level of service and are, therefore, more expensive. When the quality of the solution is decreased (scenario 3) the participation of hybrid solution rises.

Finally, the study proposes using output maps, for different time steps, in order to make interpretation of results as clear as possible.

5.2 Managerial and societal implications

The subject of access to energy is a longstanding issue in Angola and the choices and investments that have been made, in the last 5 years, mostly focused on generation and grid intensification while grid expansion has been mostly stagnant. A rethinking of this national strategy could contribute decisively to achieve the ambitious goals, for the coverage of the grid, that have been subscribed for 2030 (ECCAS - Economic Community of Central African States, 2014; MINEA, 2018).

The use of this methodology can shorten the time to produce reports and improve the quality of decision-making. The consequence of this may be the introduction, in the supply mix, of the proposed hybrid systems, which will have an immense impact in people’s lives and on the country as a whole. For the impoverished population of rural Angola, especially in the south, it represents the opportunity for a different way of life beyond mere survival due to the economic opportunities that become available with access to electricity.

5.3 Research limitations

This research represents the first step in a journey. Quality data is a problem for research in Angola since there is no culture of data gathering in the country. This means some inputs, like geolocation have to be extracted or inferred from databases that present some issues of scale, currency and temporal agreement (Leyk et al., 2019). This is mitigated using the methods described in the methodology chapter, but the issues remain present, as in any research of this kind.

Another issue is that generation cost is not included, since it was not available, so the results are somewhat skewed in favor of grid extension.

The study works on the assumption that majority of population changes will occur in the same settlements, not through the extinction and creation of new ones. At the time of writing, this assumption seems more realistic since provinces have been in lockdown for a year and half due to COVID-19 pandemic. How long this will last is uncertain and currently few people are vaccinated.

Technological advances can impact the results for example, when solid-state batteries are introduced in the market, the conclusions may vary significantly, in favor of increasing the contribution of hybrid solutions. It is not expected that this technology becomes widely available in the period of this research.

5.4 Avenues for future research

There is a sincere hope that new research that focuses on Angola, surfaces. In terms of future research, some potential avenues are:

- Automating the extraction of ancillary data from OpenStreetMap, google earth, and similar databases since it would surely improve results;
- Developing research on hydro potential, especially regarding mini-grids, including underground water layers. There are some studies on this (Korkovelos et al., 2018; Szabó et al., 2013), but for the case of Angola, more needs to be done;
- Modeling the volatility of oil prices since this influences the results;
- Modeling the national grid generation cost;
- Conducting research on the degree of acceptance of populations for hybrid solutions, sometimes seen as an inferior alternative. This could originate campaigns to prepare communities before connection, to avoid resistance to the technology, in order to maximize the potential success of these projects.

- Application of the principles and structure that have guided this research to other sectors that affect human quality of life;

Finally, as a recommendation, it would be desirable that research starts being conducted in Angola by Angolans, since they know and live the reality of the country better than anyone and, despite all the difficulties, show a strong desire to carry out this work.

Page intentionally left blank.

References

Aldersey-Williams, J., & Rubert, T. (2019). Levelised cost of energy – A theoretical justification and critical assessment. *Energy Policy*, *124*(February 2018), 169–179. https://doi.org/10.1016/j.enpol.2018.10.004

Alkire, S., & Foster, J. (2011). Counting and multidimensional poverty measurement. *Journal of Public Economics*, *95*(7–8), 476–487. https://doi.org/10.1016/j.jpubeco.2010.11.006

Bakillah, M., Liang, S., Mobasheri, A., Jokar Arsanjani, J., & Zipf, A. (2014). Fine-resolution population mapping using OpenStreetMap points-of-interest. *International Journal of Geographical Information Science*, *28*(9), 1940–1963. https://doi.org/10.1080/13658816.2014.909045

Batista e Silva, F., Gallego, J., & Lavalle, C. (2013). A high-resolution population grid map for Europe. *Journal of Maps*, *9*(1), 16–28. https://doi.org/10.1080/17445647.2013.764830

Bertheau, P., Cader, C., & Blechinger, P. (2016). Electrification Modelling for Nigeria. *Energy Procedia*, *93*(March), 108–112. https://doi.org/10.1016/j.egypro.2016.07.157

Bertheau, P., Oyewo, A. S., Cader, C., Breyer, C., & Blechinger, P. (2017). Visualizing national electrification scenarios for sub-saharan African countries. *Energies*, *10*(11), 1–20. https://doi.org/10.3390/en10111899

Bhatia, M., & Angelou, N. (2015). Beyond Connections Energy Access Redefined. World Bank. Energy Sector Management Assistance Program (ESMAP). *The World Bank*, 228. https://openknowledge.worldbank.org/handle/10986/24368

Bhattacharyya, S. C., & Timilsina, G. R. (2010). Modelling energy demand of developing countries: Are the specific features adequately captured? *Energy Policy*, *38*(4), 1979–1990. https://doi.org/10.1016/j.enpol.2009.11.079

Blechinger, P., Cader, C., & Bertheau, P. (2019). Least-Cost Electrification Modeling and Planning - A Case Study for Five Nigerian Federal States. *Proceedings of the IEEE*, *107*(9), 1923–1940. https://doi.org/10.1109/JPROC.2019.2924644

Blume, S. W. (2007). *Electrical power for the nonelectrical professional*.

