COMILLAS
UNIVERSIDAD PONTIFICIA

ICAI

MASTER THESIS

A GEOSPATIAL APPROACH TO CLEAN COOKING PLANNING IN RWANDA

Autor: Pénélope Schuwer

Director: Pablo Dueñas Martínez

Madrid

Declaro, bajo mi responsabilidad, que el Proyecto presentado con el título

A geospatial approach to clean cooking planning in Rwanda

en la ETS de Ingeniería - ICAI de la Universidad Pontificia Comillas en el curso académico 2022/23 es de mi autoría, original e inédito y no ha sido presentado con anterioridad a otros efectos. El Proyecto no es plagio de otro, ni total ni parcialmente y la información que ha sido tomada de otros documentos está debidamente referenciada.

P Schuwer

Fdo.: Pénélope Schuwer Fecha: 22/08/2023

Autorizada la entrega del proyecto

EL DIRECTOR DEL PROYECTO

Fdo.: Pablo Dueñas Martínez Fecha: 22/08/2023

COMILLAS
UNIVERSIDAD PONTIFICIA

ICAI

MASTER THESIS

A GEOSPATIAL APPROACH TO CLEAN COOKING PLANNING IN RWANDA

Autor: Pénélope Schuwer

Director: Pablo Dueñas Martínez

Madrid

UN ENFOQUE GEOESPACIAL PARA LA PLANIFICACIÓN DE "CLEAN COOKING" EN RUANDA

Autor: Schuwer, Pénélope
Director: Dueñas Martínez, Pablo
Entidad Colaboradora: MIT/Comillas-IIT Universal Energy Access Lab

RESUMEN DEL PROYECTO

En un contexto en el que cocinar sigue suponiendo una grave amenaza para la salud y el medioambiente en el mundo, la electricidad es una de las principales soluciones limpias. El objetivo de este proyecto es, mediante el uso de información geoespacial relevante, definir planes realistas de "clean cooking" en una región determinada. El planteamiento se aplicará a Ruanda, donde se están llevando a cabo importantes iniciativas de "clean cooking".

Palabras clave: Clean cooking, electrificación, datos geoespaciales, Ruanda

1. Introducción

La cocina, esencial en la vida diaria, plantea riesgos para la salud y el ambiente. El "dirty cooking", que usa cocinas y combustibles dañinos, genera humos causando graves enfermedades que afectan sobre todo a mujeres y niños. Generando emisiones de CO_2 y carbono negro, también acelera el calentamiento global y la deforestación. Urge adoptar soluciones más limpias para enfrentar estos desafíos.

Las soluciones de "clean cooking" han evolucionado para ofrecer métodos de cocción domésticos modernos, más sanos y menos contaminantes. La evaluación del "clean cooking" implica dos criterios principales: el aspecto técnico, que evalúa las emisiones de las cocinas, pero también la experiencia del cliente, que valora atributos como la seguridad, la comodidad y la asequibilidad, basándose en datos de los hogares. La electricidad, el GLP y las cocinas modernas son tres motores importantes que pueden ayudar a promover el "clean cooking". Estas soluciones reducen las emisiones y la contaminación, garantizando entornos de cocina más saludables y respetuosos con el medio ambiente.

2. Definición del proyecto

El objetivo principal del proyecto es crear un método para calcular los gastos asociados a la introducción de electricidad para cocinar de forma limpia. Para ello se emplea un modelo de electrificación que requiere el tratamiento de datos de entrada y salida. Después, el proyecto pretende aplicar esta metodología para determinar los costes de la electricidad en un contexto real -Ruanda en este caso- para validar el enfoque utilizando datos reales. Los resultados del análisis se presentarán en forma de mapas de calor que muestran las variaciones de costes en la zona estudiada.

A continuación se presenta un cronograma que muestra la organización de las distintas etapas a lo largo del periodo de tiempo dedicado al proyecto (22 semanas, de febrero a principios de julio 2023).

MONTH	FEBRUARY				MARCH				APRIL				MAY					JUNE				
PROJECT WEEK	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
STEPS	Defining topic		Writing Annex B		Review of litterature		Gathering data			Cleaning and processing data			Running electrification model			Analyzing results			Visualizat ion		Writing report	

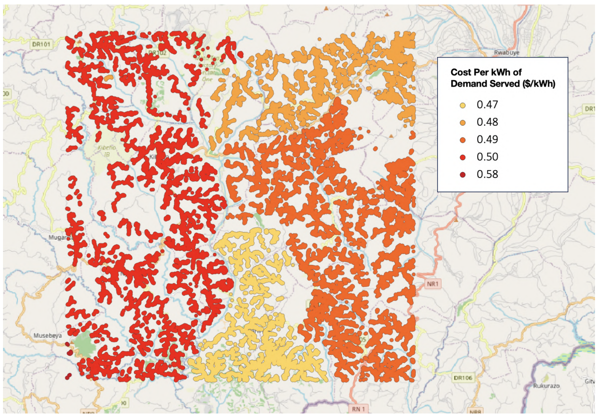
3. Metodología

El paso más crucial del proyecto es la recopilación y preparación de datos. Se ha trabajado con datos geoespaciales específicos de Ruanda para recopilar información detallada sobre edificios, carreteras y líneas eléctricas. Se han investigado diferentes fuentes de datos para nuestra zona de interés, y luego se han limpiado, formateado, analizado e interpretado los datos, utilizando herramientas de Python como Pandas y Geopandas, junto con Jupyter Noteboo y, Visual Studio Code para procesar los datos, así como QGIS para la visualización. Visualizar los datos ha ayudado a comprenderlos mejor y ofrece una perspectiva más amplia. También se han eliminado diferentes fuentes de datos y combinado solapamientos para crear un conjunto de datos completo. El objetivo es preparar los datos para introducirlos en un modelo de electrificación, garantizando resultados precisos en las fases de análisis y modelización.

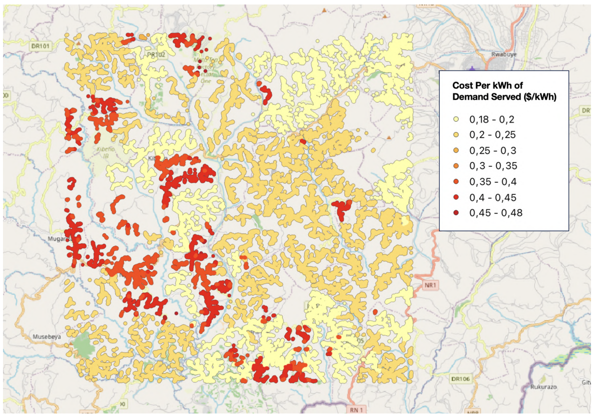
Tras procesar y refinar los datos, el siguiente paso consiste en introducir estos datos preparados en el Modelo de Electrificación de Referencia (REM) del MIT/Comillas-IIT Universal Energy Access Lab. Esta fase es central en el proyecto y cumple varios propósitos clave: en primer lugar, familiarizarse con el modelo y comprender su funcionamiento. Después, permite la experimentación con varios escenarios de entrada, facilitando la exploración de diferentes posibilidades y configuraciones. Finalmente, el proceso proporciona valiosas salidas que pueden explotarse. El análisis de estos resultados y sus visualizaciones proporcionan resultados para elaborar un plan de electrificación.

4. Resultados

Tras procesar los resultados del modelo de electrificación, se pudieron obtener y visualizar resultados significativos, elaborando mapas de los costes de la electrificación en Ruanda, mostrando diferentes indicadores, considerando o no la demanda de cocina eléctrica. El siguiente mapa de calor utiliza gradualmente colores para indicar los costes por kWh de demanda atendida sólo para la demanda de cocina eléctrica, en una zona de interés.



Mapa del coste por kWh de demanda servida por edificio sin la demanda de cocción eléctrica (QGIS)



Mapa del coste por kWh de demanda servida por edificio con la demanda de cocción eléctrica (QGIS)

La comparación de los costes de la electricidad con y sin cocina eléctrica muestra que el uso de la cocina eléctrica reduce significativamente los costes energéticos por unidad y cambia la distribución de los costes. Estos resultados ponen de relieve las ventajas económicas de integrar la cocina en los planes de electrificación y subrayan la importancia de comprender sus repercusiones en las estrategias de electrificación.

Por otra parte, los resultados que muestran el coste acumulado de proporcionar cocina eléctrica a la población ruandesa indican un crecimiento estable hasta el 80% de la población, y luego un pico de coste más allá de este punto, lo que hace difícil cubrir a todo Ruanda. Este dato orienta la planificación de una electrificación rentable del 80% de la población ruandesa, por un importe estimado de 463 millones de dólares. Para cubrir a toda la población, parece necesario explorar diferentes soluciones limpias para cocinar.

5. Conclusiones

Este trabajo aborda los retos de la recopilación, el procesamiento, el análisis y la visualización de datos para elaborar un plan de "clean cooking" en Ruanda. Las visualizaciones creadas para este proyecto desempeñaron un papel crucial en la comprensión de los impactos del desarrollo de la cocina eléctrica en Ruanda.

El trabajo logró con éxito sus objetivos principales: desarrollar una metodología para determinar los costes de electrificación utilizando el REM del MIT/Comillas-IIT Universal Energy Access Lab, y aplicarla a Ruanda, dando como resultado mapas de costes de electricidad. Mejorando la comprensión de la dinámica de costes de electrificación, los resultados destacan el impacto de la cocina eléctrica en los costes energéticos globales y los beneficios de alinear el "clean cooking" con la planificación de la electrificación. Sin embargo, la comparación de las soluciones de "clean cooking" fue incompleta. Queda un gran potencial para seguir investigando otras soluciones adecuadas, como el GLP, y elaborar un plan de "clean cooking" basado en una combinación de diferentes soluciones.

A GEOSPATIAL APPROACH TO CLEAN COOKING PLANNING IN RWANDA

Author: Schuwer, Pénélope
Supervisor: Dueñas Martínez, Pablo
Collaborating Entity: MIT/Comillas-IIT Universal Energy Access Lab

ABSTRACT

In a context where cooking still poses a serious health and environmental threat to billions of people around the world, electricity is one of the major drivers that can help to promote clean cooking. The objective of this project is, by using relevant geospatial information, to define realistic clean cooking plans in a given region. The approach will be applied to Rwanda where relevant clean cooking initiatives are happening.

Keywords: Clean cooking, electrification, geospatial data, Rwanda

1. Introduction

Cooking, a vital aspect of daily life, presents a crucial challenge, posing significant health and environmental risks to billions globally. "Dirty cooking", using harmful stoves and fuels generating toxic smoke, leads to serious health issues, notably household air pollution (HAP)-linked diseases, causing millions of annual deaths, predominantly affecting women and children. Dirty cooking also contributes to CO_2 and black carbon emissions, accelerating global warming and deforestation. To tackle these issues, a shift to cleaner cooking solutions urges.

Clean cooking solutions have evolved to offer modern, healthier, and less polluting household cooking methods. The assessment of clean cooking involves two main criteria: the technical aspect, evaluating the emissions of the cookstoves, but also the customer experience, rating attributes such as safety, convenience, and affordability, based on household-level data. Three major drivers that can help to promote clean cooking are electricity, LPG and modern stoves. These solutions offer reduced emissions and pollution, ensuring healthier and environmentally friendly cooking environments.

2. Definition of the project

The project's primary aim is to create a method for calculating the expenses associated with introducing electricity for clean cooking. This involves employing an electrification model, requiring the processing of input and output data. Then, the project aims to apply this methodology to determine electricity costs in a real-world context—Rwanda in this instance—to validate the approach using actual data. The resulting analysis outcomes are intended to be presented as heat maps showcasing cost variations within the studied area.

Below is a timeline showing the organization of the different steps over the period of time dedicated to the project (22 weeks, from February to beginning of July 2023).

MONTH	FEBRUARY				MARCH				APRIL				MAY					JUNE				
PROJECT WEEK	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
STEPS	Defining topic		Writing Annex B		Review of litterature		Gathering data			Cleaning and processing data			Running electrification model			Analyzing results			Visualizat ion		Writing report	

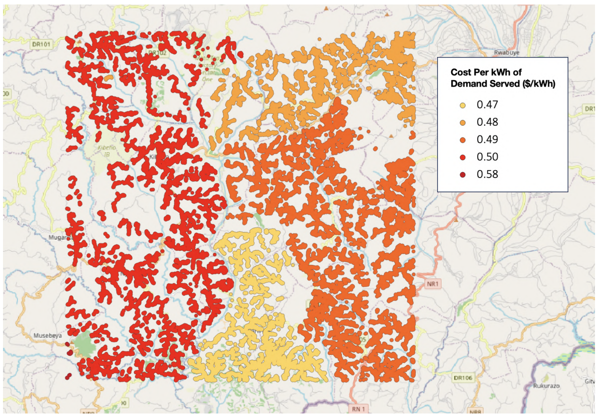
3. Methodology

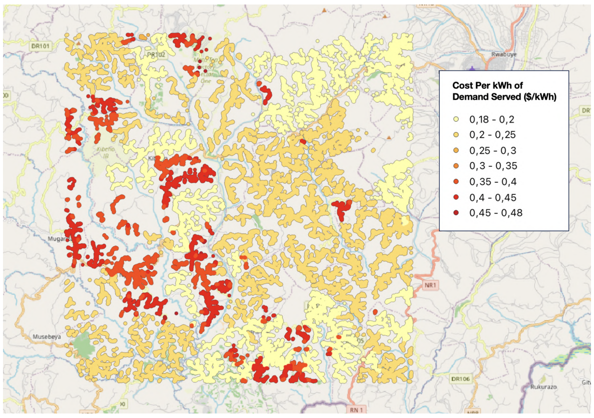
The most crucial step of the project is collecting and preparing data. I have been working with specific geospatial data for Rwanda to gather detailed information about buildings, roads, and power lines. I have researched different data sources for our focus area, then cleaned, formatted, analyzed, and interpreted the data, using Python tools like Pandas and Geopandas, along with Jupyter Notebook, Visual Studio Code to process the data, as well as QGIS for visualization. Visualizing the data helps understand it better, and offers a broader perspective. I also combined different data sources and removed overlaps to create a complete dataset. The goal is to prepare the data for input into an electrification model, ensuring accurate and insightful outcomes in the analysis and modeling stages.

After processing and refining the data, the following step involves inputting this prepared data into the MIT/Comillas-IIT Universal Energy Access Lab Reference Electrification Model (REM). This phase is central in the project and serves several key purposes: firstly, getting familiar with the model and understanding its functioning. Then, it allows the experimentation with various input scenarios, facilitating the exploration of different possibilities and configurations. Ultimately, through the different experimentations, the process provides valuable outputs that can be exploited. The analysis of these outputs and its visualizations gives insights and outcomes to elaborate an electrification plan.

4. Results

After processing the outputs of the electrification model, I was able to obtain and visualize meaningful results, producing maps of the costs of electrification in Rwanda, showing different indicators, both considering the electric cooking demand or not. The following heat map gradually uses colors to indicate the costs per kWh of demand served only for the electric cooking demand, in an area of interest.





Map of Cost per kWh of Demand Served per Building without considering the electric cooking demand

Map of Cost per kWh of Demand Served per Building considering the electric cooking demand (QGIS)

Comparing the costs of electricity with and without electric cooking shows that using electric cooking significantly lowers energy costs per unit and changes the cost distribution, highlighting the financial benefits of integrating cooking into electrification plans. These results provide valuable insights into the economic advantages of electric cooking and emphasizing the importance of understanding its impacts in electrification strategies.

Moreover, the results showing the cumulative cost of providing electric cooking for the Rwandan population indicate stable growth up to 80% of the population, then a cost spike beyond this point, making it challenging to cover everyone in Rwanda. This insight guides planning for cost-effective electrification of about 80% of Rwanda's population, for an estimated amount of $463 million. To cover the entire population, it seems necessary to explore different clean cooking solutions.

5. Conclusion

This master's thesis tackles challenges in data collection, processing, analysis, and visualization to achieve its goals of elaborating a clean cooking plan in Rwanda. The visualizations created for this master thesis played a crucial role in understanding the impacts of developing electric cooking in Rwanda. Maps not only enhance the understanding of the cost dynamics, but also facilitate identifying areas that require particular attention or potential optimization.

The master's thesis successfully achieved its primary objectives: developing a methodology to determine electrification costs using the MIT/Comillas-IIT Universal Energy Access Lab Reference Electrification Model, and applying it to Rwanda, resulting in an electricity cost map. The results emphasize electric cooking's impact on overall energy costs and the benefits of aligning clean cooking with electrification planning. However, the comparison of clean cooking solutions was incomplete. Significant potential remains for further research on other suitable solutions including LPG, and elaborating a clean cooking plan based on a combination of different solutions.

Table of contents

1. Introduction

1.1. Context of the project

The cooking issues

Cooking is an essential part of daily life and yet it poses a serious health and environmental threat to billions of people around the world. According to the Clean Cooking Alliance (CCA), 2.4 billion people are still using what we might call “dirty cooking”: cooking on open fire or inefficient stoves, using biomass fuels such as wood, charcoal, or other polluting fuels like coal or kerosene. These cooking practices generate toxic smoke, black carbon emissions, and other pollutants, endangering the families’ health and well-being and contributing to global warming and deforestation.

Open fires or inefficient stoves put households at risk for serious diseases that can lead to death. According to the CCA, the smoke from cooking is responsible for 4 million deaths each year and particularly affects women and children. It is estimated that 450,000 children die each year due to health issues related to household air pollution (HAP). Health issues associated with HAP include respiratory diseases, childhood pneumonia, chronic obstructive pulmonary disorder, ischemic heart disease, stroke, and lung cancer. Moreover, HAP causes pregnant women exposed to it to have higher blood pressure. Because of that, their babies are more likely to experience stillbirth, low birthweight and gestational age at delivery, and decreased lung functions.

Dirty cooking also has a severe environmental impact. Residential fuel burning represents 1 gigaton of carbon dioxide equivalent every year - which is equivalent to 2% of global CO_2 emissions - and 58% of global black carbon emissions, according to the CCA. We know today that the use of more efficient stoves can reduce fuel consumption by 30-60%. These considerable emissions are heavily contributing to global warming and deforestation.

Furthermore, the cooking issues accentuate gender inequalities as their dramatic consequences are largely endured by women. Indeed, cooking responsibilities are usually held by women and girls, who are more exposed to unhealthy environments due to the more time spent in the kitchen and house.

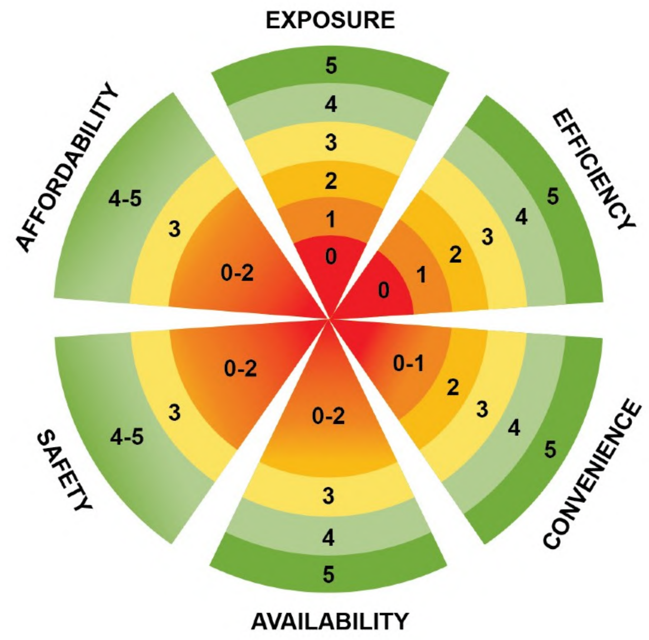
Inefficient cooking also represents great economic loss due to the impacts on health, climate, lost productivity, and deforestation. To limit these costs, large investments and public finance are necessary. As reported by the *Energizing Finance* series, research and analysis conducted by Sustainable Energy for All, USD 150 billion is required annually to reach universal access to more efficient and cleaner cooking solutions by 2030, whereas the cost of using polluting stoves is estimated to be $2.4 trillion each year. In spite of these numbers, the investments and fundings allocated to provide energy access and modern cooking solutions to vulnerable populations are still by far insufficient.

The clean cooking solutions

Clean cooking solutions have evolved over the years to help billions of people access a more modern, healthier, and less polluting way of cooking in their household. Clean cooking involves both the devices and the fuel types used to cook.

How do we measure clean cooking?

On one hand, the World Bank has elaborated the Multi-Tier Framework (MTF) for cooking, a user-oriented approach to cooking energy access that aims to compare and evaluate clean cooking solutions based on household-level data. The MTF for cooking includes six attributes: exposure, efficiency, convenience, safety, affordability, and fuel availability, rated from 0 to 5.



World Bank's Multi-Tier Framework for cooking

A household can be considered to have gained access to modern energy cooking services if it scores on Tier 4 or above on all six attributes. However, intermediate situations are common. The Netherland Enterprise Agency – involved in several programmes to improve access to clean cooking – defines different transitional cooking situations. Indeed, as only Tiers 4 and 5 are considered modern and clean, transitional Tiers 2 and 3 are already considered as an improved access to cooking services.

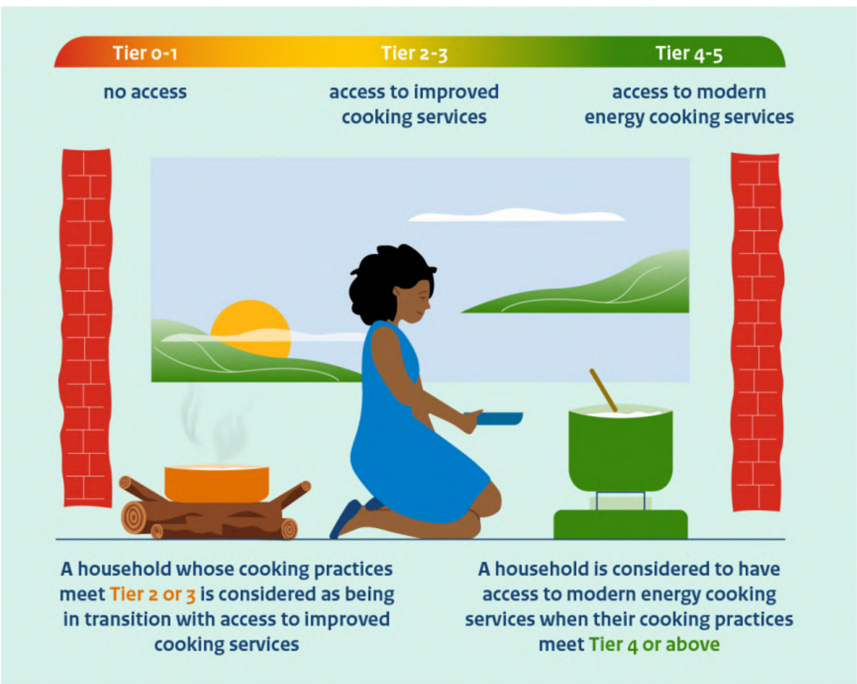


Illustration of the access to cooking services tiers from the The Netherland Enterprise Agency

On the other hand, the World Health Organization (WHO) has established guidelines to qualify a clean cookstove. These guidelines set a carbon monoxide (CO) and $PM_{2.5}$ emission level recommendation. A cookstove can be considered clean if it scores a 4 or a 5 in $PM_{2.5}$ emissions and a 5 in CO emissions.

ISO VPT Tier	WHO Category for CO	WHO Category for $PM_{2.5}$
5	Clean	Clean
4	Transitional	Clean
3	Transitional	Transitional
2	Polluting	Polluting
1	Polluting	Polluting
0	Polluting	Polluting

World Health Organization Guidelines for cooking

To sum up, a clean cooking situation for a household encompasses two important aspects: a clean cookstove functioning with clean fuel, and a high score on the Multi-Tier Framework. Firstly, the cookstove used must score a five on the WHO guidelines mentioned previously. The specific types of stoves with a clean score, as well as the different fuels used to make them work, will be detailed in the next section. Secondly, the Multi-Tier Framework is a comprehensive assessment tool that evaluates whether the household has reliable access to clean cooking, utilizes them regularly, and enjoys the associated health and environmental benefits. Together, a technically clean solution and a high Multi-Tier Framework score ensure a safe and sustainable cooking environment for households, improving health outcomes and reducing environmental impact.

What are concrete clean cooking solutions?

Clean cookstoves are designed to reduce the use of polluting fuels and emissions of smoke and maximize efficiency. According to the Netherland Enterprise Agency, examples of clean cookstoves (respecting the WHO requirements) include LPG stoves, electric stoves using hot plates or induction, electric rice cookers, ethanol stoves, biogas, and specialized biomass stoves. To make clean cooking more accessible, transitional stoves (tier 2-3, which do not meet WHO requirements) can be used as an inclusive and temporary solution. These cookstoves burn biomass more efficiently.

Clean fuels are fuels with very low levels of polluting emissions when used. The Netherland Enterprise Agency lists examples of clean fuels, including:

- Biogas, a mixture of methane, carbon dioxide, and other trace gas, that if produced at large scale, can also be used to generate electricity for other purposes.
- LPG (Liquified Propane Gas), a by-product of crude oil refining and natural gas processing that produces low CO_2 emissions. Unlike natural gas, LPG can easily be liquified and transported in cylinders.
- Electricity, the cleanest fuel when being used. The main challenge is generating it.
- Ethanol, produced from biomass sources and a renewable source of energy.
- Natural gas, a fossil fuel extracted from oil and gas reserves.
- Solar power (BLEENS), a renewable energy source.
- Pellets, a clean fuel if the pellets come from a renewable source of biomass. They have low emissions and a high energy value, but the production and distribution of pellets is challenging.

All of these fuels can be considered clean in the sense that they pollute little when burnt, and therefore can constitute a solution for clean cooking. However, for each of them, it is important to pay attention to how they were generated and delivered, as the pollution from these processes can be significant and make them dirty fuels in some cases.

1.2. State of the art

Main actors promoting clean cooking

The Clean Cooking Alliance

The Clean Cooking Alliance (CCA) is a non-profit organization working to bring clean cooking solutions to those without it, essentially to households in developing countries. The organization was launched in 2010. Supported by the United Nations Foundation, the CCA works with a global network of partners, including governments, businesses, non-profit organizations, and research institutions. It aims to promote clean cooking in developing countries, with the goal of achieving universal access to clean cooking by 2030.

The CCA's work is based on three pillars: raising awareness to the cooking problem in developing countries to help changing behaviors and raise the clean cooking demand;

searching for investment to develop networks to deliver clean cooking solutions to households in need; and developing a favorable environment for industry growth by advocating for enabling policies and providing reliable data.

In 2020, the CCA launched the Cooking Industry Catalyst (CIC) program, which aims to demonstrate the attractiveness and viability of clean cooking businesses. The program has three main components: Venture Catalyst, Demand Catalyst, and Market Catalyst. The Venture Catalyst supports selected companies to help them improve their visibility and facilitate their access to investment and growth capital. The Market Catalyst focuses on creating an environment enabling the development of clean cooking solution businesses. To do so, the CCA supports a wide range of investors and donors, sponsors and hosts relevant events, and helps the development of policies promoting the development of clean cooking. The Demand Catalyst aims to understand the consumers’ situations and needs, and raise clean cooking awareness among populations to drive demand. Thereby, the CIC program’s interventions cover the three core pillars of the CCA’s action and are complementary, employing diverse approaches to promote clean cooking.

The Clean Cooking Explorer tool by the CCA

The CCA, in collaboration with the World Research Institute, has developed the Clean Cooking Explorer, an online, open-source, and interactive spatial platform to facilitate clean cooking access. Through the analysis of geospatial datasets, this platform enables the identification of areas with high expansion potential and others in need of assistance. To understand where the development of clean cooking technology is needed, the CCE combines spatial data, economic factors and sophisticated analysis tools.

The Clean Cooking Explorer (CCE) tool shows three panels to the user. The left panel contains the collection of datasets available for the analysis, the center panel contains the interactive map, and the right panel is used to present the results.

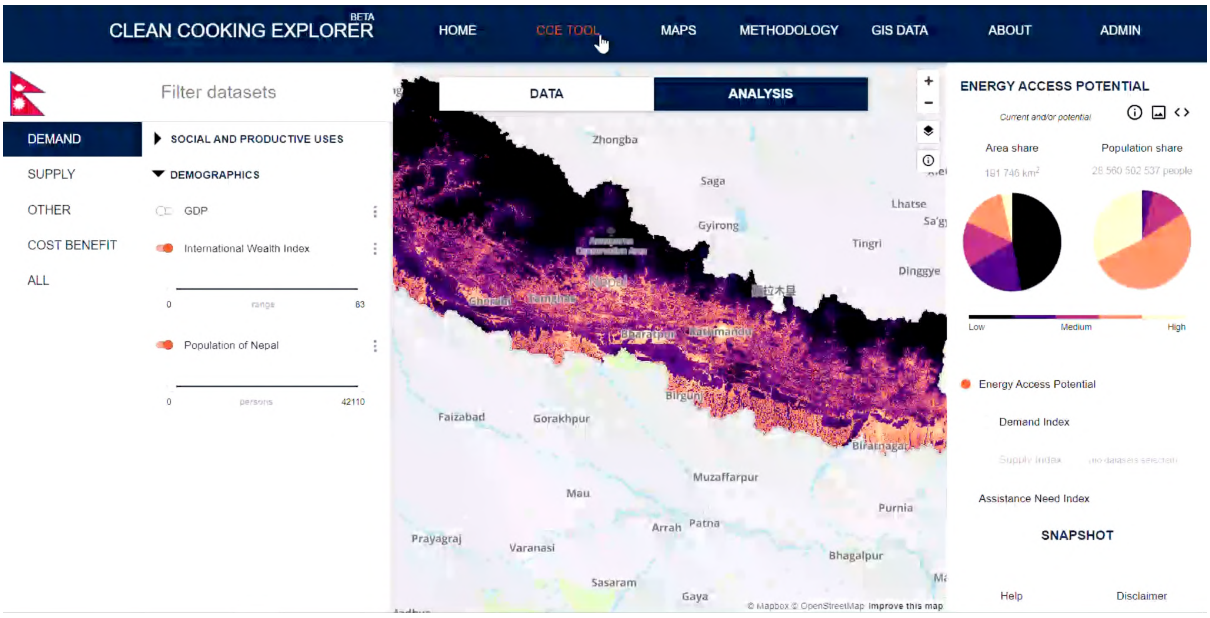


Image of the tool from the CCE demo video

To begin using the tool, in the left panel, you can choose from various layers of data available, such as demographics, social and productive uses, and supply networks. For the supply data, you have the option to specify energy sources like electricity, LPG, or biogas. For each, you can further specify, for example by choosing only MV (medium voltage) lines for electricity. The tool allows you to adjust the importance or weighting of each dataset according to your judgment. This customization feature helps tailor the analysis to your specific needs.

As you add layers of data, they will be displayed on both the map in the center panel and the right panel, where the visualized data will be presented.

To conduct an analysis, you can click on the analysis feature provided by the tool. The results of the analysis will appear on the right panel. They will include different metrics, such as the demand index, which highlights the potential demand in different regions and indicates areas with higher or lower demand. The demand index is calculated by combining all the layers of data chosen. Similarly, the tool also provides a supply index that utilizes the supply datasets to identify regions with better supply potential for different types of stoves or where the supply is more easily accessible. The results of the analysis can be aggregated by population share and area share, providing additional insights into the distribution and impact of clean cooking solutions.

In summary, the Clean Cooking Explorer tool allows users to select and analyze different datasets, visualize them on an interactive map, and generate metrics such as the demand index and supply index to understand the potential demand and supply of clean cooking solutions in various regions. Thereby, Clean Cooking Explorer facilitates planning, coordination, and decision making for policy makers, supply companies, and other stakeholders.

Furthermore, the CCA is currently working with KTH Royal Institute of Technology on the development of OnSSTOVE, a geospatial modeling tool to estimate the benefits of different clean cooking solutions and identify the best option for each region.

Sustainable Energy for All

Sustainable Energy for All is another international organization, working towards achieving access to affordable, reliable, sustainable and modern energy for all (Sustainable Development Goal 7) by 2030, in line with the Paris Agreement on climate. SEforALL is conducting a Clean Cooking Access Programme to raise awareness to cooking issues and to help governments and investors recognize the benefits of clean cooking.

Firstly, SEforALL acts as an advocate, urging governments and partners to increase their commitment and prioritize access to clean cooking solutions.

SEforALL also contributes to informed decision-making and investments in the clean cooking sector. They provide reliable data on consumer preferences and usage patterns regarding different clean cooking solutions. This information helps stakeholders allocate resources to the most viable and effective clean cooking technologies and interventions.

Another crucial aspect of SEforALL's work is quantifying the social, environmental, and economic benefits of clean cooking access. By highlighting the wide-ranging impacts of clean cooking investments, policymakers gain a deeper understanding of the value and importance of prioritizing clean cooking solutions.

SEforALL further supports the clean cooking sector by providing technical input to research projects and initiatives undertaken by stakeholders. Their expertise and guidance contribute to the development and implementation of effective clean cooking strategies.

Additionally, SEforALL actively identifies and fosters opportunities to mobilize finance for clean cooking initiatives. They explore existing or planned initiatives, such as results-based financing, to unlock financial resources and ensure the sustainable growth of the clean cooking sector.

By engaging in these various activities, SEforALL and the Clean Cooking Access Programme play a significant role in promoting clean cooking.

Moreover, in 2020, SEforALL launched the Clean Cooking Data for All project, an initiative that consists in valuing household insights to enhance policy-making and incite clean cooking investment. The project intends to create a pathway for reliable data to travel from their source, the households, up to the information systems of decision-making. The project offers a new approach to clean cooking, offering valuable insights on what customers need and like, which will guide decisions for investors, enterprises and governments in direction that corresponds to the actual demand.

The World Bank

The World Bank is also taking action to support clean cooking, essentially by financing projects and raising awareness in developing countries. The World Bank is currently lending over $350 million in clean cooking and heating projects across 21 countries, helping over 3.6 million households and 18 million people.

The World Bank has also established a partnership with 24 partners: the Energy Sector Management Assistance Program (ESMAP): the aim is to help low and middle-income countries reduce poverty and boost growth through sustainable energy solutions. In 2019, ESMAP launched its $500M Clean Cooking Fund, with important contributions from the Netherlands, Norway, and the UK. This fund aims to scale up investments in the clean cooking sector, and is expected to generate US$2 billion.

By scaling up its investments in the clean cooking sector, mobilizing political commitments, and promoting innovations, the World Bank has become a major actor in the development of clean cooking in the past several years.

Clean cooking in Rwanda

Rwanda is also a key partner in the clean cooking transition. Today, charcoal and wood are the main sources of energy for cooking, responsible for major health and deforestation issues. According to KOKO Networks, a clean ethanol fuel and stove company operating in different countries in Africa including Rwanda, the household air pollution from dirty cooking fuels leads to more than 7,000 deaths annually, and over 50% of these deaths occur in children under the age of five. Upgrading Rwandan households to modern sustainable cooking fuel is an urgent priority for family health. Moreover, the demand for charcoal in Kigali, the capital city, has resulted in the destruction of over 19,000 hectares of forest each year. These numbers highlight the urgent need for cleaner cooking solutions to address both the significant health risks and the environmental degradation associated with traditional cooking practices in Rwanda. According to the Clean Cooking Alliance, Rwanda aims to achieve universal access to clean cooking by 2030.

The Energy Access Quality Improvement Project in Rwanda is the first Clean Cooking Fund project, started in 2020. The project is expected to expand access to clean cooking to 500,000 households across Rwanda and leverage US$30 million in investment.

According to the government institution Rwanda Development Board, in 2022, the Government of Rwanda and KOKO Networks have reached an agreement to develop a nationwide renewable cooking fuel utility. Households will be able to purchase the KOKO Cooker & Canister, a bioethanol cooking solution, designed by working and experimenting with families.

Photo of the KOKO Cooker & Canister

The government is in charge of creating an enabling policy environment including removal of value-added tax and import duties on equipment and ethanol fuel, while KOKO Networks will provide the necessary technology and expertise to make this project successful. This agreement is an important step in the effort to provide clean and renewable cooking solutions to households.