Branker, K., Pathak, M. J. M., & Pearce, J. M. (2011). A review of solar photovoltaic levelized cost of electricity. *Renewable and Sustainable Energy Reviews*, *15*(9), 4470–4482. https://doi.org/10.1016/j.rser.2011.07.104

Çalışkan, M., & Anbaroğlu, B. (2020). Geo-MST: A geographical minimum spanning tree plugin for QGIS. *SoftwareX*, *12*, 100553. https://doi.org/10.1016/j.softx.2020.100553

Candelise, C., Chan, N. L. A., Gambhir, A., Foster, S., Nelson, J., Ekins-Daukes, N., Sandwell, P., Buckle, S. J., Nagpal, D., & Emmott, C. J. M. (2016). Off-grid solar photovoltaic systems for rural electrification and emissions mitigation in India. *Solar Energy Materials and Solar Cells*, *156*, 147–156. https://doi.org/10.1016/j.solmat.2016.04.030

Ciller, P., & Lumbreras, S. (2020). Electricity for all: The contribution of large-scale planning tools to the energy-access problem. *Renewable and Sustainable Energy Reviews*, *120*(December 2019), 109624. https://doi.org/10.1016/j.rser.2019.109624

Corigliano, S., Carnovali, T., Edeme, D., & Merlo, M. (2020). Holistic geospatial data-based procedure for electric network design and least-cost energy strategy. *Energy for Sustainable Development*, *58*, 1–15. https://doi.org/10.1016/j.esd.2020.06.008

Cuisinier, E., Penz, B., Bourasseau, C., Ruby, A., & Lemaire, P. (2020). *Techno-economic planning of local energy systems through optimization models : a survey of current methods*. *August*. https://doi.org/10.1002/er.6208

Dagnachew, A. G., Lucas, P. L., Hof, A. F., Gernaat, D. E. H. J., de Boer, H. S., & van Vuuren, D. P. (2017). The role of decentralized systems in providing universal electricity access in Sub-Saharan Africa – A model-based approach. *Energy*, *139*, 184–195. https://doi.org/10.1016/j.energy.2017.07.144

de Roo, A., Siyal, S., Scholtz, E., Taliotis, C., Mentis, D., Rogner, H., Oudalov, A., Arderne, C., Howells, M., Korkovelos, A., Tanvez, Y., Bazilian, M., & Zepeda, E. (2017). Lighting the World: the first application of an open source, spatial electrification tool (OnSSET) on Sub-Saharan Africa. *Environmental Research Letters*, *12*(8), 85003. https://doi.org/10.1088/1748-9326/aa7b29

Deichmann, U., Meisner, C., Murray, S., & Wheeler, D. (2011). The economics of renewable energy expansion in rural Sub-Saharan Africa. *Energy Policy*, *39*(1), 215–227. https://doi.org/10.1016/j.enpol.2010.09.034

Diário da República II Série - N.º 54. (2019). Critérios de enquadramento de pagamentos dos seguintes encargos constantes no Regulamento do Fornecimento de Energia Eléctrica. *Instrutivo Presidencial, N.º 2/19 de 26 de Abril*.

ECCAS - Economic Community of Central African States. (2014). *Livre Blanc de la CEEAC et de la CEMAC Politique régionale pour un accès universel aux services énergétiques modernes et le développement économique et social. 2014 - 2030*.

Ellman, D. (2015). The Reference Electrification Model: A Computer Model for Planning Rural Electricity Access. *Physics, Princeton University*, *2009*, 109. https://dspace.mit.edu/bitstream/handle/1721.1/98551/920674644-MIT.pdf?sequence=1

Florczyk, A. J., Melchiorri, M., Zeidler, J., Corbane, C., Schiavina, M., Freire, S., Sabo, F., Politis, P., Esch, T., & Pesaresi, M. (2020). The Generalised Settlement Area: mapping the Earth surface in the vicinity of built-up areas. *International Journal of Digital Earth*, *13*(1), 45–60. https://doi.org/10.1080/17538947.2018.1550121

Freire, S., Schiavina, M., Florczyk, A. J., MacManus, K., Pesaresi, M., Corbane, C., Borkovska, O., Mills, J., Pistolesi, L., Squires, J., & Sliuzas, R. (2020). Enhanced data and methods for improving open and free global population grids: putting 'leaving no one behind' into practice. *International Journal of Digital Earth*, *13*(1), 61–77. https://doi.org/10.1080/17538947.2018.1548656

Garcia, M. M., & Hawes, S. (2021). Africa Energy Outlook 2019. *Energy Outlook*. https://webstore.iea.org/africa-energy-outlook-2019

Gaughan, A. E., Stevens, F. R., Linard, C., Patel, N. N., & Tatem, A. J. (2015). Exploring nationally and regionally defined models for large area population mapping. *International Journal of Digital Earth*, *8*(12), 989–1006. https://doi.org/10.1080/17538947.2014.965761

Girres, J. F., & Touya, G. (2010). Quality Assessment of the French OpenStreetMap Dataset. *Transactions in GIS*, *14*(4), 435–459. https://doi.org/10.1111/j.1467-9671.2010.01203.x

Grippa, T., Linard, C., Lennert, M., Georganos, S., Mboga, N., Vanhuysse, S., Gadiaga, A., & Wolff, E. (2019). Improving urban population distribution models with very-high resolution satellite information. *Data*, *4*(1). https://doi.org/10.3390/data4010013