Moreover, one of the Rwandan government's key priorities is to increase access to electricity. To achieve this, they have been making significant investments in electrification projects. As a result, according to the World Bank, by 2020, 46.6% of Rwandan households had access to electricity, while only 15.2% of them had access in 2013. Access to electricity is improving the households' health and standards of living, including by bringing a clean cooking solution.

1.3. Motivation

Bringing clean cooking solutions to developing countries is a crucial matter driven by the need to address a range of serious health, environment, and gender issues related to the wood-based and traditional cooking methods currently used. Still today, the majority of Rwandans rely on inefficient, polluting, and dangerous fuels for cooking. This has led to high levels of indoor air pollution and health risks, including respiratory and cardiovascular diseases, as well as a range of environmental problems, including deforestation, land degradation, and climate change. Additionally, the reliance on solid fuels has tended to disproportionately burden women, who are often responsible for gathering firewood and managing the cooking process.

Moreover, the government of Rwanda's recent initiatives in collaboration with international organizations include promising projects intending to expand access to clean cooking and achieve universal electrification in Rwanda. This hopeful action opens up encouraging perspectives.

In line with these initiatives, this project examines different clean cooking solutions, focusing on electricity as part of the electrification effort that is already underway. Through geospatial analysis and modelization, the study intends to evaluate the cost of delivering reinforced electricity to the different areas of the country. This geospatial approach to clean cooking in Rwanda is in the continuity of the recent initiatives that have been launched in the past few years, but also intends to offer new methods and purposes to clean cooking planning.

1.4. Objectives of the project

The main objective of this project is to elaborate a methodology to determine the cost of implementation of electricity for clean cooking. To obtain these numbers, the use of two models for electricity will be necessary. In order to run these two models, the processing of the input and output data is an important part of the project.

The second objective is to apply the methodology developed to evaluate costs of electricity (for the previous objective) to a real and specific area. In this project, we are applying our model to Rwanda. This allows us to put our methodology to the test, using the realistic value of an actual country. In the end, the results of this analysis should take the form of heat maps showing the variation of costs over space inside a studied area.

The last objective of this project is to come up with a methodology to define the best clean cooking solution, depending on the area taken into consideration. The best solution for a specific area depends not only on the price estimated in the first objective, but also on the solution that has been settled for in other areas around. In some areas, we could imagine a combination of electricity and other fuels, such as LPG. In this case, the proportions would have to be defined.

1.5. Alignment with sustainable development goals

Looking to implement healthy and environment-friendly solutions to everyday life matters, this project aligns with many sustainable development goals:

Affordable and clean energy

Expanding access to electricity ensures sustainable energy, for cooking as well as for other household activities.

Good health and well-being

Cooking with electricity or low-emission fuels reduces diseases, due to less exposition to smoke and polluted air inside the household.

Climate action

By reducing the use of polluting fuels and emissions of smoke and maximizing efficiency, clean cooking reduces climate damage.

Gender equality

Clean cooking reduces the gender inequality in health issues, due to the exposure to harmful smoke while cooking and spending time in the house.

No poverty

Delivering access to modern cooking to households helps consumers save time and money.

Sustainable cities and communities

The use of clean cooking ensures better household and ambient air as well as a better global resource efficiency.

Decent work and economic growth

With the development of clean cooking comes job creation and inclusive economic growth.

Industry, innovation, and infrastructure

Delivering electricity to the different regions of Rwanda will allow the development of electricity networks and infrastructures.

1.6. Structure of document

Introduction

The document follows a well-organized structure, beginning with an introduction that sets the stage for the project. The introduction consists of several sections, starting with "Context of the project," which provides a brief overview of the project's background and the environment in which it operates. By explaining the cooking issues worldwide, clearly defining clean cooking, and mentioning concrete clean cooking solutions, this section aims to provide necessary context and establish the relevance of the project within a broader framework.

Following that, the "State of the art" section delves into the existing knowledge and advancements related to clean cooking. It offers a comprehensive review of the current state of research, technologies, and practices in the field, mentioning encouraging initiatives that are taking place in developing countries worldwide, as well as describing the current context in Rwanda. This section helps establish the project's position in relation to the existing cooking situation in developing countries.

The "Motivation" section comes next, explaining the driving factors behind the project. It outlines the reasons and incentives for undertaking the research or development initiative, emphasizing both the importance of developing clean cooking and reducing health and environmental issues, and the encouraging context developing in Rwanda thanks to recent promising initiatives.

The "Objectives of the project" section outlines the specific goals that the project aims to achieve, including developing a methodology to determine the cost of implementation of electricity for clean cooking, and applying this methodology to a specific area of study, Rwanda.

Next, the document addresses the project's alignment with sustainable development goals. This part demonstrates how the project contributes to broader sustainability objectives, such as expanding access to clean energy, reducing health issues related to HAP, or reducing pollution fuels emissions and therefore climate change.

Methodology

Moving on, the document transitions into the methodology section, which describes the approach and techniques used to conduct the project. The "Data collection" subsection explains how relevant data was gathered, giving detail on the information, format, and source of all the data, ensuring transparency and reproducibility.

The "Data processing" subsection outlines the steps taken to clean, analyze, and interpret the collected data. It includes parts about understanding the data, as well as preprocessing, cleaning, and transformation of different datasets in order to make better use of the collected data.

The document then introduces the "Reference Network Model" (RNM) and the "Reference Electrification Model" (REM), which serve as key tools for the project. These sections provide global information about the use of these two models as well as detail on their inputs and outputs.

The "Preparing the REM input data" subsection follows, elaborating on the steps taken to prepare the input data for the REM. This section details specific processing of datasets in order to obtain the specific data and format needed to input in REM, emphasizing the importance of data accuracy in ensuring reliable and meaningful results.

Results

The document proceeds to the results section, where the outputs of the REM are described in detail in the "Description of the outputs" subsection. This subsection presents a comprehensive overview of the outputs generated by REM.

The "Analyzing results" subsection then delves into the finding of the project's execution, and interpretation and analysis of the obtained results. It explains the processing of the outputs of REM in order to analyze them, then provides different types of visualization of results and discusses the main information discovered in the data, providing insights and drawing conclusions based on the analysis of the output of REM.

Conclusion

The conclusion section follows, summarizing the main challenges, issues, and findings of the project. It goes back to the objectives of the project to conclude on the success of the study. This section wraps up the overall outcomes and significance of the research, as well as the potential future directions of the project

References

Finally, the document concludes with a list of references, providing citations and acknowledgments for the sources and materials used throughout the document.

2. Methodology

The purpose of this section is to describe the detailed methods used throughout the different steps of the project. This includes the data collection methods, as well as the tools and techniques used to process the data collected and analyze the results. The processed data is injected into the electrification model to determine the costs of delivering these clean cooking solutions. This part also describes these two models.

2.1. Data collection

The data used in this project is geospatial data in the area of Rwanda. The goal is to find information on the Rwandan buildings, roads, and power lines. I have studied the different sources and types of geospatial data available for our area of interest.

I collected geospatial data concerning the following elements in Rwanda, from various open sources:

Data	Sources
Buildings	OpenStreetMap, Google Open Buildings, Humanitarian Data Exchange
Roads	OpenStreetMap
Power lines	Zenodo, dataset created by Christopher Arderne, Conrad Zorn, Claire Nicolas, and Elco E. Koks
Settlements	Geo-Referenced Infrastructure and Demographic Data for Development
Provinces boundaries	Humanitarian Data Exchange
District boundaries	Humanitarian Data Exchange
Rwandan border	Natural Earth

Two major sources of data for buildings are available in Rwanda: OpenStreetMap and Google Open Buildings. As both datasets include a non-exclusive record of buildings, I chose to use the data from both sources, and later merge them.
OSM is an open, free, collaborative geographic database that contains data about buildings, roads, shops, railways, and much more, all over the world. OSM offers to download all data for a specific country, which includes a shapefile with only buildings. I downloaded this dataset for Rwanda.
Google Open Buildings is a large-scale open dataset with outlines of buildings derived from high-resolution satellite imagery. It offers to download the data from a rectangle area of your choice (of around 4 degrees latitude and 6 degrees longitude), as a csv file. I downloaded the data for the area including Rwanda.
Additionally, the Humanitarian Data Exchange, an open-data platform managed by the United Nations Office for the Coordination of Humanitarian Affairs, provides a geographic dataset with information on health facilities, including geolocation with a 5 to 10 meters accuracy. I collected this data for the area of Sub-saharan Africa, as they give us additional information on buildings. They might give us more precise information about buildings related to health, or even give us the location of buildings that are not registered in the OSM or Google datasets.

Two different sources have open data on roads in Rwanda: OSM and the Global Biodiversity model for policy support (GLOBIO), developed by the PBL Netherlands Environmental Assessment Agency. GLOBIO provides an open dataset of roads in the form of one shapefile for the continent of Africa. I downloaded that data from GLOBIO, and compared it to the roads shapefile from OSM (I had the entire dataset available for Rwanda that I had downloaded with the buildings). I was able to compare the two datasets by visualizing them in QGIS, and observed two things: the OSM dataset was more complete and included more roads, and the GLOBIO data was in a different projection system. Therefore, I chose to continue working with the OSM data for roads.

Screenshot from QGIS showing a portion of the roads datasets from OSM (orange) and GLOBIO (blue)

I also collected data concerning power lines in Rwanda. Zenodo, an open repository developed under the European OpenAIRE program, includes an open source dataset created by Christopher Arderne, Conrad Zorn, Claire Nicolas, and Elco E. Koks predicting the location of electricity network lines, using night-time lights, satellite imagery, and OpenStreetMap data. I downloaded the shapefile with the power lines data for the entire world.

Data on the settlements in Rwanda was provided by the Geo-Referenced Infrastructure and Demographic Data for Development (GRID3). The dataset is a shapefile for Rwanda that consists of polygons representing areas where there is likely a human settlement based on the presence of buildings detected in satellite imagery, with population estimates. The settlements are categorized in three types: built-up areas (BUA), small settlement areas (SSA), and hamlets. The settlement types will later be used to identify types of buildings. For example, buildings located in hamlets will most likely be residential.

Finally, I needed the administrative boundaries of Rwanda, its provinces, and its districts. Rwanda is divided into five provinces, based mostly on geography: the Northern Province, the Southern Province, the Eastern Province, the Western Province, and the Municipality of Kigali in the center. These five provinces are themselves divided in thirty districts. I downloaded the data containing the geography of these provinces and districts from the Humanitarian Data Exchange datasets. I then collected the population of each province from the City Population database, a population statistics website offering maps and charts for cities, agglomerations and administrative divisions of all countries of the world. Natural Earth, a public domain map dataset, provides geospatial data for borders of countries, from which I downloaded a dataset with the entire world.

The data collection phase was satisfying as I was able to download data concerning all of the major elements of interest for the area of Rwanda, that came from reliable sources, and was in a format easy to exploit (mostly in shapefile format, and if not, in csv). However, open data collaborative sources as OSM and others are not complete. For example, I was missing buildings, and lots of fields giving information (such as the type of building) about the buildings that were in the database were empty.

The collected data will then be analyzed in order to get information on Rwanda's geospatial organization and existing infrastructure. It will be processed in order to make the transformation of format and content necessary to be inputted in the electrification and LPG models.

2.2. Data processing

This part explains the methods I decided to use to analyze and interpret the data. It includes information on the tools used, and data cleaning and analysis techniques employed.
The main resources I used were:

- Python 3 to develop the project, especially to manipulate data, working with the Pandas and Geopandas libraries.
- Jupyter Notebook for computing.
- Visual Studio Code for code editing.
- QGIS for visualization of geographic data.

Data reading and preprocessing

The first step was to open the data in a Jupyter Notebook to discover what it looks like: fields, size, missing data, etc. I used the Python library Geopandas to work with geospatial data. The shapefile were already formatted as geographic data and could directly be read as GeoDataFrames by Geopandas, whereas csv files had to be read as DataFrames by Pandas, and then transformed into GeoDataFrames. To do that transformation, I changed the geometry column, that was a string object containing a WKT (Well-Known Text) representation of a geometry, into a geometric object, using the function wkt.loads from the Shapely library. Then, I was able to create a GeoDataFrame from the DataFrame that included a geometry column (with geometric objects), using Geopandas' GeoDataFrame function.

OPEN BUILDINGS

```python
df_openbuildings = pd.read_csv('/Users/penelopeschuwer/Documents/TFM jupyter/DATA/buildings/open_buildings/19d_buildings.csv')
df_openbuildings.head()
```

	latitude	longitude	area_in_meters	confidence	geometry	full_plus_code
0	-2.902912	30.455017	30.7198	0.6499	POLYGON((30.4550459396869 -2.90293260058396, 3...	6G9G3FW4+R2JG
1	-3.071905	29.126861	36.2631	0.7298	POLYGON((29.1268977863354 -3.07192477607673, 2...	6G8FW4HG+6PQP
2	-2.116668	29.403169	51.6235	0.8392	POLYGON((29.4032125947254 -2.11668346461412, 2...	6G9FVCM3+87M7
3	-2.042385	30.577078	24.7522	0.8112	POLYGON((30.5770980251588 -2.04241052193389, 3...	6G9GXH5G+2RWM
4	-2.522023	30.459097	7.3252	0.6330	POLYGON((30.4591086751409 -2.5220357640513, 30...	6G9GFFH5+5JX2

```python
len(df_openbuildings)
```

```
20065296
```

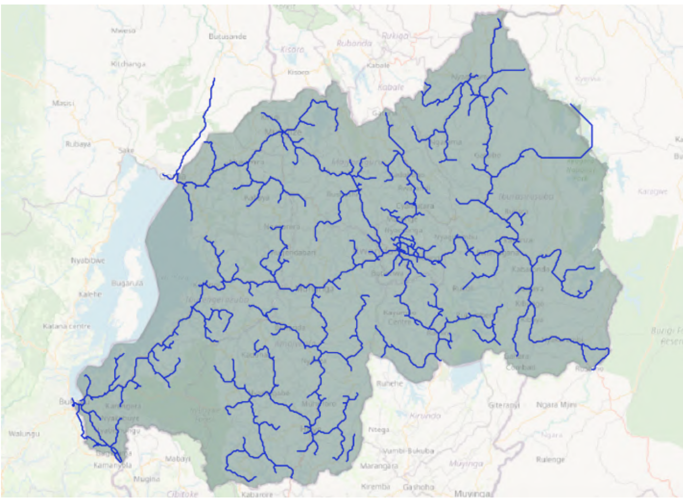
```python
df_openbuildings['geometry'] = df_openbuildings['geometry'].apply(wkt.loads)
gdf_openbuildings=gpd.GeoDataFrame(df_openbuildings, crs="EPSG:4326")
```

Code to create a GeoDataFrame from a csv with Python

Visualizing all data in a Jupyter Notebook and in the same format - GeoDataFrames - allowed me to better understand the data I was dealing with and gain a better perspective on it. This phase helps guide subsequent steps, such as figuring out which type of processing will have to be applied and which kind of insights can be drawn from it.

Another way to better understand the data is to visualize it on a map. I used QGIS as my main tool for mostly viewing but also editing the data.

In order to more fully comprehend the data, I also decided to use QGIS to visualize it on a map. QGIS is a free, open-source application that supports viewing, editing, printing, and analysis of geospatial data. This helped me better understand the data, as the visual representation allowed me to see the actual geometries and projections of the geographic objects, and to have a look at the dataset as a whole, and not row by row like in the Jupyter Notebook. Below is an example for the power lines dataset.



Visual representation of the Rwandan power lines dataset in QGIS

I also cleaned the data, keeping only the rows concerning our area of interest, Rwanda. From the Natural Earth dataset with countries' borders, I saved only the geometry of the Rwandan territory into a new GeoDataFrame in my Notebook, and shapefile in my computer.

```
rwanda = countries[countries["SOVEREIGNT"]=="Rwanda"][['SOVEREIGNT', 'geometry']]
rwanda.to_file('rwanda_border.shp')
```

Code to save the geometry of Rwanda

Then, I joined this Rwanda GeoDataFrame with the others to keep only the data concerning our area of interest. Below is an example of the transformation, by realizing a spatial join of the GeoDataFrame with the health facilities of Sub-saharan Africa with the Rwanda GeoDataFrame. By default, the parameter "predicate" of the function is set to "intersects", meaning that I kept all the buildings with geometries that intersect with the geometry of Rwanda. I repeated that transformation for all data.

```
#Create gdf from shp
gdf_health = gpd.read_file('/Users/penelopeschuwer/Documents/TFM jupyter/DATA/health facilities/suhsharan_health_faciliti

#Keep only geometry inside Rwanda
gdf_health_rwanda=gpd.sjoin(gdf_health, rwanda)

#Save as shp
gdf_health_rwanda.to_file('/Users/penelopeschuwer/Documents/TFM jupyter/DATA/health facilities/health_facilities.shp')
```

Code to create a roads GeoDataFrame with only the data for Rwanda

Processing and transformation of the buildings datasets

Merging data from different sources

The main part of the processing phase dealt with the buildings datasets. The first step was merging the different sources of data: OpenStreetMap, Open Buildings, and the Humanitarian Data Exchange for health facilities. When a building was in both datasets, I had to remove it from one of them. The method that I chose was to delete all buildings from the OSM dataset that overlapped with a building from the Open Buildings dataset, using a spatial join on intersection. This means that for any row of the OSM dataset, if the polygon of the building intersects with any polygon of the Open Buildings dataset, that row will be deleted from the OSM dataset. After completing that operation, I concatenated the new OSM dataset with the Open Buildings one. Below is the detailed code used for this transformation.