Haklay, M. (2010). How good is volunteered geographical information? A comparative study of OpenStreetMap and ordnance survey datasets. *Environment and Planning B: Planning and Design*, *37*(4), 682–703. https://doi.org/10.1068/b35097

Halsnæs, K., & Garg, A. (2011). Assessing the Role of Energy in Development and Climate Policies-Conceptual Approach and Key Indicators. *World Development*, *39*(6), 987–1001. https://doi.org/10.1016/j.worlddev.2010.01.002

Hay, S. I., Noor, A. M., Nelson, A., & Tatem, A. J. (2005). *The accuracy of human population maps for public health application*. *10*(10), 1073–1086. https://doi.org/10.1111/j.1365-3156.2005.01487.x

Herfort, B., Li, H., Fendrich, S., Lautenbach, S., & Zipf, A. (2019). Mapping human settlements with higher accuracy and less volunteer efforts by combining crowdsourcing and deep learning. *Remote Sensing*, *11*(15). https://doi.org/10.3390/rs11151799

INE - Instituto Nacional de Estatística. (2016). Angola 2014 Census. *Ine*.

INE - Instituto Nacional de Estatística. (2018). Projecção Da População 2015-2050. *Ine*.

INE - Instituto Nacional de Estatística. (2019). Pobreza multidimensional nos municípios de Angola. *Ine*.

International Monetary Fund. (2015). Angola: Selected Issues. *IMF Staff Country Reports*, *15*(302), 1. https://doi.org/10.5089/9781513547855.002

IRENA. (2018). *Measurement and estimation of off-grid solar , hydro and biogas energy*.

Jordan, D. C., & Kurtz, S. R. (2013). Photovoltaic degradation rates - An Analytical Review. *Progress in Photovoltaics: Research and Applications*, *21*(1), 12–29. https://doi.org/10.1002/pip.1182

Kemausuor, F., Adkins, E., Adu-Poku, I., Brew-Hammond, A., & Modi, V. (2014). Electrification planning using Network Planner tool: The case of Ghana. *Energy for Sustainable Development*, *19*(1), 92–101. https://doi.org/10.1016/j.esd.2013.12.009

Khan, M., Kabir, F., & Parra, J. P. (2007). Financing Large Projects: Using Project Finance Techniques and Practices. In *Prentice Hall*. Elsevier B.V. https://doi.org/10.1016/j.proeng.2015.11.019

Korkovelos, A., Khavari, B., Sahlberg, A., Howells, M., & Arderne, C. (2019). The role of open access data in geospatial electrification planning and the achievement of SDG7. An onsset-based case study for Malawi. *Energies*, *12*(7). https://doi.org/10.3390/en12071395

Korkovelos, A., Mentis, D., Siyal, S. H., Arderne, C., Rogner, H., Bazilian, M., Howells, M., Beck, H., & De Roo, A. (2018). A geospatial assessment of small-scale hydropower potential in sub-saharan Africa. *Energies*, *11*(11). https://doi.org/10.3390/en11113100

Kruskal, J. B. (1956). On the Shortest Spanning Subtree of a Graph and the Traveling Salesman Problem. *Proceedings of the American Mathematical Society*, *7*(1), 48. https://doi.org/10.2307/2033241

Kumar, A., Sah, B., Singh, A. R., Deng, Y., He, X., Kumar, P., & Bansal, R. C. (2017). A review of multi criteria decision making (MCDM) towards sustainable renewable energy development. *Renewable and Sustainable Energy Reviews*, *69*(June 2016), 596–609. https://doi.org/10.1016/j.rser.2016.11.191

Leyk, S., Gaughan, A. E., Adamo, S. B., De Sherbinin, A., Balk, D., Freire, S., Rose, A., Stevens, F. R., Blankespoor, B., Frye, C., Comenetz, J., Sorichetta, A., Macmanus, K., Pistolesi, L., Levy, M., Tatem, A. J., & Pesaresi, M. (2019). The spatial allocation of population: a review of large-scale gridded population data products and their fitness for use. *Earth System Science Data*, *11*(3), 1385–1409. https://doi.org/10.5194/essd-11-1385-2019

Li, X., & Zhou, Y. (2017). Urban mapping using DMSP/OLS stable night-time light: a review. *International Journal of Remote Sensing*, *38*(21), 6030–6046. https://doi.org/10.1080/01431161.2016.1274451

Linard, C., Gilbert, M., & Tatem, A. J. (2011). Assessing the use of global land cover data for guiding large area population distribution modelling. *GeoJournal*, *76*(5), 525–538. https://doi.org/10.1007/s10708-010-9364-8

López-González, A., Ferrer-Martí, L., & Domenech, B. (2019). Sustainable rural electrification planning in developing countries: A proposal for electrification of isolated communities of Venezuela. *Energy Policy*, *129*(February), 327–338. https://doi.org/10.1016/j.enpol.2019.02.041

Mahbub, M. S., Viesi, D., Cattani, S., & Crema, L. (2017). An innovative multi-objective optimization approach for long-term energy planning. *Applied Energy*, *208*(September), 1487–1504. https://doi.org/10.1016/j.apenergy.2017.08.245