MERGING OSM AND OPENBUILDINGS

```
#join on intersection to identify osm buildings that overlap
gdf_buildings_joined=gpd.sjoin(gdf_openbuildings_rwanda, gdf_osm_rwanda)

#list of osm buildings to delete = osm_id of all the osm buildings that overlap with open buildings
osm_buildings_to_delete=gdf_buildings_joined.osm_id.unique()
```

```
#creating a function that returns whether a row has to be deleted or not
def to_del(id, i):
    print("0")
    if id in osm_buildings_to_delete: #if the osm_id is in the list of id to be deleted
        return True
    return False
```

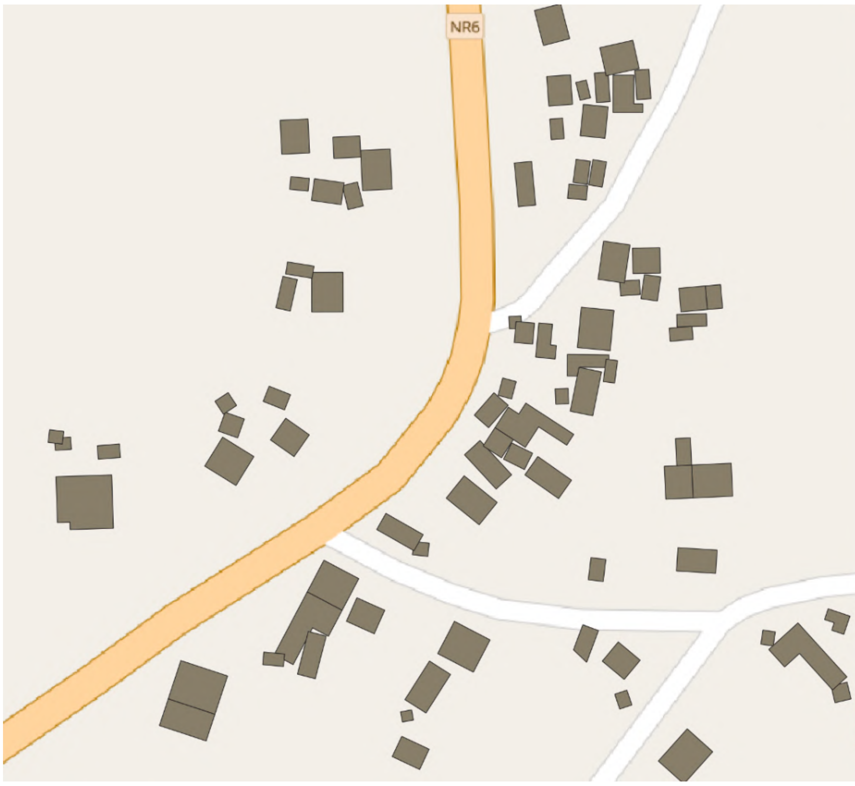
```
#Creating new gdf without the overlaping buildings
gdf_osm_rwanda_cleaned=gdf_osm_rwanda.copy(deep=False)

#Applying to_del function to each row of gdf
gdf_osm_rwanda_cleaned['del'] = gdf_osm_rwanda_cleaned.apply(lambda row : to_del(row['osm_id'], row.index), axis = 1)

#Creating new gdf with all buildings
gdf_osm_rwanda_cleaned=gdf_osm_rwanda_cleaned[gdf_osm_rwanda_cleaned['del']==False] #deleting overlapping rows
gdf_buildings=pd.concat([gdf_openbuildings_rwanda, gdf_osm_rwanda_cleaned]) #concat openbuildings and osm cleaned
gdf_buildings.to_file('buildings_rwanda.shp') #saving as shp
```

Code to merge the OSM and Open Buildings dataset without overlapping

In the same way, I merged the new buildings GeoDataFrame with the health facilities GeoDataFrame, from the Humanitarian Data Exchange. Using another Geopandas spatial join, I was able to delete all the health facilities which had a geometry that intersected with any of the geometries of the other buildings. Then, I concatenated the GeoDataFrame with remaining health facilities and the one with the other buildings. In the end, I obtained a shapefile containing 5.7 million buildings.

Cleaning the resulting dataset

I opened the new shapefile in QGIS, with all the buildings in Rwanda merged from my three sources (Open Buildings, OSM, and HDX). By zooming in to see the shape of buildings, I noticed that buildings were still overlapping.



Buildings in QGIS, from merging Open Buildings, OSM, and HDX datasets

I had removed all the buildings from one source that overlapped with a building from another source. This meant that within a dataset coming from a unique source, buildings also overlapped.

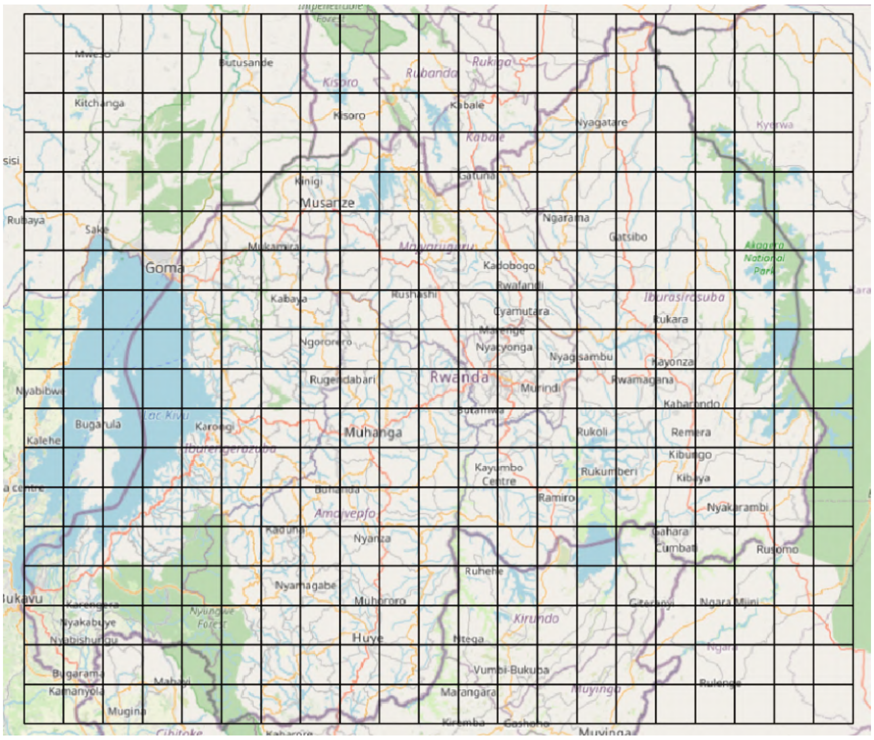
To remove these overlapping buildings, I had to compare the geometry of each building to the 5.7 million geometries of the other buildings. As it was a very heavy operation, I divided the area of Rwanda into smaller squares, in order to compare each building only to the buildings of the same area. To create the grid, I used the total_bounds function, which gives the bounds of the geometry of a GeoSeries as a whole, applied to the GeoSeries containing all the buildings of Rwanda. Then I divided the outer rectangle into squares of 0.1 degree longitude by 0.1 degree latitude, which gave me a grid containing 378 squares (some of them being outside of Rwanda and therefore empty), as shown in the image below.

```
gdf_buildings_geom.geometry.total_bounds
✓ 2.7s

array([28.86721175, -2.8232108 , 30.87265839, -1.0514816 ])
```

```
li_squares = []
li_id = []
id=0
for x in np.arange(28.86721174, 30.87265840, 0.1):
    for y in np.arange(-2.8232109, -1.0514815, 0.1):
        li_id.append(id)
        li_squares.append(Polygon([(x, y), (x+0.1, y), (x+0.1, y+0.1), (x, y+0.1)]))
        id=id+1
✓ 0.0s
```

```
squares = gpd.GeoSeries(li_squares)
d = {'id': li_id, 'geometry': squares}
gdf_squares = gpd.GeoDataFrame(data=d, crs="EPSG:4326")
✓ 0.0s
```

Code to create a grid dividing Rwanda into smaller squares



Visualization of the grid in QGIS

Then, I assigned each building to the square it was in, using a spatial join with the predicate “within”, so that a building could not be assigned to two different squares. Using this method, buildings which were on the border between two squares were removed. In this operation, I lost about 8,000 buildings out of the 5.7 million, which I considered acceptable.

```python
gdf_buildings_geom_square = gpd.sjoin(gdf_buildings_geom, gdf_squares, predicate="within")
```
✓ 58.1s

Code to assign each building to its corresponding square

Then, inside each square, I compared the geometry of each building to each of the other geometries of the buildings of the area, in order to remove overlapping buildings. For each square, as shown in the code below, I created a function to_intersects that returns a boolean indicating whether a geometry intersects with any of the following buildings of the corresponding GeoDataFrame. This means that if two buildings intersect, only the first one appearing in the GeoDataFrame will remain. I applied to_intersects to each row of the GeoDataFrame of the square and stored the result in a new column, named interects. Then I created a new GeoDataFrame, "gdf_buildings_geom_cleaned", containing only the rows where the "intersects" column indicated False, meaning that the building did not intersect with any of the following. Finally, I concatenated the GeoDataFrame to the global one containing the geometries of the buildings for all the squares.

```python
for sq in buildings_per_square.index[1:]: #square number

    print(sq) #square number

    #create gdf of buildings within square
    gdf_buildings_geom = gdf_buildings_geom_square[gdf_buildings_geom_square['id']==sq]
    gdf_buildings_geom = gdf_buildings_geom.reset_index()
    gdf_buildings_geom=gdf_buildings_geom.drop(columns=['index', 'index_right'])
    gdf_buildings_geom['id_building'] = pd.Series(range(0,len(gdf_buildings_geom)))

    #creating a function that returns whether a geometry overlaps with another or not (bool)
    def to_intersects(geom, i):
        li =  gdf_buildings_geom.geometry.iloc[i+1:].intersects(geom) #list of bools, true if geom intersects with following geometries
        for el in li:
            if el==True:
                return True
        return False

    #apply to_intersects function to each row
    gdf_buildings_geom['intersects'] = gdf_buildings_geom.apply(lambda row : to_intersects(row['geometry'], row.id_building), axis = 1)

    #keep only rows that do not intersect with other
    gdf_buildings_geom_cleaned = gdf_buildings_geom[gdf_buildings_geom['intersects']==False]

    #concat with gdf rwanda
    gdf_buildings_geom_cleaned_rwanda = pd.concat([gdf_buildings_geom_cleaned_rwanda, gdf_buildings_geom_cleaned])
```

Code to remove all the overlapping buildings

As a result, I obtained a shapefile containing 4.6 million buildings, meaning I had removed more than a million overlapping buildings. I viewed the dataset in QGIS and, by comparing it with the initial dataset, I observed that many buildings had been removed, and more importantly that buildings did not overlap anymore. Two views of the dataset in QGIS are displayed below.

Cleaned version of the view shown in Figure #

Comparison between the initial and cleaned datasets in QGIS

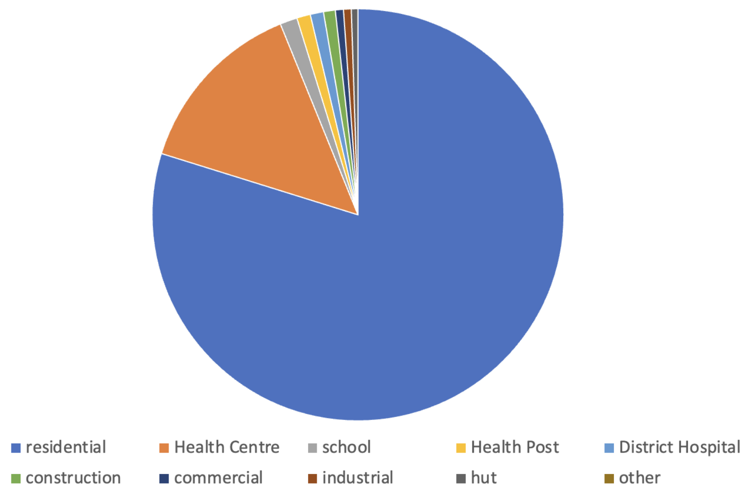
As a result, I successfully merged data obtained from various sources, resulting in the creation of a complete dataset containing the geometries of all buildings in Rwanda. Furthermore, I employed diverse methods, such as spatial joins and the implementation of a grid system, to meticulously clean the dataset. By using these techniques, I ensured that the information was accurate, precise, and nearly exhaustive.

Exploratory analysis of the types of buildings and settlements

Exploring the types of buildings

Inside the new buildings database, I focused on the types of buildings. The Open Buildings dataset did not have a field for the type. Only the OSM and health facilities datasets did, and I merged them into a single column in the new GeoDataFrame. Only 3,500 buildings out of almost six million had a type.

Below is a representation of the proportions of the types of buildings:

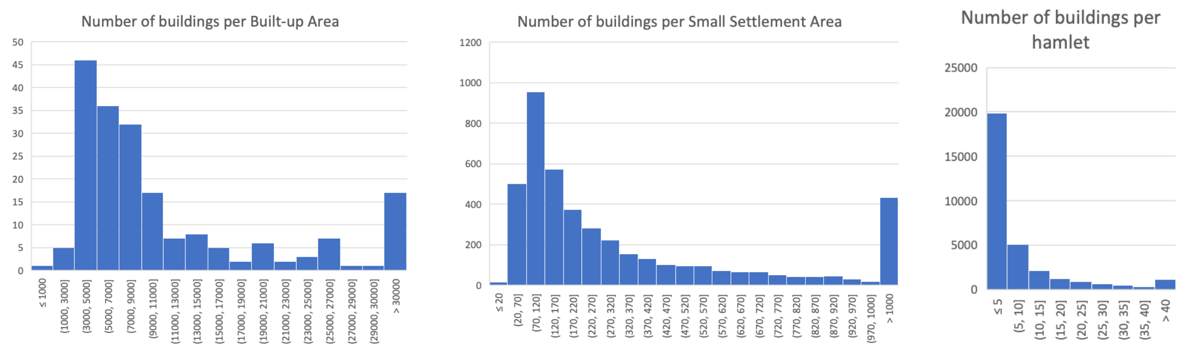


Pie chart of types of buildings

This chart gave me an idea of the different types of buildings that were represented in the dataset. However, with less than 1 building with type out of 1,000, the proportion did not give enough information to conclude anything.

Joining the buildings with the settlements

I joined the buildings with the settlements, in order to have more information on each building. I used a geopandas spatial join, but this time using the predicate “within”, to make sure that buildings were only assigned to one settlement. I looked at how many buildings were in a settlement, for each type of settlement: built-up area, small settlement area, and hamlet. Below are three histograms, one for each type of settlement, showing the number of buildings (from the buildings dataset) per settlement.



Histograms representing the number of buildings per settlement for the 3 types of settlement

From these graphs, I observed that built-up areas generally included over 2,000 buildings, small settlement areas between 40 and 1,000 buildings, and hamlets less than 20. I also determined the median number of buildings per settlement for each type: 7594 for built-up areas, 189 for small settlement areas, and 4 for hamlets. I assumed the hamlets were all residential. In order to fill in some of the missing type values, I chose to assign the residential type to all buildings located in a hamlet.

I also identified the settlements with the most buildings with a type, in order to find areas that could be representative of the proportions of types of buildings in Rwanda, and potentially extrapolated to the rest of the country.

The settlement with the most buildings with type (674 out of more that 900,000 buildings) was a built-up area located in the South of Rwanda. However, I looked at the types of buildings, and it turned out that this settlement had only residential buildings registered. I found another settlement, another built-up area in the South of Rwanda, close to the border with Burundi, with a more interesting range of building types. It had 18 commercial buildings out of 103 buildings with a type. I chose to focus on this area, as it seemed to have a rather high density and variety of buildings with type.

Definition of a square area of interest

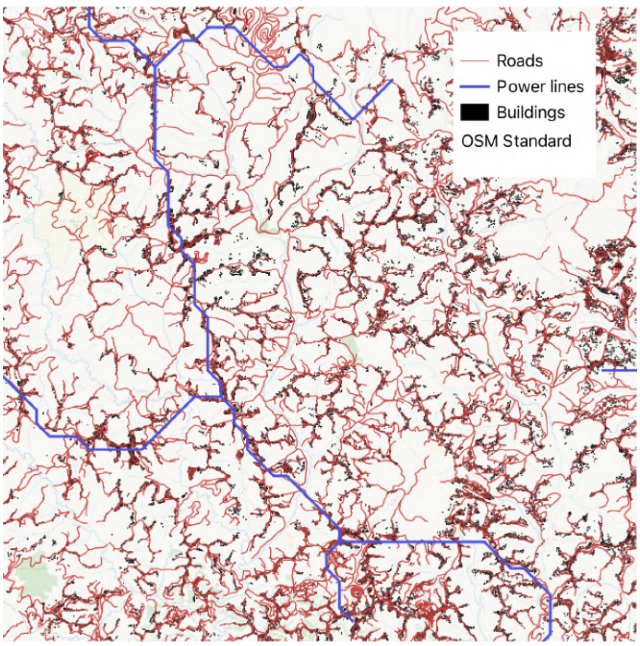
To test the model, I only needed the data inside a square of 20 km. I chose a square that best included the settlement I wanted. I defined a 20x20km square that had part of the settlement in question, but was still fully in the area of Rwanda, as the settlement is close to the border. The square has the following coordinates: [29.5, -2.775, 29.7, -2.575].

Map of Rwanda with the location of the square [29.5, -2.775, 29.7, -2.575] in QGIS

This square has 71699 buildings and 1231 settlements. 8349 of these buildings have a type, including:

- 8332 residential buildings, that I may have identified as so because they were in hamlets
- 13 commercial buildings
- 3 constructions
- 1 toilet

To test the model, we will now be using the data concerning roads, buildings, and powerlines inside this specific area. The geometries of these elements are represented in the map below.



Map of the roads, buildings and powerline in the square area

2.3. The Reference Network Model

The Reference Network Model (RNM) is a very large-scale planning tool, which plans the electrical distribution network using the GPS coordinates and power of every single customer and distributed energy resource. The model takes into account geographical factors like the roads map and topography of the area. Its purpose is not to construct the actual network, but instead to construct a reference network for which the cost gives an indication of the cost efficiency of constructing a network.

RNM is a proprietary mathematical model of the Institute de of Technological Research (IIT: Instituto de Investigación Tecnológica), a research center that is part of the ICAI School of Engineering of the Comillas Pontifical University. I am using this model for my work, without making any modification to it.

The output of the model that I focused on in the context of this project is the location of the substations. For each substation found, the RNM generates one folder, named after the substation, containing 3 files:

- subest: location of substations
- customers: information about the connected buildings
- customers_isolated: information about buildings that are not connected to the substation

This output of the RNM will then have to be converted to the right format to input in the Reference Electrification Model.