Malczewski, J., & Rinner, C. (2015). Multicriteria Decision Analysis in Geographic Information Science. In *Analysis methods* (Issue Massam 1993). http://www.amazon.com/Multicriteria-Decision-Analysis-Geographic-Information/dp/3540747567/ref=sr_1_1?ie=UTF8&qid=1430864854&sr=8-1&keywords=Multicriteria+decision+analysis+in+geographic+information+science

Mamaghani, A., Avella Escandon, S. A., Najafi, B., Shirazi, A., & Rinaldi, F. (2016). Techno-economic feasibility of photovoltaic, wind, diesel and hybrid electrification systems for off-grid rural electrification in Colombia. *Renewable Energy*, *97*, 293–305. https://doi.org/10.1016/j.renene.2016.05.086

Mentis, D., Siyal, S. H., Korkovelos, A., & Howells, M. (2017). Estimating the spatially explicit wind generated electricity cost in Africa - A GIS based analysis. *Energy Strategy Reviews*, *17*, 45–49. https://doi.org/10.1016/j.esr.2017.07.002

Mentis, Welsch, M., Fuso Nerini, F., Broad, O., Howells, M., Bazilian, M., & Rogner, H. (2016). A GIS-based approach for electrification planning-A case study on Nigeria. *Energy for Sustainable Development*, *29*, 142–150. https://doi.org/10.1016/j.esd.2015.09.007

MINEA. (2018a). Angola Energia 2025. *Ministério Da Energia e Águas » Contéudo » Geração*, 13. http://www.angolaenergia2025.com/pt-pt/conteudo/geracao-0

MINEA. (2018b). PLANO DE ACÇÃO DO SECTOR DE ENERGIA E ÁGUAS 2018-2022. *MINEA » Entrada » Conselho Consultivo » CC2018 » IVº Painel-Balanço Das Empresas Públicas, Comissões de Gestão e GEPE MINEA*, 15. http://www.minea.gv.ao/index.php/conselho-consultivo/category/135-painel-iv

MINISTÉRIO DAS FINANÇAS. (2018). *Despesa 2018* (pp. 17–19).

MINISTÉRIO DAS FINANÇAS. (2019). *RESUMO DA DESPESA 2019* (pp. 2–3).

MINISTÉRIO DAS FINANÇAS. (2020). *RESUMO DA DESPESA 2020*. 2–3.

MINISTÉRIO DAS FINANÇAS. (2021). *RESUMO DA DESPESA 2021*. 0–1.

Modi, V., Adkins, E., Carbajal, J., & Sherpa, S. (2013). *Liberia Power Sector Capacity Building and Energy Master Planning*. 1–52.

Moksnes, N., Korkovelos, A., Mentis, D., & Howells, M. (2017). Electrification pathways for Kenya-linking spatial electrification analysis and medium to long term energy planning. *Environmental Research Letters*, *12*(9). https://doi.org/10.1088/1748-9326/aa7e18

Moner-Girona, M., Bódis, K., Morrissey, J., Kougias, I., Hankins, M., Huld, T., & Szabó, S. (2019). Decentralized rural electrification in Kenya: Speeding up universal energy access. *Energy for Sustainable Development*, *52*, 128–146. https://doi.org/10.1016/j.esd.2019.07.009

Morrissey, J. (2019). Achieving universal electricity access at the lowest cost: A comparison of published model results. *Energy for Sustainable Development*, *53*, 81–96.

https://doi.org/10.1016/j.esd.2019.09.005

Mullen, W. F., Jackson, S. P., Croitoru, A., Crooks, A., Stefanidis, A., & Agouris, P. (2015). Assessing the impact of demographic characteristics on spatial error in volunteered geographic information features. *GeoJournal*, *80*(4), 587–605. https://doi.org/10.1007/s10708-014-9564-8

Neis, P., Zielstra, D., & Zipf, A. (2013). Comparison of Volunteered Geographic Information Data Contributions and Community Development for Selected World Regions. *Future Internet*, *5*(2), 282–300. https://doi.org/10.3390/fi5020282

Nerini, F. F., Broad, O., Mentis, D., Welsch, M., Bazilian, M., & Howells, M. (2016). A cost comparison of technology approaches for improving access to electricity services. *Energy*, *95*, 255–265. https://doi.org/10.1016/j.energy.2015.11.068

Nissen, U., & Harfst, N. (2019). Shortcomings of the traditional “ levelized cost of energy ” [LCOE] for the determination of grid parity *. *Energy*, *171*, 1009–1016. https://doi.org/10.1016/j.energy.2019.01.093

Ohiare, S. (2015). Expanding electricity access to all in Nigeria: a spatial planning and cost analysis. *Energy, Sustainability and Society*, *5*(1). https://doi.org/10.1186/s13705-015-0037-9

Ouedraogo, N. S. (2017). Modeling sustainable long-term electricity supply-demand in Africa. *Applied Energy*, *190*, 1047–1067. https://doi.org/10.1016/j.apenergy.2016.12.162

Palacios-López, D., Bachofer, F., Esch, T., Marconcini, M., Macmanus, K., Sorichetta, A., Zeidler, J., Dech, S., Tatem, A. J., & Reinartz, P. (2021). High-resolution gridded population datasets: Exploring the capabilities of the world settlement footprint 2019 imperviousness layer for the african continent. *Remote Sensing*, *13*(6). https://doi.org/10.3390/rs13061142