2.4. The Reference Electrification model

The Reference Electrification Model (REM) is a computer-based optimization tool that enables the automatic planning of the electrification of virtually any population size. It analyzes a variety of factors, including economic and geographical constraints, to identify the most efficient design solutions available in order to provide desired levels of electricity access in a given area.

Like mentioned previously for RNM, REM is another proprietary mathematical model of IIT of the Comillas Pontifical University. This model is an important resource to carry out this project. I have been working with REM, preparing the inputs and analyzing outputs, without modifying the code of the model in any way.

REM is a powerful support tool in the electrification planning process. It uses a combination of heuristic optimization, mathematical algorithmic optimization, and simulation algorithms to develop a least-cost grid/off-grid electrification plan that takes into account a range of factors. REM considers the specific demand profile of each customer, estimated yearly weather conditions, targets of quality of electricity supply, the reliability performance of local distribution feeders, voltage and capacity constraints of lines and transformers, and other data. It then compares a large number of clustering alternatives to find the best solution.

Moreover, REM has great precision and is capable of creating customized plans at various geographical scales. The model has been successfully applied to many cases globally, ranging from small regional areas with hundreds of customers, to comprehensive analyses of entire countries with millions. The model also has a particular configuration, LREM (Local REM), used specifically for villages or small regions to provide detailed electrification designs.

Despite REM's power to simulate and evaluate grid/off-grid electrification plans from both a technical and economic perspective, it is only one component of the electrification process. Political, administrative, regulatory, legal, social, and cultural factors should also be taken into consideration when deciding on an electrification plan.

Stages of utilization of REM

1. Data preparation

Data preparation is a critical step before using the model. This includes looking for valid sources of data, cleaning the data, and processing it to create a dataset containing exactly the inputs required, in the right format. It is important to condition the data in a way that the model can most effectively learn from. This step is essential to the good functioning of the model.

2. Mini-grid generation design

In the end, the model will determine the optimal generation and storage for each potential grouping of customers that is evaluated. The mini-grid generation design consists in facilitating this task: the approximative designs for some number of representative combinations of consumers are solved in advance and stored in a look-up table, in order to speed up the estimation of local generation cost for each future design.

3. Clustering

Customers are grouped into two types of clusters: off-grid and grid-extension. The algorithms for clustering involve joining customers into groups so that the expected cost of being connected to the grid as a group is lower than the expected cost of being electrified separately.

4. Final designs

With the obtained clusters, REM calls the Reference Network Model to evaluate network designs and costs. Then, REM decides the ideal combination of direct systems, mini-grids, and extensions for the grid. During this process, decisions are made on whether or not to accept suggestions for grouping clusters together and attaching clusters to the grid, based on design expenses.

5. Post-processing and reports

REM produces reports, graphs, tables, spatial datasets, and other relevant materials, delivering all the information on the cost-optimal designs determined by REM.

REM inputs description

An electrification problem is a set of files describing a region, electricity supply technologies, and the designer's preferences. This section describes the various input files needed for REM, which are all "text-style". The input files are organized as follows:

- ❖ Local information
  Information particular to a given location, including the buildings in the area, demand/supply profiles, existing network, local energy resources, and topography.
  - ➢ Demand (codes De#)

        - De1. Buildings (location and type of customer)
        - De2. Customers definition (in terms of demand patterns)
        - De3. Load profiles –samples- for a demand pattern (one file for each demand pattern)
        - De4. Supply patterns
        - De5. Target pre-design points definition (for the look-up table of designs)
    - Existing network (codes En#)
        - En1. Existing MV network.
        - En2. MV feeder types.
    - Topography and geography (codes Tg#)
        - Tg1. Solar power
        - Tg2. Terrain elevation (raster map)
        - Tg3. Penalized areas (polygons)
        - Tg4. Village boundaries (polygons)

- Equipment catalog

  Specifies the technical and economical parameters of the available electrification technologies.
    - Network design (codes Nd#)
        - Nd1. Lines
        - Nd2. Transformers
    - Microgrid generation design (codes Mg#)
        - Mg1. AC/DC converters
        - Mg2. Charge controllers
        - Mg3. Diesel generators
        - Mg4. Batteries
        - Mg5. PV panels

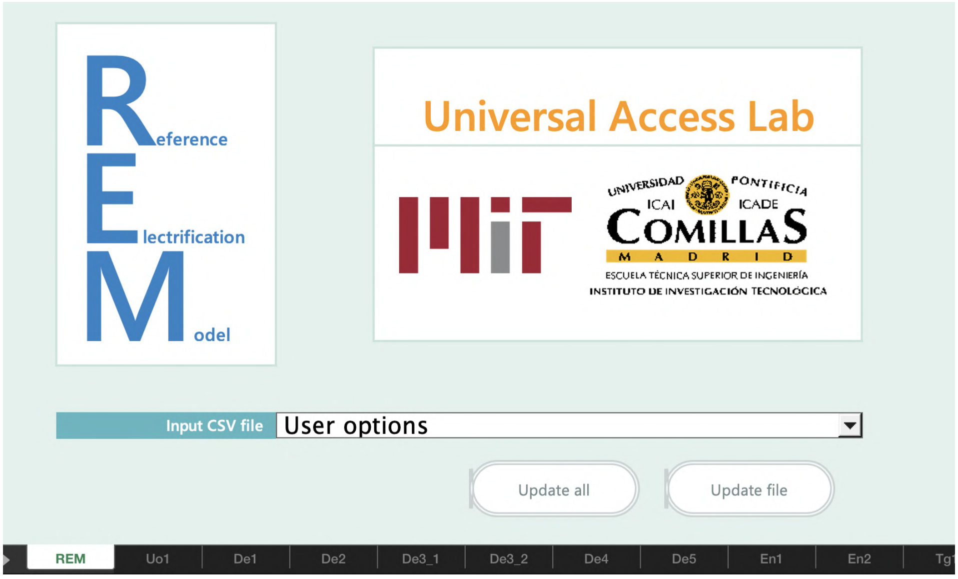
- User options

  Single file that specifies values for design parameters and references to the rest of the input files, including workspace definition, electrification problem, electrification criteria, etc.
    - User options
        - Uo1. User options

As I have been working on the input files De1 and En1, specifying the location of the buildings in need of electrification and of the substations, these two input files will be detailed in the next section. The other input files will not be more detailed as it is not essential to the comprehension of the project.

Inserting the inputs

To be able to test the model on different datasets, I used an excel file that was already created with the model to insert inputs easily. The excel file has one sheet per csv input file required to run REM. Excel is a way to have all the inputs in the same place. When I needed to change input data, I modified the corresponding sheet, and updated with the new data in the REM sheet, as shown below.



Screenshot of the REM input Excel file

2.5. Preparing the REM input data

Buildings and MV network from RNM to REM

Concerning the substations of the reference network constructed by RNM, I had to process the output data of the RNM into REM input, which involved converting the data into the correct format to be read by the REM. This process included creating a loop to run a program for all substations found by RNM, code the modifications in content and format that had to be made to translate the RNM output data into REM input, create two specific files with the corresponding information for each substation, and saving these files in the substation folders created by RNM.

As stated previously, the output of RNM was one folder per substation, each containing 3 csv files, with the following fields:

subest.csv	
field	description
x	X coordinate of the substation in UTM for the area of Rwanda
y	Y coordinate of the substation in UTM for the area of Rwanda
id	Identifier of the substation assigned by RNM
voltage	Voltage level of the substation

customers.csv	
field	description
x	X coordinate of the building connected to the substation in UTM for the area of Rwanda
y	Y coordinate of the building connected to the substation in UTM for the area of Rwanda
z	Altitude of the building
id	Identifier of the building assigned by RNM
voltage	Voltage level required

customers_isolated.csv	
field	description
x	X coordinate of the building connected to the substation in UTM for the area of Rwanda
y	Y coordinate of the building connected to the substation in UTM for the area of Rwanda
z	Altitude of the building

id	Identifier of the building assigned by RNM
voltage	Voltage level required

These files have to be transformed into input files for REM. For each substation, I created two new csv files, presented in the previous section:

- En1: location of substation
- De1: location of the buildings (connected and not connected)

The specific fields required for these 2 files are the following:

En1.csv		
field	format	description
id	integer	Identifier of the substation (assigned by RNM)
x1	double	Latitude of the substation 1
x2	double	Latitude of the substation 2
y1	double	Longitude of the substation 1
y2	double	Longitude of the substation 2
segment_type	integer	Type of the segment. 1 for all.

De1.csv		
field	format	description
id	integer	Identifier of the building (assigned by RNM)
x	double	X coordinate of the building in Longitude degrees
y	double	Y coordinate of the building in Latitude degrees
customer_type	integer	Associates the building with a type of customer based on the consumption profile (voltage level required)
elec_status	boolean	Electrification status: indicates whether a building is currently electrified or not. 0 for all as they are not

		yet electrified

I used Python in a Jupyter Notebook to achieve the transformation to the data from the RNM output files and the creation of the new En1 and De1 csv files, for each substation. I wrote all of the operations inside a loop “for subset in [*list of substations*]”, to repeat the same steps for all substations.

The En1 files, containing the information about the substation, are created from the information given in the RNM output file “subest”. The first step is to load the subest file and create a DataFrame with the data it contains, using the Pandas library.

```
#get substation file
df_substation = pd.read_csv('/Users/penelopeschuwer/Documents/TFM jupyter/RNM output/' + subset + '/subest.txt', sep=";", header=None)
df_substation.columns = ["x", "y", "id", "voltage_level", 'col5', 'col6', 'col7']
```

Code for uploading the data from the RNM output “subest.txt” file

Then, from the DataFrame df_substation, I created a new DataFrame (df_En1) that I modified to have the correct format for the En1 file of the REM input described previously, preceding with the following steps:

- The only columns from df_substation that I needed to keep were the identifier of the substation (id), its longitude (x), and latitude (y).

- The longitude and latitude from the RNM output are in UTM for the area of Rwanda. Python has a utm library including the to_latlon function, that returns longitude and latitude from UTM coordinates and the UTM zone. Rwanda is located in the UTM zone 35M. Using this function, I was able to convert all the coordinates to longitudes and latitudes.

- The REM takes a segment as the input for the location of the substations, and not a point. Therefore, I had to create small lines artificially, right around the point that was given by the RNM output.

- I set the segment type to 1.

- I set the identifier to the name of the substation (i.e. the name of the folder created by RNM).

Finally, I saved the new En1.csv file inside the corresponding substation folder.

```
#to En1 format

df_En1 = df_substation[["id", "x", "y"]]
#conversion to latlong
x = df_En1["x"]
y = df_En1["y"]
y2 = copy.copy(y)
df_En1["x"]=utm.to_latlon(x, y, 35, 'M')[0]
df_En1["y"]=utm.to_latlon(x, y2, 35, 'M')[1]
#get line
df_En1["x1"] = df_En1["x"] - 0.1
df_En1["x2"] = df_En1["x"] + 0.1
df_En1["y1"] = df_En1["y"] - 0.1
df_En1["y2"] = df_En1["y"] + 0.1
df_En1["segment_type"] = 1
df_En1 = df_En1[["id", "x1", "y1", "x2", "y2", "segment_type"]]

#change id to name of subset
df_En1["id"] = subset

#add row with nb of elements
df_length = pd.DataFrame(data={'id': len(df_En1), 'x1': None, 'y1': None, 'x2': None, 'y2': None, 'segment_type': None}, index=[0])
df_En1=pd.concat([df_length, df_En1])

#save cvs En1
df_En1.to_csv('/Users/penelopeschuwer/Documents/TFM jupyter/RNM output/' + subset + '/En1.csv', header=False, index=False)
```

Code for the creation of the En1 file from the subset.txt file of the RNM output

In the same way, I used Pandas to read the customers and customers_isolated files and create DataFrames. As the De1 file has to include the data concerning both connected and not connected buildings, I concatenated both DataFrames.

Then, I created a new DataFrame df_De1 and followed these steps to modify it, in order to have the required format for the De1 file described previously:

- I used the utm.to_latlon function to convert the coordinates from UTM for the area of Rwanda (35M) to latitude and longitude

- I created a function volt_type(n) that returns ‘1’ when a value in column 4 is 0.42 and ‘2’ when a value in column 4 is 12.47. Then I applied that function to the column voltage_level, in order to have two categories of customers, 1 for those who require a voltage level of 0.42 V, or 2 for those who require a voltage level of 12.47 V.

- I set the electrification status to 0 for all rows.

- I added a row with the number of elements of the table, to have easy access to the number of customers of a substation.

Finally, I saved the new De1.csv file inside the corresponding substation folder.

```
#get customers file

df_customers_1 = pd.read_csv('/Users/penelopeschuwer/Documents/TFM jupyter/RNM output/' + subset + '/customers.txt', sep=";", header=None)
df_customers_2 = pd.read_csv('/Users/penelopeschuwer/Documents/TFM jupyter/RNM output/' + subset + '/customers_isolated.txt', sep=";", header=None)
df_customers=pd.concat([df_customers_1, df_customers_2])
df_customers.columns = ["x", "y", "z", "id", 'voltage_level', 'col6', 'col7']

#to De1 format

df_De1 = df_customers[["id", "x", "y", "voltage_level"]]
x = df_De1["x"]
y = df_De1["y"]
y2 = copy.copy(y)
df_De1["x"]=utm.to_latlon(x, y, 35, 'M')[0]
df_De1["y"]=utm.to_latlon(x, y2, 35, 'M')[1]
df_De1["voltage_level"] = df_De1["voltage_level"].map(lambda x: volt_type(x))
df_De1["elec_status"] = 0
#add row with nb of elements
df_length = pd.DataFrame(data={'id': len(df_De1), 'x': None, 'y': None, 'voltage_level': None, 'elec_status': None}, index=[0])
df_De1=pd.concat([df_length, df_De1])

#save cvs De1

df_De1.to_csv('/Users/penelopeschuwer/Documents/TFM jupyter/RNM output/' + subset + '/De1.csv', header=False, index=False)
```

Code for uploading the data from the RNM output “customers.txt” and “customers_isolated.txt” files and creating the De1 file from it

Buildings and MV network from open source data to REM

As seen in the “Data Collection” section, I had previously collected data for the existing installations and infrastructure in Rwanda. In addition to the information from the RNM output, giving locations of substations from a reference network, I was able to create REM input files De1.csv and En1.csv from the information on the existing buildings and power lines I had collected.

In order to create a De1.csv file from the buildings shapefile, I kept only the relevant information, the coordinates and the type of buildings. I created a GeoDataFrame from the shapefile and realized some changes, in order to create the right format file for the REM input, as follow:

- The coordinates were included in the original data.

- The building type, based on the consumer’s profile, had to be determined. I considered that if the building type is either empty or residential, the building requires a voltage level of 0.42 V, and therefore the customer_type column contains a 1. Otherwise, for other types of buildings (commercial, hospital, etc.), the building requires a voltage level of 12.47 V, and contains a 2 in the customer_type column. The function to_customer_type, which appears in the code below, converts the type of building to a demand level.

- The electrification status was 0 for all buildings as they were considered not yet electrified.

- I added a row with the number of elements of the table.

With all this information, I created a De1.csv file corresponding to the buildings files I was working on. Below is part of the code in Python to perform these operations. I tried this code on the square I had been working on previously, with the following coordinates: [29.5, -2.775, 29.7, -2.575].

```python
#to De1

df_De1 = gdf_buildings[['longitude', 'latitude', 'building_t']]
df_De1["building_t"] = df_De1["building_t"].map(lambda x: to_customer_type(x))
df_De1 = df_De1.rename(columns={"building_t": "customer_type"})
df_De1["elec_status"] = 0

#add row with nb of elements
df_length = pd.DataFrame(data={'longitude': len(df_De1), 'latitude': None, 'customer_type': None, 'elec_status': None}, index=[0])
df_De1=pd.concat([df_length, df_De1])

df_De1.to_csv('/Users/penelopeschuwer/Documents/TFM jupyter/DATA/square/REM_input/De1.csv', header=False, index=True)
```

Code for the creation of a De1 file from data of a buildings shapefile

Similarly, I used the power lines shapefile to create the En1.csv REM input file. After uploading the shapefile into a GeoDataFrame, I adapted the GeoDataFrames to have exactly the input needed for REM, in the En1.csv files, with the following operations:

- To obtain the minimum and maximum latitudes and longitudes of the segment of power line, I applied the bounds function to the linestring. The bounds function returns a (minx, miny, maxx, maxy) tuple that bounds the object.

- I set the segment type to 1.

- As previously, I added a row with the number of elements of the table.

Then, I saved the En1.csv file for each district.

```python
#to En1

df_En1=gdf_lines
df_En1["x1"] = df_En1["geometry"].map(lambda x: x.bounds[0])
df_En1["y1"] = df_En1["geometry"].map(lambda x: x.bounds[1])
df_En1["x2"] = df_En1["geometry"].map(lambda x: x.bounds[2])
df_En1["y2"] = df_En1["geometry"].map(lambda x: x.bounds[3])
df_En1['segment_type']=1
df_En1=df_En1.drop(columns=['geometry'])

#add row with nb of elements
df_length = pd.DataFrame(data={'x1': len(df_En1), 'y1': None, 'x2': None, 'y2': None, 'segment_type': None}, index=[0])
df_En1=pd.concat([df_length, df_En1])

df_En1.to_csv('/Users/penelopeschuwer/Documents/TFM jupyter/DATA/square/REM_input/En1.csv', header=False, index=True)
```

Code for the creation of a En1 file from data of a power lines shapefile

Applying this code to the buildings and power lines inside the square of interest, I got two csv files, De1 and En1, as shown below.