Parshall, L., Pillai, D., Mohan, S., Sanoh, A., & Modi, V. (2009). National electricity planning in settings with low pre-existing grid coverage: Development of a spatial model and case study of Kenya. *Energy Policy*, *37*(6), 2395–2410. https://doi.org/10.1016/j.enpol.2009.01.021

Patela, N. N., Angiuli, E., Gamba, P., Gaughan, A., Lisini, G., Stevens, F. R., Tatem, A. J., & Trianni, G. (2015). Multitemporal settlement and population mapping from landsatusing google earth engine. *International Journal of Applied Earth Observation and Geoinformation*, *35*(PB), 199–208. https://doi.org/10.1016/j.jag.2014.09.005

Prim, R. C. (November 1957). Shortest connection networks And some generalizations. *Bell System Technical Journal, 36 (6): 1389–1401,*

Reed, F. J., Gaughan, A. E., Stevens, F. R., Yetman, G., Sorichetta, A., & Tatem, A. J. (2018). Gridded population maps informed by different built settlement products. *Data*, *3*(3). https://doi.org/10.3390/data3030033

Republic of Angola Ministry of Energy and Water. (2017). *Atlas and NATIONAL STRATEGY FOR THE NEW RENEWABLE ENERGIES*. 242.

http://www.angolaenergia2025.com/sites/default/files/editor/atlas_e_estrategia_nacional_para_as_novas_energias_renovaveis.pdf

Riva, F., Tognollo, A., Gardumi, F., & Colombo, E. (2018). Long-term energy planning and demand forecast in remote areas of developing countries: Classification of case studies and insights from a modelling perspective. *Energy Strategy Reviews*, *20*, 71–89. https://doi.org/10.1016/j.esr.2018.02.006

Sanoh, A., Parshall, L., Sarr, O. F., Kum, S., & Modi, V. (2012). Local and national electricity planning in Senegal: Scenarios and policies. *Energy for Sustainable Development*, *16*(1), 13–25. https://doi.org/10.1016/j.esd.2011.12.005

Senaratne, H., Mobasheri, A., Ali, A. L., Capineri, C., & Haklay, M. (Muki). (2017). A review of volunteered geographic information quality assessment methods. *International Journal of Geographical Information Science*, *31*(1), 139–167. https://doi.org/10.1080/13658816.2016.1189556

Sievert, M., & Steinbuks, J. (2020). Willingness to pay for electricity access in extreme poverty : Evidence from sub-Saharan Africa. *World Development*, *128*, 104859. https://doi.org/10.1016/j.worlddev.2019.104859

Sklar-Chik, M. D., Brent, A. C., & de Kock, I. H. (2016). Critical review of the levelised cost of energy metric. *South African Journal of Industrial Engineering*, *27*(4), 124–133. https://doi.org/10.7166/27-4-1496

Smil, V. (2010). *Energy Transitions - History, Requirements, Prospects*.

Sorichetta, A., Hornby, G. M., Stevens, F. R., Gaughan, A. E., Linard, C., & Tatem, A. J. (2015). High-resolution gridded population datasets for Latin America and the Caribbean in 2010, 2015, and 2020. *Scientific Data*, *2*, 1–12. https://doi.org/10.1038/sdata.2015.45

Stevens, F. R., Gaughan, A. E., Nieves, J. J., King, A., Sorichetta, A., Linard, C., & Tatem, A. J. (2020). Comparisons of two global built area land cover datasets in methods to disaggregate human population in eleven countries from the global South. *International Journal of Digital Earth*, *13*(1), 78–100. https://doi.org/10.1080/17538947.2019.1633424

Suganthi, L., & Samuel, A. A. (2012). Energy models for demand forecasting - A review. *Renewable and Sustainable Energy Reviews*, *16*(2), 1223–1240. https://doi.org/10.1016/j.rser.2011.08.014

Sutton, P., Roberts, D., Elvidge, C., & Baugh, K. (2001). Census from Heaven: An estimate of the global human population using night-time satellite imagery. *International Journal of Remote Sensing*, *22*(16), 3061–3076. https://doi.org/10.1080/01431160010007015

Szabó, S., Bódis, K., Huld, T., & Moner-Girona, M. (2013). Sustainable energy planning: Leapfrogging the energy poverty gap in Africa. *Renewable and Sustainable Energy Reviews*, *28*, 500–509. https://doi.org/10.1016/j.rser.2013.08.044

Tatem, A. J., Noor, A. M., von Hagen, C., Di Gregorio, A., & Hay, S. I. (2007). High resolution population maps for low income nations: Combining land cover and census in East Africa. *PLoS ONE*, *2*(12). https://doi.org/10.1371/journal.pone.0001298

Thomson, D. R., Gaughan, A. E., Stevens, F. R., Yetman, G., Elias, P., & Chen, R. (2021). Evaluating the Accuracy of Gridded Population Estimates in Slums: A Case Study in Nigeria and Kenya. *Urban Science*, *5*(2), 48. https://doi.org/10.3390/urbansci5020048

Tobler, W., Deichmann, U., Gottsegen, J., & Maloy, K. (1997). World population in a grid of spherical quadrilaterals. *International Journal of Population Geography*, *3*(3), 203–225. https://doi.org/10.1002/(sici)1099-1220(199709)3:3<203::aid-ijpg68>3.0.co;2-c

Trotter, P. A., Cooper, N. J., & Wilson, P. R. (2019). A multi-criteria, long-term energy planning optimisation model with integrated on-grid and off-grid electrification – The case of Uganda. *Applied Energy*, *243*(March), 288–312. https://doi.org/10.1016/j.apenergy.2019.03.178

Trotter, P. A., McManus, M. C., & Maconachie, R. (2017). Electricity planning and implementation in sub-Saharan Africa: A systematic review. *Renewable and Sustainable Energy Reviews*, *74*(June 2016), 1189–1209. https://doi.org/10.1016/j.rser.2017.03.001

United Nations Development Programme. (2015a). *Global Tracking Framework*.