En1

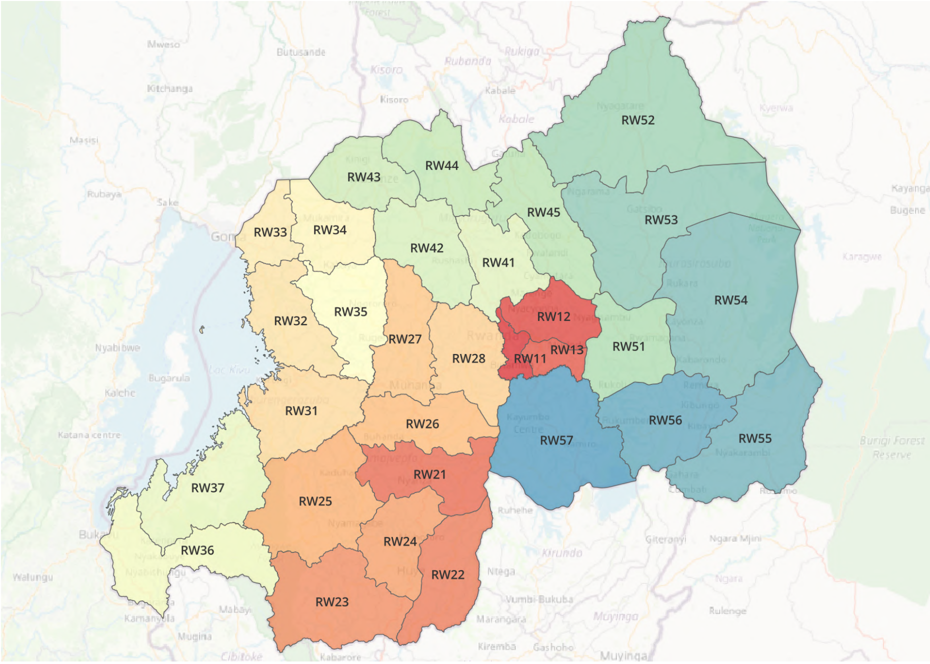
0	12.0				
0	29.52708356955577	-2.5937500207504343	29.54791690305575	-2.575	1
1	29.54791690305575	-2.610416687550426	29.62291690365572	-2.5812500206504296	1
2	29.68958357085569	-2.6895833548503845	29.7	-2.6895833548503845	1
3	29.54791690305575	-2.693750021550386	29.568750236555758	-2.5937500207504343	1
4	29.564583569855756	-2.6979166882503876	29.568750236555758	-2.6979166882503876	1
5	29.568750236555758	-2.6979166882503876	29.568750236555758	-2.693750021550386	1
6	29.5	-2.7145833550503795	29.564583569855756	-2.6916664521945908	1
7	29.568750236555758	-2.7395833552503603	29.606250236855715	-2.6979166882503876	1
8	29.606250236855715	-2.743750021950362	29.606250236855715	-2.7395833552503603	1
9	29.606250236855715	-2.743750021950362	29.610416903555716	-2.743750021950362	1
10	29.60208357015574	-2.768750022150357	29.610416903555716	-2.743750021950362	1
11	29.610416903555716	-2.775	29.672916904055683	-2.743750021950362	1

De1

0	71699.0			
0	29.62131964	-2.66495472	1	0
1	29.62518549	-2.66741167	1	0
2	29.61758177	-2.67800417	1	0
3	29.62107372	-2.66519581	1	0
4	29.61911025	-2.67372789	1	0
5	29.62999632	-2.66749075	1	0
6	29.62589304	-2.66843872	1	0
7	29.62951734	-2.66730212	1	0
8	29.63148482	-2.6640068	1	0
9	29.6397631	-2.65608493	1	0
10	29.62008258	-2.66787836	1	0
11	29.63145274	-2.66332027	1	0
12	29.61959425	-2.66863825	1	0

En1 and De1 csv files resulting from the code

After testing the process on the specific square, I needed all the data for Rwanda. As the buildings and power lines shapefiles for the entire country were heavy, I separated the data for each district of Rwanda. I used the data for administrative borders of districts from the Humanitarian Data Exchange dataset. Below is a map of Rwanda showing the thirty districts in QGIS, using the HDX data.



Map of the 30 districts of Rwanda in QGIS

I displayed the distribution of the buildings per district, as shown below, to check both the consistency of the dataset and the coherence of the operation, separating buildings per district.

```
buildings_per_adm2=gdf_buildings_adm2.groupby(['ADM2_PCODE', 'ADM2_EN'])['ADM2_PCODE'].count()
pd.set_option('display.max_rows', None)
print(buildings_per_adm2)

ADM2_PCODE  ADM2_EN
RW11        Nyarugenge     152740
RW12        Gasabo         409891
RW13        Kicukiro       222555
RW21        Nyanza         152618
RW22        Gisagara       128835
RW23        Nyaruguru      114852
RW24        Huye           169694
RW25        Nyamagabe      137677
RW26        Ruhango        151441
```

Code to show the distribution of buildings per district in Rwanda

As the results of this verification were satisfying, with a consistent order of magnitude of the number of buildings per district, I continued the process of dividing the buildings dataset.

In order to create a De1.csv file and a En1.csv file in the right format for each district of Rwanda, I followed the steps detailed previously, adapting the code to have one De1 and one En1 per district. I created a “for” loop for all of the district codes, “ADM2_PCODE”.

```python
for adm2 in gdf_buildings_adm2.ADM2_PCODE.unique():
    gdf_buildings_RW = gdf_buildings_adm2[gdf_buildings_adm2['ADM2_PCODE']==adm2]
    df_De1 = gdf_buildings_RW[['longitude', 'latitude', 'building_t']]
    df_De1["building_t"] = df_De1["building_t"].map(lambda x: to_customer_type(x))
    df_De1 = df_De1.rename(columns={"building_t": "customer_type"})
    df_De1["elec_status"] = 0

    #add row with nb of elements
    df_length = pd.DataFrame(data={'longitude': len(df_De1), 'latitude': None, 'customer_type': None, 'elec_status': None}, index=[0])
    df_De1=pd.concat([df_length, df_De1])

    df_De1.to_csv('/Users/penelopeschuwer/Documents/TFM jupyter/DATA/final/De1_per_adm2/De1_' + adm2 + '.csv', header=False, index=False)
```

Code for the creation of De1 files for each district from the data of the buildings dataset

```python
for adm2 in gdf_lines_adm2.ADM2_PCODE.unique():
    gdf_lines_RW = gdf_lines_adm2[gdf_lines_adm2['ADM2_PCODE']==adm2]
    df_En1=gdf_lines_RW[['geometry']]
    df_En1["x1"] = df_En1["geometry"].map(lambda x: x.bounds[0])
    df_En1["y1"] = df_En1["geometry"].map(lambda x: x.bounds[1])
    df_En1["x2"] = df_En1["geometry"].map(lambda x: x.bounds[2])
    df_En1["y2"] = df_En1["geometry"].map(lambda x: x.bounds[3])
    df_En1['segment_type']=1
    df_En1=df_En1.drop(columns=['geometry'])

    #add row with nb of elements
    df_length = pd.DataFrame(data={'x1': len(df_En1), 'y1': None, 'x2': None, 'y2': None, 'segment_type': None}, index=[0])
    df_En1=pd.concat([df_length, df_En1])

    df_En1.to_csv('/Users/penelopeschuwer/Documents/TFM jupyter/DATA/final/En1_per_adm2/En1_' + adm2 + '.csv', header=False, index=False)
```

Code for the creation of En1 files for each district from the data of the power lines dataset

After making these modifications, I had all the input data necessary to use the REM. In addition to the data coming from the RNM, giving locations of substations from a reference network, I had transformed the data concerning the existing buildings and power lines of Rwanda, creating one En1.csv file and one De1.csv file for each district of the country.

Electric demand

In order to use the electrification model effectively, it is crucial to input the electric demand for the designated area. As the De1 file, which has been detailed previously, gives the information about the location and type of the buildings of the area of interest, this part focuses on the De2 and De3 files, presented in the REM input description section. By incorporating information about electric demand into the model, we can accurately estimate the costs associated with providing electricity to each building. This step ensures that the electrification model takes into account realistic characteristics and energy requirements of different structures, enabling a more precise analysis of the financial aspects of electrification.

Demand patterns

The electric demand patterns are detailed in the De3_1 and De3_2 files. The De3_1 file gives the number of kW needed for an average household in Rwanda, without considering electric cooking, for each hour of the day, with two different levels of priority. I did not personally work on this electric demand pattern.

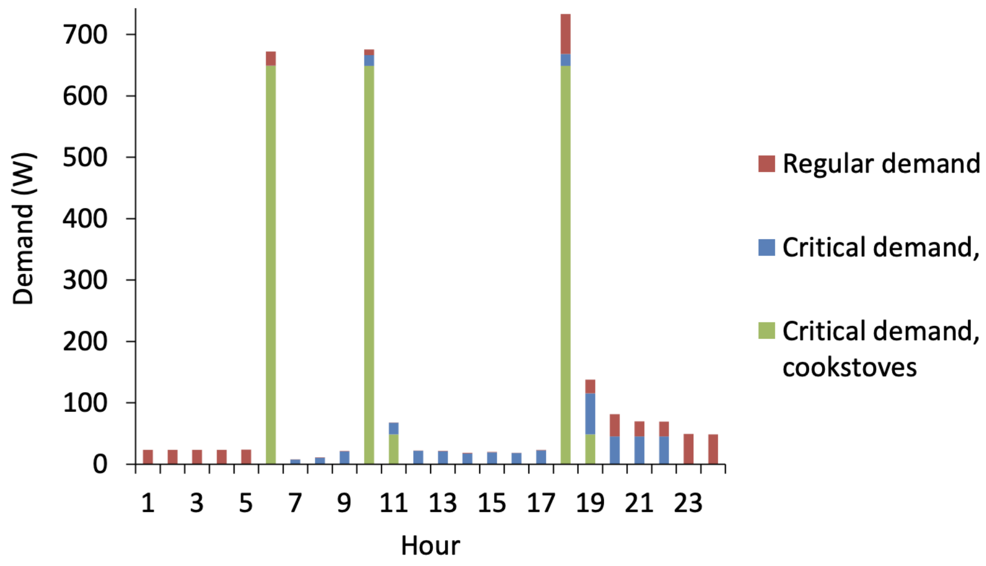
Demand patterns – [kW]

Power factor	0.80		
# of profiles	3		
# of consumers	1	100	1000

Demand samples →	High priority	Low priority
1	0.0000	0.0116
2	0.0000	0.0079
3	0.0000	0.0079
4	0.0000	0.0079
5	0.0000	0.0079
6	0.0000	0.0195
7	0.0000	0.0267
8	0.0000	0.0208

Screenshot of the De3_1 file from the input excel file

In the same way, the De3_2 file gives the information for the number of kW used by an average Rwandan household for each hour of the day, only for cooking with electricity. The electric cooking demand patterns used as inputs for the REM were obtained from the article titled "Investigating the Necessity of Demand Characterization and Stimulation for Geospatial Electrification Planning in Developing Countries", published by the MIT Center for Energy and Environmental Policy Research. This article presents their research on electric demand in developing countries, explores the benefits of coordinating clean cooking and electrification planning, and forecasts electric demand patterns. The article offers the following graph for the daily electric demand for an average household, including electric cooking, in a developing country.



Electric demand profile for residential consumers that have adopted electric cooking, from the MIT Center for Energy and Environmental Policy Research

We observe that the electric demand increases significantly three times a day, each rise corresponding to the cooking of a meal. I transcribed the information on the cooking demand, given by the green bars in the graph, into the De3_2 file used as REM input. The electric demand pattern, representing the anticipated consumption of electricity by consumers within our area of interest, is an essential input element to allow REM to accurately estimate electrification costs.

Demand profiles

Once the electric demand patterns files have been completed correctly, another step is to associate each building, entered in the De1 file, with its corresponding demand pattern, detailed in the De3 files. The De2 file defines the different types of consumer's profile and thus makes the link between the demand patterns and the consumers. As seen previously, each building is associated with a type, either household (type 1) or commercial building (type 2). In the De2 file, we associate these two types with an electric demand: we consider that a household's demand is equal to the demand profile established in the De3 files, as they were estimated for an average household, and that a commercial building's demand is ten times the demand of a household. To be able to evaluate the costs of electrification either for consumers who have adopted electric cooking, or for consumers who have not, the electric demand has been divided into separate files. This way, by changing the multiplying factors in the De3_2 column, we are able to run the model considering the electric cooking demand, or not (shown below).

Type-of-customers definition and load profiles composition

# of demand patterns	2	Coherent with sheet Uo1, cell C16
# of basic demand patterns	2	
# of types of customers	2	Coherent with sheet Uo1, cell C14

Demand patterns →	De3_1	De3_2
Household	1	0
Commercial	10	0

Screenshot of the De2 file from the input excel file, without considering the electric cooking demand

Type-of-customers definition and load profiles composition

# of demand patterns	2	Coherent with sheet Uo1, cell C16
# of basic demand patterns	2	
# of types of customers	2	Coherent with sheet Uo1, cell C14

Demand patterns →	De3_1	De3_2
Household	1	1
Commercial	10	10

Screenshot of the De2 file from the input excel file, considering the electric cooking demand

In conclusion, the process of preparing data for input into the electrification model is a crucial step in ensuring accurate and meaningful results. By carefully gathering, processing, and organizing the relevant input information, we establish a clear and solid foundation for conducting successful runs of REM. Through meticulous data preparation, we ensure to input precisely the information needed by REM to run correctly, coming from reliable sources, and in the right format. A clear set up for a run, organizing all the data in the input excel file, facilitates the next steps, including numerous runs of the model, testing various combinations of inputs. This meticulous approach to data preparation sets the stage for informed analysis of the results given by REM, that is to say the costs of electrification in the different areas of Rwanda.

3. Results

The electrification model employed in this study has provided valuable insights into the costs associated with providing electricity to each building within the designated input area. By analyzing the outputs of the model, we can gain a comprehensive understanding of the financial implications of electrification and assess the feasibility of extending electrical access to all structures in the region. This results section aims to present and analyze the outputs of REM, examining the cost distribution among different areas of Rwanda and highlighting key observations. By analyzing these outputs, we can obtain important information that will inform decision-making processes related to clean cooking planning.

3.1. Description of the outputs

This section contains a comprehensive overview of the outputs generated by the electrification model (REM). These outputs include a wide range of indicators giving different information on the costs of electrification in the different areas studied. I explored different indicators in order to understand the specific information given by each of them, and which were the most relevant ones to analyze in our particular situation.

The main outputs generated by REM are:

- Results by customer: in a csv file, each entry on the table corresponds to an individual customer with its electrification mode.
- Results per system: in a csv file, each entry on the table corresponds to an independent system (either mini-grid or grid extension).
- Results summary: in a csv file, overall results for each type of electrification mode.

The first output, the results by customer, associate each building to a system type, either a mini-grid or grid extension, as shown below. This output is crucial as it gives specific information for each of the buildings of the area of interest, for which we have a precise location.

resultsByCustomer_Test1

X Coordinate	Y Coordinate	System Type	Customer Type	Cell Number
m	m			
791475.838	9705130.1217	ext_1	1	1
791905.4202	9704857.3467	ext_1	1	1
791056.841	9703687.1086	ext_2	1	1
791448.4171	9705103.5038	ext_1	1	1
791227.9266	9704159.9084	ext_1	1	1
792440.7144	9704847.4532	ext_1	1	1

Example of a REM output Results By Customer csv file

The second output gives more specific information on each system through many different indicators. The scale is larger, as this output does not contain data on any specific building or location, but only on types of systems. However, through its various figures, this output allows us to understand the different financial aspects of electrification. Below is a table showing all the fields contained in the results by group output file.

System Type	Folder	Total Number of Customers	Number of Customers per System	Centroid (lat)(deg)	Centroid (lon)(deg)	PV Array Size (kWp)	Battery Bank Size (kWh)
Generator Size (kW)	Percentage of diesel used (%)	Annual Demand (kWh)	Fraction of Demand Served (p.u.)	Peak Demand (kW)	NPV System Cost ($)	NPV Network Cost ($)	NPV Energy Cost ($)
NPV Connection Cost ($)	NPV per Customer System Cost ($)	NPV per Customer Network Cost ($)	NPV per Customer Energy Cost ($)	NPV per Customer Connection Cost ($)	Annual Total Investment and Operation Cost ($/yr)	Annual Generation Cost ($/yr)	Annual Network Cost ($/yr)
Annual Connection Cost ($/yr)	Annual Non-served Energy Cost ($/yr)	Annual Administrative Cost ($/yr)	Cost Per kWh of Demand Served ($/kWh)	Cost Per kWh of Total Demand ($/kWh))	Annual System Cost per Customer ($/yr)	Annual Non-served Energy Cost per Customer ($/yr)	Annual Administrative Cost per Customer ($/yr)
GE Annual Total Cost ($/yr) ($/year}	OG Annual Total Cost ($/yr) ($/year}	GE Annual Investment and Operation Cost ($/yr) ($/year}	OG Annual Investment and Operation Cost ($/yr) ($/year}	CAPEX (Total) ($/yr)	OPEX (Total) ($/yr)	CAPEX (Network) ($/yr)	CAPEX (Generation) ($/yr)
CAPEX (Connection) ($/yr)	OPEX (Network) ($/yr)	OPEX (Generation) ($/yr)	OPEX (Management) ($/yr)	CAPEX Project (Total) ($)	CAPEX Project (Network) ($)	CAPEX Project (Generation) ($)	CAPEX Project (Connection) ($)
Overnight Cost (Total) ($)	Overnight Cost (Network) ($)	Overnight Cost (Generation) ($)	Overnight Cost (Connection) ($)	Region Name			

Fields of the Results By Group REM outputs

The third output provides a summary of the final electrification solution, in total and by type of electrification mode: mini-grids, isolated systems, and grid extensions. The table gives information about the distribution of the different modes over the area of study, through the number and fraction of customers for each mode, and then different global cost indicators for the entire area of study, grouped by electrification mode. Below is an example of a Results Summary table.

	Mini-grids	**Isolated**	**Grid Extensions**	**All**
Number of Customers	**8,903**	**17,148**	**26,658**	**52,709**
Fraction of Customers	**0.17**	**0.33**	**0.51**	**1.00**
CAPEX Per Customer ($/yr)	**191.69**	**173.38**	**94.23**	**136.44**
OPEX Per Customer ($/yr)	**49.83**	**49.50**	**141.73**	**96.20**
Non-served Energy Cost Per Customer ($/yr)	**5.28**	**53.99**	**66.90**	**52.29**
Final Cost Per Customer ($/yr)	**246.80**	**276.87**	**302.87**	**284.94**
Total CAPEX ($/yr)	**1,706,576**	**2,973,177**	**2,512,065**	**7,191,818**
Total OPEX ($/yr)	**443,677**	**848,847**	**3,778,261**	**5,070,785**
Total Non-served Energy Cost ($/yr)	**47,041**	**925,818**	**1,783,549**	**2,756,408**
Final Cost ($/yr)	**2,197,293**	**4,747,841**	**8,073,876**	**15,019,011**
Fraction of Demand Served (p.u.)	**0.986**	**0.903**	**0.900**	**0.911**
Cost Per kWh of Demand Served ($/kWh)	**0.312**	**0.313**	**0.206**	**0.245**

Example of a REM output Results Summary csv file

The three outputs generated by the electrification model (REM) work together, complementing each other to provide a comprehensive understanding of the costs of electrification across the entire study area. Each output offers unique insights and information that, when combined, form a holistic view of the electrification landscape. Integrating data from these outputs will allow us to gain comprehensive knowledge about the costs associated with the different electrification initiatives, along with specific details on areas of focus.

3.2. Analyzing results

This section presents the results derived from the outputs of the electrification model, which provides a comprehensive analysis of the potential electrification strategies and impacts, according to the area of study. The REM generates a wealth of valuable insights into the financial aspects associated with electrification, through the various indicators it gives, particularly in the Results per Groups outputs seen previously. The results of REM applied to Rwanda offer a good understanding of the impact of cooking on the electrification process, as well as a global vision of the distribution of the costs of electrification on the entire area of Rwanda.

To understand the impact of cooking on the costs of electrification, we run the model on a chosen region, first without considering the electric cooking demand, and then considering it, using the electric cooking demand patterns explained in the previous section.

I ran the model with the following inputs:

- The area of study was set on the specific square that has already been defined and studied in previous parts of the thesis. The buildings (De1) and existing power lines (En1) inside this specific area were inputted in the input excel file.
- Once without considering the electric cooking demand, then once considering it. To choose whether to consider the electric cooking demand or not, I modified the De2 file, as shown in the Demand Profiles section.
- The other inputs stayed the same as the defaults set ups.

Processing outputs

Once the model has run in Matlab, I obtained the expected output files, and focused on the Results By Customer and the Results By Group. As the Results By Customer assigned a system to each customer, and the Results By Group assigned an important number of indicators to each system, I was able to join both tables and assign all the cost indicators to each customer. Using Python, I created one DataFrame with the data from the Results By Customers, and another with the data from the Results By Group, and joined them on their Group attribute, as shown below.

```
df_customers_groups_wo = df_customers_wo.set_index('Group').join(df_groups_wo.set_index('Group'))
✓ 0.0s                                                                                     Python
```

```
df_customers_groups_wo.head()
✓ 0.0s                                                                                     Python
```

Group	X Coordinate	Y Coordinate	Customer Type	Cell Number	lat	long	System Type	Total Number of Customers	Number of Customers per System	Centroid (lat) (deg)	...	OPEX (Management) ($/yr)
isolated_type_1	778013.0874	9715075.0466	1.0	1.0	-2.575321	29.500147	NaN	NaN	NaN	NaN	...	NaN
isolated_type_1	778020.1171	9715073.1896	1.0	1.0	-2.575338	29.500210	NaN	NaN	NaN	NaN	...	NaN
isolated_type_1	778071.6243	9715099.4083	1.0	1.0	-2.575100	29.500672	NaN	NaN	NaN	NaN	...	NaN
isolated_type_1	778077.5723	9715094.076	1.0	1.0	-2.575148	29.500726	NaN	NaN	NaN	NaN	...	NaN
isolated_type_1	784013.467	9693076.8521	1.0	1.0	-2.774035	29.554483	NaN	NaN	NaN	NaN	...	NaN

5 rows × 58 columns

Code joining the data from Results By Customers and Results By Group

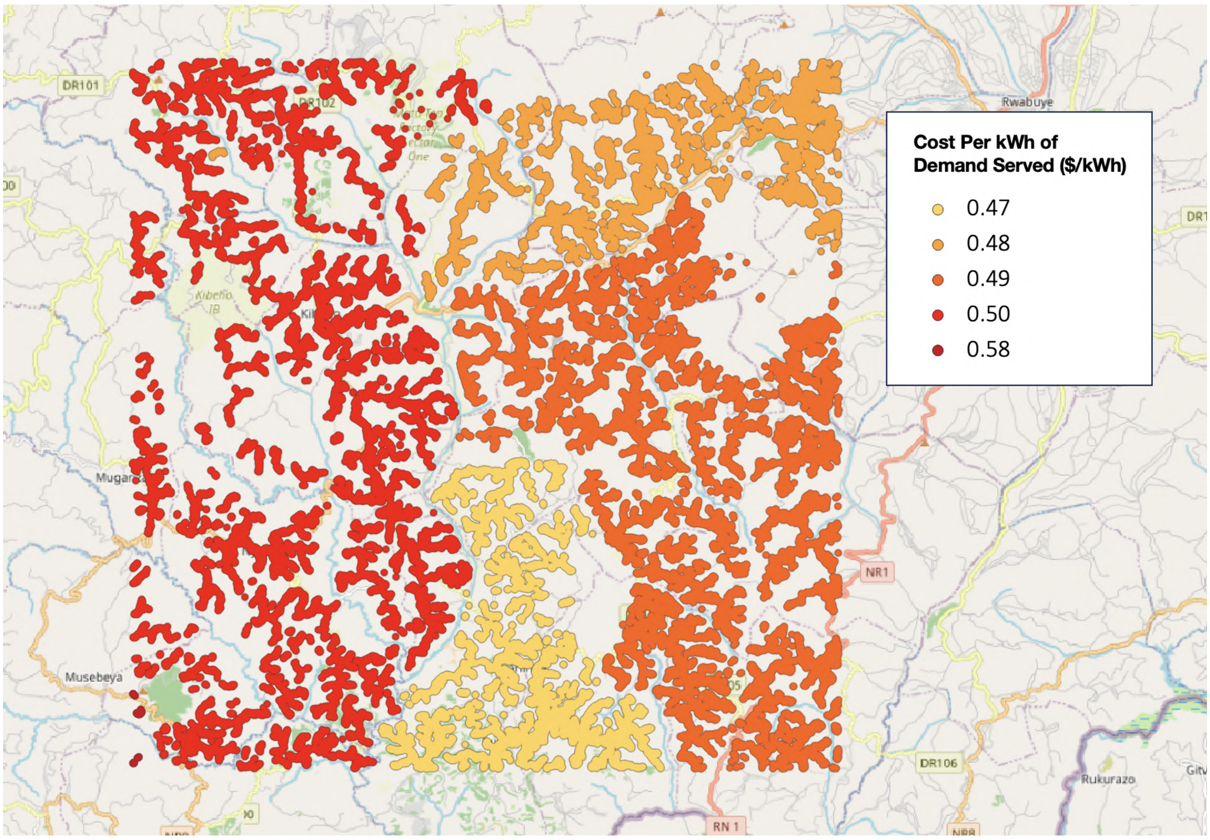
I then transformed these DataFrames into GeoDataFrames and exported them as shapefiles, in order to visualize them in QGIS.

Visualizations and analysis

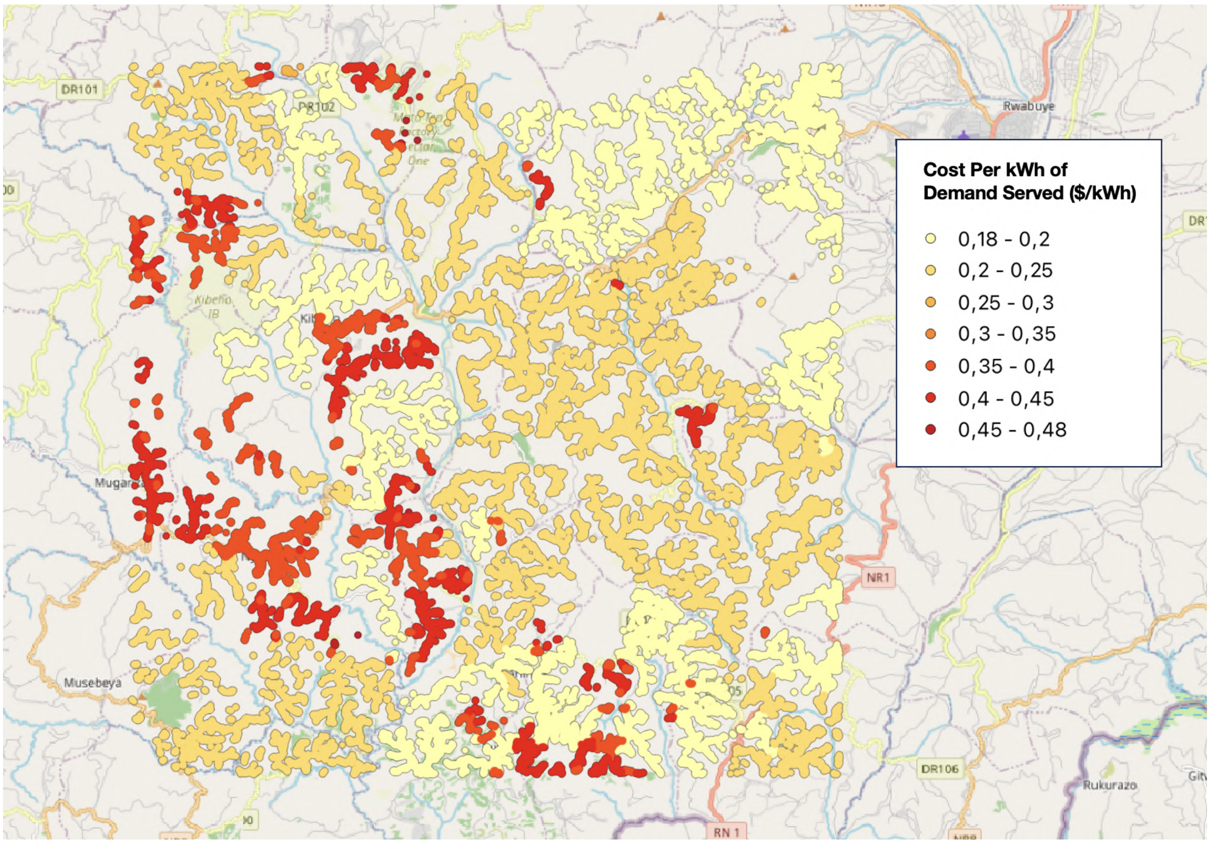
First, I visualized the cost of electricity per unit of energy served, an intuitive, easily understood, and revealing metric. Below is a visualization from QGIS of the results on a map showing the entire country, to highlight the position of our square of study. It shows without detail the cost of delivering electricity per unit of demand served, without considering the electric cooking demand. Maps showing more detail on the area of the square are presented below. This map underlines the location of our study in a mostly rural area.

Map of Rwanda showing the position of our square of interest (QGIS)

Then, the two following maps are closer visualizations of the results I also made in QGIS. They show the square of interest, gradually using colors to indicate the costs per kWh of demand served, without and with the electric cooking demand .



Map of Cost per kWh of Demand Served per Building without considering the electric cooking demand (QGIS)



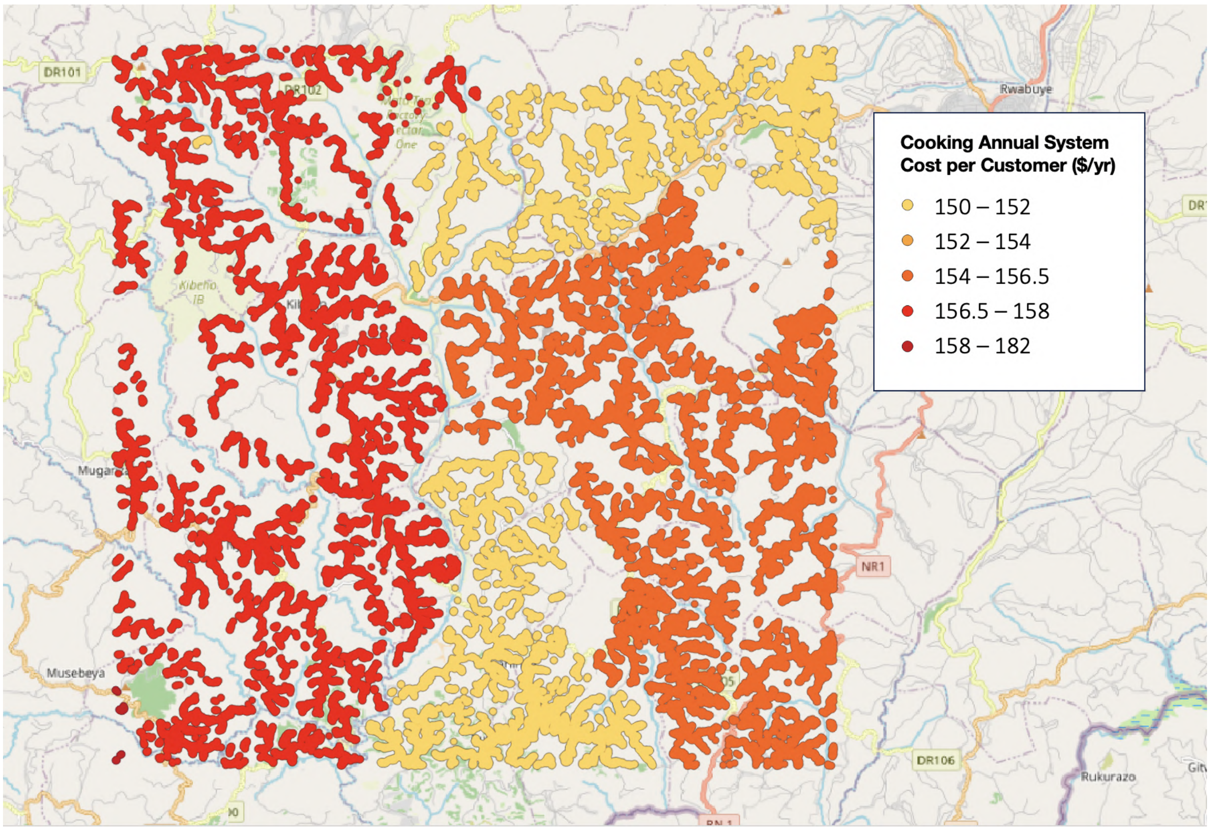
Map of Cost per kWh of Demand Served per Building considering the electric cooking demand (QGIS)

These first visualizations of REM results reveal that the adoption of electric cooking technologies significantly reduces costs per unit of energy consumed. This finding underscores the potential economic benefits of transitioning from traditional cooking methods to electric alternatives, and of coordinating this electric cooking alternative with the electrification planning. This observation puts forward a more cost-effective approach to meet clean cooking energy needs.

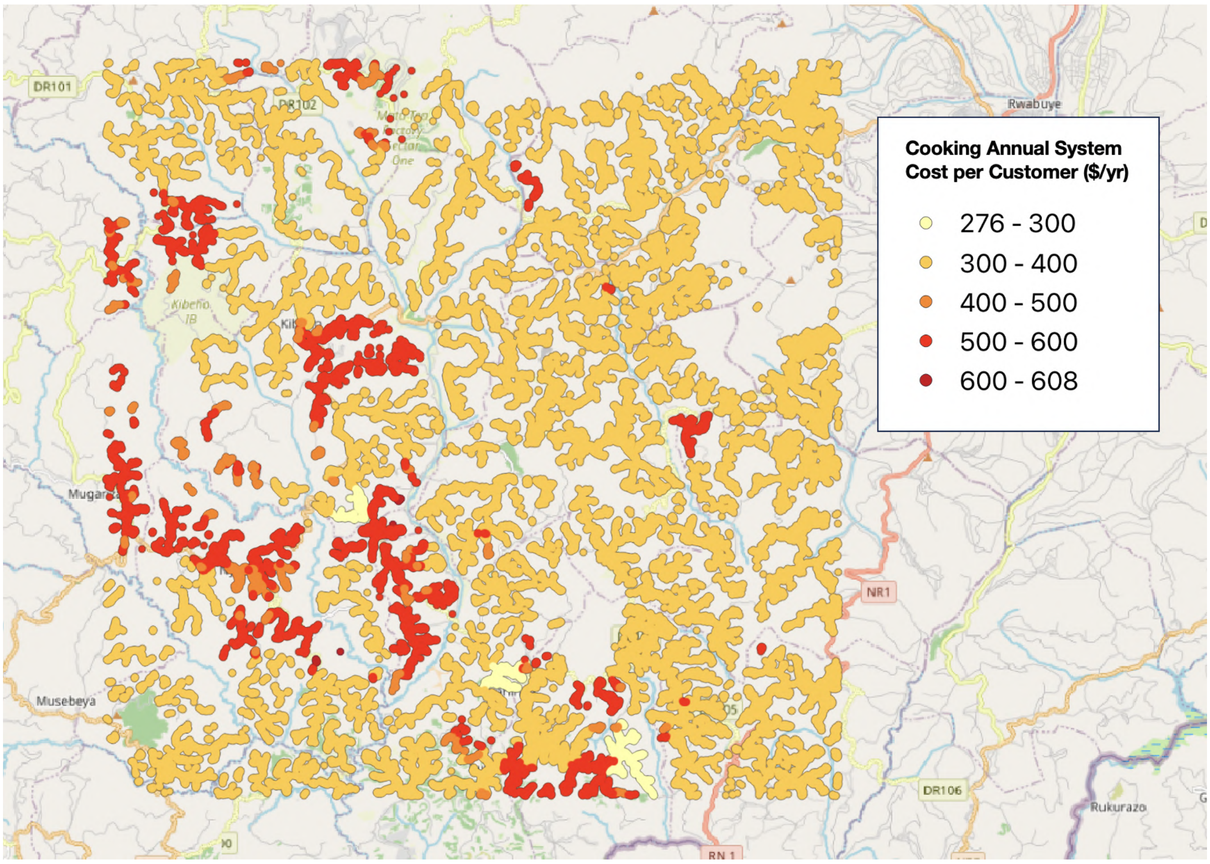
Moreover, the visualization showcases a distinct disparity in energy distribution patterns when comparing scenarios with and without electric cooking. This disparity implies that the introduction of electric cooking introduces a notable shift in the allocation and utilization of energy resources, emphasizing the importance of understanding the distributional impacts of electrification initiatives.