United Nations Development Programme. (2015b). *Rapid Assesment Gap Analysis: Angola*.

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. https://doi.org/10.1016/j.energy.2011.11.037

Weiss, D. J., Nelson, A., Gibson, H. S., Temperley, W., Peedell, S., Lieber, A., Hancher, M., Poyart, E., Belchior, S., Fullman, N., Mappin, B., Dalrymple, U., Rozier, J., Lucas, T. C. D., Howes, R. E., Tusting, L. S., Kang, S. Y., Cameron, E., Bisanzio, D., … Gething, P. W. (2018). inequalities in accessibility in 2015. *Nature Publishing Group*, *553*(7688), 333–336. https://doi.org/10.1038/nature25181

Zeyringer, M., Pachauri, S., Schmid, E., Schmidt, J., Worrell, E., & Morawetz, U. B. (2015). Analyzing grid extension and stand-alone photovoltaic systems for the cost-effective electrification of Kenya. *Energy for Sustainable Development*, *25*, 75–86. https://doi.org/10.1016/j.esd.2015.01.003

Zhao, Y., Li, Q., Zhang, Y., & Du, X. (2019). Improving the accuracy of fine-grained population mapping using population-sensitive POIs. *Remote Sensing*, *11*(21), 11–13.

APPENDICES

Appendix A – Multidimensional index of poverty (MPI), by municipality, by quintile.

Ord.	Municipalit	MPI	Quintile	Ord.	Municipal	MPI	Quintile	Ord.	Municipality	MPI	Quintile	Ord.	Municipalit	MPI	Quintile	Ord.	Municipal	MPI	Quintile
1	Luanda	0.029		34	Baía Farta	0.370		67	Quibala	0.464		100	Mucaba	0.548		133	Cuangar	0.607	
2	Cazenga	0.030		35	Luena	0.374		68	Mussende	0.465		101	Cubal	0.553		134	Capenda-	0.608	
3	Kilamba	0.058		36	Menongu	0.374		69	Balombo	0.470		102	Caconda	0.553		135	Namacun	0.612	
4	Talatona	0.073		37	Tomboco	0.377		70	Mucari	0.473		103	Cuango	0.554		136	Jamba	0.615	
5	Viana	0.091		38	Cacuso	0.379		71	Mavinga	0.476		104	Lucapa	0.556		137	Gambos	0.618	
6	Lobito	0.110		39	Caála	0.380		72	Bembe	0.478		105	Quilenda	0.556		138	Cunda-	0.618	
7	Cabinda	0.127		40	Quiculung	0.381		73	Andulo	0.483		106	Calandula	0.557		139	Cuvango	0.620	
8	Benguela	0.141		41	Icolo e	0.383		74	Ngonguembo	0.484		107	Alto	0.561		140	Léua	0.622	
9	Catumbela	0.147		42	Cuimba	0.384		75	Chikala	0.487		108	Quela	0.562		141	Cameias	0.622	
10	Cazengo	0.148		43	Saurimo	0.387		76	Luau	0.493		109	Caluquemb	0.563		142	Muconda	0.624	
11	Cacuaco	0.150		44	Cela	0.395		77	Catabola	0.496		110	Camanong	0.572		143	Cacolo	0.625	
12	Belas	0.160		45	Negage	0.396		78	Mungo	0.496		111	Humpata	0.572		144	Bibala	0.627	
13	Huambo	0.160		46	Amboim	0.396		79	Ganda	0.497		112	Seles	0.575		145	Cacula	0.630	
14	Namibe	0.169		47	Longonjo	0.402		80	Sanza Pombo	0.500		113	Chitembo	0.575		146	Rivungo	0.633	
15	Tômbwa	0.191		48	Dembos-	0.404		81	Cuanhama	0.504		114	Cambulo	0.575		147	Caungula	0.636	
16	Uíge	0.206		49	Ecunha	0.407		82	N'harea	0.506		115	Caimbamb	0.575		148	Cuilo	0.636	
17	Cacongo	0.210		50	Ukuma	0.414		83	Matala	0.506		116	Buengas	0.583		149	Chicomba	0.639	
18	Malanje	0.212		51	Ambuíla	0.417		84	Canganda	0.507		117	Cahama	0.585		150	Cuito	0.642	
19	Mbanza	0.223		52	Quitexe	0.418		85	Camacupa	0.511		118	Xá-Muteba	0.587		151	Chibia	0.643	
20	Pango-	0.223		53	Samba	0.420		86	Cassongue	0.511		119	Bundas	0.589		152	Luacano	0.644	
21	Soyo	0.229		54	Bolongon	0.420		87	Bula-Atumba	0.514		120	Chongoroi	0.590		153	Luquemb	0.646	
22	Dande	0.243		55	Bailundo	0.425		88	Chipindo	0.517		121	Dirico	0.591		154	Massango	0.650	
23	Lubango	0.272		56	Songo	0.426		89	Ambaca	0.523		122	Quimbele	0.593		155	Lóvua	0.656	
24	Cambambe	0.283		57	Quissama	0.435		90	Puri	0.525		123	Cuemba	0.595		156	Cuchi	0.661	
25	Sumbe	0.294		58	Cunhinga	0.438		91	Maquela do	0.526		124	Nancova	0.603		157	Cambundi	0.661	
26	Chitato	0.307		59	Libolo	0.440		92	Conda	0.529		125	Quipungo	0.603		158	Ombadja	0.663	
27	Lucala	0.319		60	Banga	0.451		93	Damba	0.529		126	Cuvelai	0.603		159	Marimba	0.673	
28	Cuito	0.322		61	Catchiung	0.451		94	Ebo	0.530		127	Cangola	0.603		160	Luchazes	0.683	
29	Golungo	0.337		62	Chinguar	0.452		95	Bungo	0.536		128	Milunga	0.603		161	Quirima	0.684	
30	Ambriz	0.339		63	Londuimb	0.453		96	Quiwaba-	0.537		129	Dala	0.604		162	Virei	0.702	
31	Porto	0.340		64	Nóqui	0.455		97	Bocoio	0.541		130	Cahombo	0.605		163	Camucuio	0.712	
32	Nzetu	0.345		65	Tchinjenje	0.458		98	Calai	0.544		131	Lubalo	0.607		164	Curoca	0.753	
33	Buco Zau	0.365		66	Belize	0.462		99	Nambuangon	0.544		132	Quilengues	0.607					

Adapted from INE - Instituto Nacional de Estatística (2019)

Appendix B – Python implementation of Prim’s algorithm

```
@dataclass(eq=False)
class Vertex :
   idnum : int

@dataclass
class Graph :
   source  : int
   adjlist : dict

   def PrimsMST(self):

       priority_queue = { Vertex(self.source) : 0 }
      added = [False] * len(self.adjlist)
      min_span_tree_cost = 0

      while priority_queue :

         vertex = min(priority_queue, key=priority_queue.get)
         cost = priority_queue[vertex]

         del priority_queue[vertex]

         if added[vertex.idnum] == False :
            min_span_tree_cost += cost
            added[vertex.idnum] = True
            print("Added Vertex : " + str(vertex.idnum) + ", cost now : "+str(min_span_tree_cost))

            for item in self.adjlist[vertex.idnum] :
               adjvertex = item[0]
               adjcost = item[1]
               if added[adjvertex] == False :
                  priority_queue[Vertex(adjvertex)] = adjcost
```

```
        return min_span_tree_cost

def main() :

    g1_edges_from_vertex = {}

    # Outgoing edges from the vertex: (adjacent_vertex, cost) in graph 1.
    # This is a small test adjacency list, not our own.

    g1_edges_from_vertex[0] = [ (1,1), (2,2), (3,1), (4,1), (5,2), (6,1) ]
    g1_edges_from_vertex[1] = [ (0,1), (2,2), (6,2) ]
    g1_edges_from_vertex[2] = [ (0,2), (1,2), (3,1) ]
    g1_edges_from_vertex[3] = [ (0,1), (2,1), (4,2) ]
    g1_edges_from_vertex[4] = [ (0,1), (3,2), (5,2) ]
    g1_edges_from_vertex[5] = [ (0,2), (4,2), (6,1) ]
    g1_edges_from_vertex[6] = [ (0,1), (2,2), (5,1) ]

    g1 = Graph(0, g1_edges_from_vertex)
    cost = g1.PrimsMST()
    print("Cost of the minimum spanning tree in graph 1 : " + str(cost) +"\n")

if __name__ == "__main__" :
    main()
```

Appendix C – Python code for < code_Adjacency list>

```
# adjacency list representation of a Graph in Python
import collections
class Graph:
   def __init__(self):
      self.graph = collections.defaultdict(dict)

   def add_edge(self, u, v, weight = 1, directed = True):
      self.graph[u][v] = weight
      if not directed:
         self.graph[v][u] = weight

   def __str__(self):
      to_return = ''
      for vertex in self.graph:
         to_return  += '    g1_edges_from_vertex[' + str(vertex) + ']=[ '
         for edge in self.graph[vertex]:
            to_return +=  '(' + str(edge) + ', ' + str(self.graph[vertex][edge]) + ')'
            to_return += ', '
         to_return += '] \n'
      return to_return
if __name__ == '__main__':
   g = Graph()
   g.add_edge(1,2638,430675)
   ... <All edges should be added here>
    g.add_edge(4557,4556,337788)

   ###line by line save to file of output in edges and vertices format
   f = open("Adjacency.txt", "a")
   print(g, file=f)
   f.close()
```