These two key observations provide valuable insights into the economic advantages of electric cooking and the consequential changes it involves in energy distribution over the country.

However, as the cost per kWh does not show the actual quantity of electricity served, but only a cost per unit of energy, this metric does not represent the variations in the electric demand. As we want to understand the impact of electric cooking on electrification planning and costs, I looked at another metric: the Annual System Cost per Customer, which varies with the variation of the electric demand. I visualized this metric in QGIS, as previously, and obtained the following two maps.



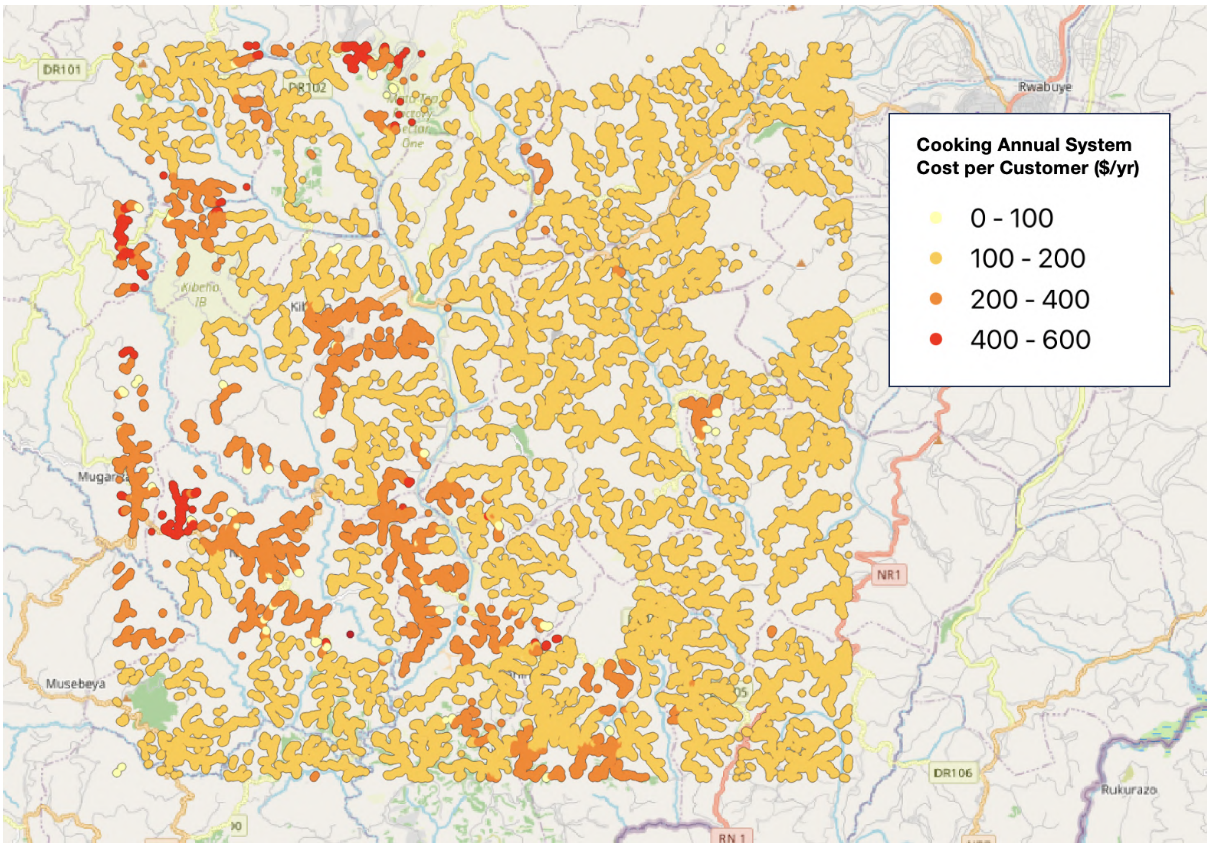
Map of Cooking Annual System Cost per Customer without considering the electric cooking demand (QGIS)



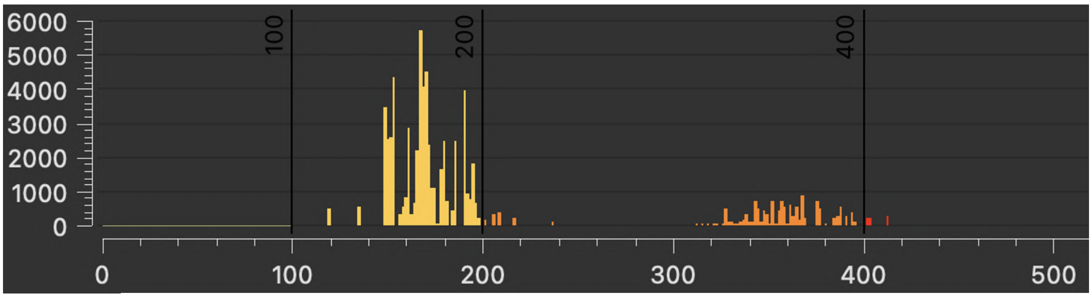
Map of Cooking Annual System Cost per Customer considering the electric cooking demand (QGIS)

The two maps depicting the Annual System Cost per Customer show similar distribution patterns to the maps illustrating the cost per kWh. However, in these two maps, the additional electric cooking demand appears more clearly. We see in the legend of the map that the inclusion of electric cooking demand reveals a notable increase in the annual cost per customer, with the costs at least doubling compared to scenarios without electric cooking. This observation highlights the significant impact that electric cooking can have on overall system costs, emphasizing the need to carefully consider and assess the implications of integrating electric cooking technologies into electrification strategies.

By subtracting the annual cost per customer considering electric cooking to the annual cost per customer without considering electric cooking, I obtained for each location the annual cost per consumer of electric cooking, represented in the map below. The histogram represents the data and the same metric as the map.



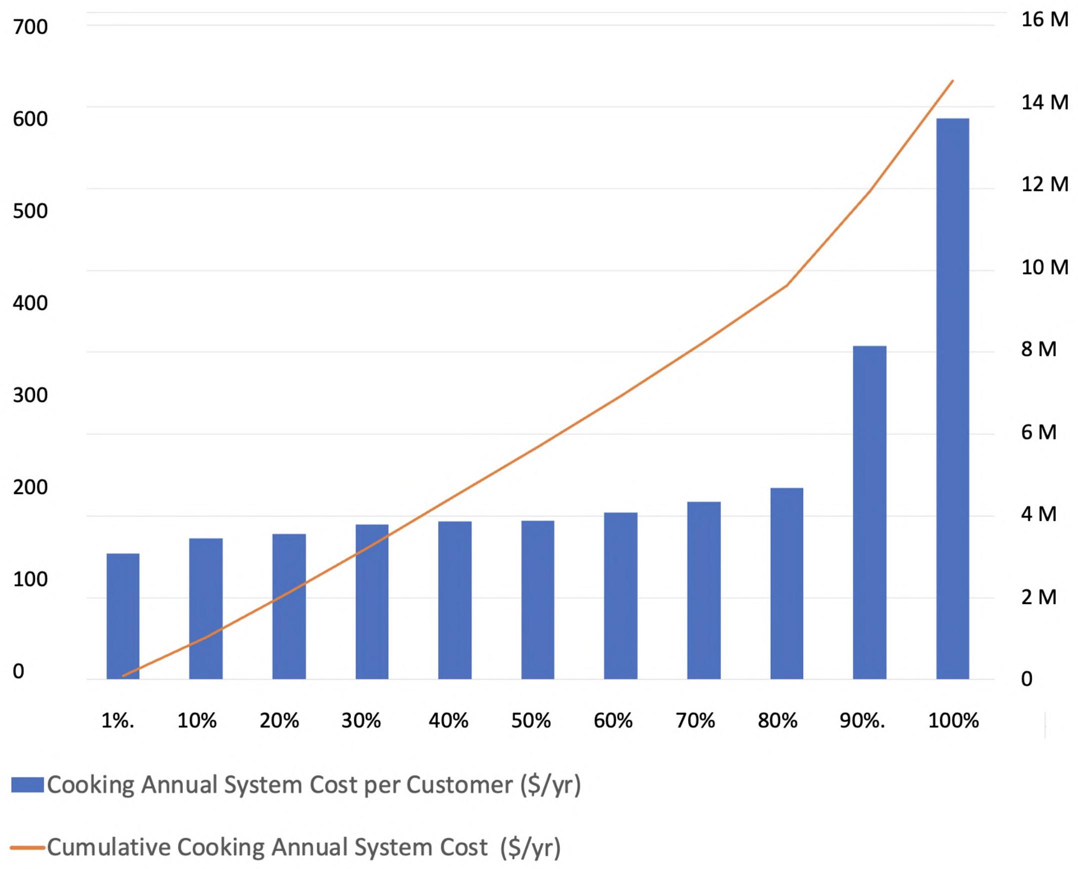
Map of Annual System Cost per Customer of the electric cooking demand (QGIS)



Histogram of the Annual System Cost per Customer of the electric cooking demand for each building

The map and histogram depicting the Annual System Cost per Customer of electric cooking provide visual insights into the costs by households of cooking daily with electric cooking technologies. The analysis reveals that for the majority of households, the annual cost per consumer of electric cooking falls within the range of $150 to $200 per year. This range aligns with the order of magnitude observed for the annual cost per consumer of electric demand without factoring in cooking. Once again, this observation highlights the significance of electric cooking as a contributing factor to the overall energy costs of the average Rwandan household.

I also represented the cumulative cost of providing electricity for clean cooking in Rwanda, from 0% to 100% of the population. The following graph shows the annual system cost per consumer per customer ($/year) in ascending order using bars, while a corresponding curve represents the cumulative annual system cost for a certain percentage of the Rwandan population ($/year). The X-axis is dedicated to the percentage of accumulated consumers, indicating the progress of consumers included. On the other hand, the two Y-axes are employed to represent the cost per consumer, each with its own distinct scale: the left axis for the cost per consumer, and the right axis for the cumulative cost.



Graph of the cumulative Annual System Cost of providing electric cooking for each percentage of the Rwandan population

The diagram combines two revealing metrics that go with one another. The bars provide a clear visual representation of the individual annual system costs, ordered from lowest to highest. Simultaneously, the curve showcases the cumulative cost of providing electricity for cooking to each percentage of the population, demonstrating how the costs accumulate as more consumers are included.

Examining the access cost concerning electrification for clean cooking solutions, the graph highlights that up to 80% of the population, the access cost appears relatively stable. This indicates that electrification could be a viable and cost-effective means to provide clean cooking options for a substantial portion of households. However, beyond this threshold, the access cost experiences a substantial escalation, indicating that it will be a challenge to implement electric cooking for all Rwanda.

As the costs of delivering electricity for cooking remain relatively stable until they cover 80% of the population's cooking needs, we might consider addressing 80% of the cooking demands with electricity as a viable clean cooking plan for this area, and even for Rwanda. The graph indicates that fulfilling this requirement for the population under study adds up to an annual cost of 9.6 million dollars. As established in the Data Processing section, the area of interest encompasses approximately 71,000 buildings, with an assumed average household size of 4 people, according to census data. Therefore, the electricity cost of 9.6 million dollars covers the cooking needs of around 280,000 individuals in the study area. Considering Rwanda's total population of 13.5 million, extending this plan to cover 80% of the country's population, approximately 10.8 million people, would entail an estimated cost of around 463 million dollars per year. This projection lays the groundwork for comprehensive planning to meet the electric cooking requirements of a significant portion of Rwanda's population in a cost-effective and sustainable manner.

However, the considerable rise of the costs after a certain threshold drives us to consider alternative fuels and technologies, alongside electricity, to complement the clean cooking planning efforts. Exploring other energy sources and efficient cooking solutions becomes imperative to ensure the accessibility and affordability of clean cooking for all segments of the population. By diversifying the approach and adopting a multi-faceted strategy, we can effectively address the challenges posed by the escalating access costs of using just one clean cooking solution, and strive towards a more comprehensive clean cooking transformation.

4. Conclusion

In the process of achieving its objectives concerning determining the cost of providing electricity in Rwanda and analyzing the specific cost for cooking purposes, this master thesis explores many challenges related to data collection, processing, and visualization. I was led to dealing with different issues with the geospatial data, including merging data from various sources, addressing format inconsistencies, resolving overlapping geometries in the buildings dataset, filling missing data regarding types of buildings, and formatting the data to meet the model's requirements. All these operations involved processing a significant volume of data, and developing strategies to reduce the complexity of the algorithms applied to this data.

Moreover, the visualizations created for this master thesis played a crucial role in understanding the impacts of developing electric cooking in Rwanda. Maps were used to visualize the geographical distribution of the cost of delivering electricity throughout the country. They highlight the financial benefits of coordinating electric cooking with the Rwandan electrification planning, as adopting electric cooking reduces by more than 50% costs per unit of energy served in some areas. Complementary graphs were employed to provide a clear representation of the evolution of the costs of providing electricity for cooking for any percentage of the population. This analysis led to an estimation of the costs of expanding access to electricity for cooking to 80% of the Rwandan population to 463 million dollars. The representations of the results provide an order of magnitude of the costs of electrification for cooking as well as a comprehensive view of the spatial disparities in electricity delivery expenses. They not only enhance the understanding of the cost dynamics, but also facilitate identifying areas that require particular attention or potential optimization. Overall, the visualizations presented in this thesis were essential in presenting data in a user-friendly and accessible manner. They are an effective support to evidence-based decision-making in the field of electricity provision and clean cooking planning in Rwanda.

In conclusion, this master thesis successfully achieved its main objectives: first, elaborating a methodology to determine cost of providing electricity, using the MIT/Comillas-IIT Universal Energy Access Lab Reference Electrification Model; and second, applying this methodology to a specific area of study, Rwanda. Indeed, these processes successfully resulted in the creation of a map giving the actual cost of delivering electricity in all the areas of Rwanda, detailing total costs and specific costs for cooking purposes. Additionally, the thesis highlights the significance of electric cooking as a contributing factor to the overall energy costs for Rwandan households and emphasized the benefits of coordinating clean cooking and electrification planning.

However, it is important to note that one objective outlined in the thesis was not fully completed. The comparison of different clean cooking solutions and determining the best option based on the area of focus was not achieved. Nevertheless, the thesis acknowledges ongoing work by the MIT/Comillas-IIT Universal Energy Access Lab on developing a model to determine the cost of providing LPG. The next step proposed is to use a similar methodology, leveraging much of the geospatial data for Rwanda used in running the REM, but applying it to the LPG model. This would enable a comparison of costs based on location and facilitate an analysis to identify the optimal clean cooking solution for each specific area. It is recognized that this task is complex due to the consideration of not only the estimated prices for electricity and LPG but also the solution that has been settled for in neighboring areas. Moreover, the thesis highlights the potential of combining multiple clean cooking solutions to meet the specific needs of each area.

In this study, while the primary objectives related to determining the cost of providing electricity in Rwanda were successfully accomplished, there remains significant potential for further research and analysis in determining the most suitable clean cooking solutions for specific regions in Rwanda. The suggested continuation of the study, involving the integration of the LPG model and a comprehensive analysis considering various factors, promises to provide insights for effective clean cooking planning.

5. References

[online] *Clean Cooking Alliance*, 28 June 2023, cleancooking.org/.

[online] *Clean Cooking Explorer*, cleancookingexplorer.org/.

[online] "What Is Clean Cooking?" *Netherland Enterprise Agency*, english.rvo.nl/information/what-clean-cooking.

[online] "Sustainable Energy for All." *SEforALL*, 7 July 2023, www.seforall.org/.

[online] "Multi-Tier Framework for Cooking: A Comprehensive Assessment Method to Measure Access to Modern Energy Cooking Services." *World Bank*, 26 Oct. 2020, www.worldbank.org/en/topic/energy/brief/fact-sheet-multi-tier-framework-for-cooking.

[online] "Moving the Needle on Clean Cooking for All." *World Bank*, www.worldbank.org/en/results/2023/01/19/moving-the-needle-on-clean-cooking-for-all#:~:text=The%20World%20Bank%20has%20been,across%2030%20access%2Ddeficit%20countries.

[online] "Technology for Clean Heat and Modern Living." *KOKO Networks*, rwanda.kokonetworks.com/.

[online] "Koko Fuel Solution - Ethanol Cooker & Ethanol Cooking Stoves in Africa." *KOKO Networks | Technology for Life in the World's Fastest-Growing Cities*, 10 Jan. 2023, kokonetworks.com/koko-fuel/.

[online] Amataya, Reja, et al. "Computer-Aided Electrification Planning in Developing Countries: The Reference Electrification Model (REM)." *MIT/Comillas-IIT Universal Energy Access Lab*, 28 Sept. 2018.

[online] Lee, Stephen J., et al. "Investigating the Necessity of Demand Characterization and Stimulation for Geospatial Electrification Planning in Developing Countries." *MIT Center for Energy and Environmental Policy Research (MIT CEEPR)*, Oct. 2019.

[online] “Rwanda Integrated Clean Cooking Plan: Report on geospatial data.” *MIT/Comillas-IIT Universal Energy Access Lab*, 2022.

[online] “Download Openstreetmap Data for Rwanda.” *Geofabrik Download Server*, download.geofabrik.de/africa/rwanda.html.

[online] “Open Buildings.” *Open Buildings*, sites.research.google/open-buildings/.

[online] “Health Facilities in Sub-Saharan Africa.” *Humanitarian Data Exchange*, data.humdata.org/dataset/health-facilities-in-sub-saharan-africa.

[online] Arderne, Christopher, et al. “Data from: Predictive Mapping of the Global Power System Using Open Data.” *Zenodo*, 16 Jan. 2020, zenodo.org/record/3628142#.YmPy8tpBwuW.

[online] “Grid3 Data Hub.” *GRID3 Data Hub*, data.grid3.org/.

[online] “Rwanda - Subnational Administrative Boundaries.” *Humanitarian Data Exchange*, data.humdata.org/dataset/health-facilities-in-sub-saharan-africa.

[online] “Natural Earth " Blog Archive " Admin 0 – Countries - Free Vector and Raster Map Data at 1:10m, 1:50m, and 1:110m Scales.” *Natural Earth* , www.naturalearthdata.com/downloads/10m-cultural-vectors/10m-admin-0-countries/.