Appendix D – Python code for < Pruning_Degree_1 >

```
for x in range (20):

    import networkx as nx


    import matplotlib.pyplot as plt
    x = 0
    G = nx.Graph()
    T = nx.Graph()

    remove_G=[]
    remove_T=[]
    intersection_list=[]



    myscope = {}
    exec(open('./Edges.txt').read())

    #Os negativos seriam as diferenças positivas escolhendo RE#
    ##print(data['weight']) só um teste para aferir se está a escolher pesos negativos##

    for vertex1, vertex2, data in G.edges(data=True):
        if data['weight'] <= 0:


            T.add_edge(vertex1, vertex2,weight=data['weight'])




    #testing if all Edges with negative weight are selected. A new graph T is created#
    ###print ('The vertices with negative weights in graph G (original) are ',T.edges(data='weight'))



    remove_G  = [vertex for vertex,degree in dict(G.degree()).items() if degree ==1]
```

```
    remove_T  = [vertex for vertex,degree in dict(T.degree()).items() if degree ==1]

    intersection_set = set.intersection(set(remove_G), set(remove_T))
    intersection_list = list(intersection_set)

    ##print(G.edges)
    ###print (remove_G)
    ###print (remove_T)
    ###print (intersection_list)

    G.remove_vertices_from(intersection_list)

    ###Clear file
    f = open("Edges.txt", "a")
    f.truncate(0)

    f.close()

    ###line by line save to file of output in adges and vertices format
    for (u, v, wt) in G.edges.data('weight'):

        f = open("Edges.txt", "a")

        print('G.add_edge(%d, %d, weight=%.3f)' % (u, v, wt), file=f)

        f.close()

        ###for (u) in G.vertices():
        ### f = open("Vertices.txt", "a")
        ###print('(%d)' % (u), file=f)
        ###f.close()
for (u, v, wt) in G.edges.data('weight'):

        x = open("Pruned_degree1_edges.txt", "a")
```

```
print('G.add_edge(%d, %d, weight=%.3f)' % (u, v, wt), file=x)

x.close()
```

Appendix E – Python code – Prim’s algorithm for settlement selection

---

```
import networkx as nx
g1 = nx.Graph()

from dataclasses import dataclass, field

# Setting frozen=True and eq=True makes a class immutable and hashable.
# eq=False is needed so that dictionary can contain multiple items with
# the same key (Vertex(idnum)) but with different values (cost)
@dataclass(eq=False)
class Vertex :
   idnum : int

@dataclass
class Graph :
   source  : int
   adjlist : dict

   def PrimsMST(self):

      # Priority queue is implemented as a dictionary with
      # keys as object of 'Vertex' class and value.
      # Since the priority queue will can have multiple entries for the
      # same adjvertex but with different cost, we have to use objects as
      # keys so that they can be stored in a dictionary.
      # [As dictionary can't have duplicate keys so objectify the key]

      priority_queue = { Vertex(self.source) : 0 }
      added = [False] * len(self.adjlist)
      min_span_tree_cost = 0
```

```
while priority_queue :
   # Choose the adjacent vertex with the least edge cost
   vertex = min(priority_queue, key=priority_queue.get)
   cost = priority_queue[vertex]

   # Remove the element from a dictionary in python
   del priority_queue[vertex]

   if added[vertex.idnum] == False :
      min_span_tree_cost += cost
      added[vertex.idnum] = True

      f = open("Prim_cost.txt", "a")
      if min_span_tree_cost <= 180000000:
         print(vertex.idnum, min_span_tree_cost, "2022")
      if ((min_span_tree_cost >180000000) and  (min_span_tree_cost <=360000000)):
         print(vertex.idnum, min_span_tree_cost, "2023")
      if ((min_span_tree_cost >360000000) and  (min_span_tree_cost <=540000000)):
         print(vertex.idnum, min_span_tree_cost, "2024")
      if ((min_span_tree_cost >540000000) and  (min_span_tree_cost <=720000000)):
         print(vertex.idnum, min_span_tree_cost, "2025")
      if ((min_span_tree_cost >720000000) and  (min_span_tree_cost <=900000000)):
         print(vertex.idnum, min_span_tree_cost, "2026")
      if ((min_span_tree_cost >900000000) and  (min_span_tree_cost <=1080000000)):
         print(vertex.idnum, min_span_tree_cost, "2027")
      if ((min_span_tree_cost >1080000000) and  (min_span_tree_cost <=1260000000)):
         print(vertex.idnum, min_span_tree_cost, "2028")
      if ((min_span_tree_cost >1260000000) and  (min_span_tree_cost <=1440000000)):
         print(vertex.idnum, min_span_tree_cost, "2029")
      if ((min_span_tree_cost >1440000000) and  (min_span_tree_cost <=1620000000)):
         print(vertex.idnum, min_span_tree_cost, "2030")



      print("Added Vertex : " + str(vertex.idnum) + ", cost now : "+str(min_span_tree_cost), file=f)
      f.close()
```

```
                for item in self.adjlist[vertex.idnum] :
                    adjvertex = item[0]
                    adjcost = item[1]
                    if added[adjvertex] == False :
                        priority_queue[Vertex(adjvertex)] = adjcost

        return min_span_tree_cost

def main() :

    g1_edges_from_vertex = {}

    # Outgoing edges from the vertex: (adjacent_vertex, cost) in graph 1.

    <New MST adjacency list>

    g1 = Graph(1, g1_edges_from_vertex)
    cost = g1.PrimsMST()

    f = open("Prim_cost.txt", "a")

    print("Cost of the minimum spanning tree in graph 1 : " + str(cost) +"\n", file=f)
    f.close()

if __name__ == "__main__" :
    main()
```