A Geographic Information Systems-based approach for the planning and evaluation of remote DC micro-grid topologies for rural electrification

Michele Tagliapietra

Technische Universiteit Delft

A Geographic Information Systems-based approach for the planning and evaluation of remote DC micro-grid topologies for rural electrification

by

Michele Tagliapietra

to obtain the degree of Master of Science
at the Delft University of Technology,
to be defended publicly on January 21, 2019.

Student number: 4633342
Project duration: November 21, 2017 – January 21, 2019
Thesis committee: Prof. dr. ir. P. Bauer, TU Delft, Chair
Dr. E. Z. Qin, TU Delft, supervisor
Dr. Ir. P. W. Heijnen, TU Delft, TBM, external committee member
MSc. N. Narayan, TU Delft, daily supervisor

This thesis is confidential and cannot be made public until January 31, 2020.

TU Delft

Preface

This bunch of pages full of words and numbers and figures is the outcome of the long and challenging, but for sure satisfying and rewarding, work of the past year. It all started while I was randomly walking in the corridor of the third floor of the EWI low building, without a thesis project, with the intention to go and annoy people who already had one. Then I saw a poster hanging outside of Nishant's office, I read its title and I thought "Hey, that's what I want to do for my thesis"... and so it all began.

Hence, thanks *Nishant* for being a careful and empathetic supervisor, always ready to help me go through hard times with suggestions, ideas and enlightening coffee catch-ups. What you did for me and for other Master students in the past months is highly appreciated and will not be forgotten.

Thanks also to Dr. *Zian Qin* for the precious insights and pieces of advice he gave me along the project. Thanks to Dr. Ir. *Petra Heijnen* for the precious material and insights on Graph Theory and Network Analysis, which are extremely useful and fascinating fields of science that I really enjoyed exploring through her work, and for being a member of my thesis committee. Thanks also to Prof. Dr. Ir. *Pavol Bauer*, for being part of the thesis committee, for having seen the potential of my thesis project and for giving me the chance to present it and to graduate.

Now, it's time to get cheesy and emotional. I remember around 8 years ago, I was in Bologna and if my friends would ask me where I wanted to study, live and work, I would answer "Bologna, where else? It's the most beautiful city in the world!". Now, after some time, here I am, graduating in the Netherlands, after two years and a half of Master in Delft, with the prospect of living and working in Arnhem at least for some years. But one thing has not changed: I still think Bologna is the most beautiful city in the world, it's where my roots and my heart are, it's the main cause of my belly, it's where my family lives and where I hope my future own family will live, it's where most of my friends are and it's the place I will always go back to when my mind needs a break, my heart needs some warmth and my stomach needs some *ragú*.

That is why my first big "thank you" goes to all of my friends from there, to the *Superamici* and to the *Ignoranti*, to *Lollo, Branca, Max, Lupo, Paolino, Kla, Michi, J, Ceci, Frabba, Matti, Brigo, Dario, Bonna,* I miss you every day and you all contributed in your own way to make me become the kind of person I am now (so blame yourselves!). I am lucky to have you guys and I love you way more than I manage to show from 1000 km.

Thanks also to *Sara*, for being my most precious friend, at the same time the most mature and most childish person I know, for being always there to support me (or criticise me) whenever I need. Please don't forget that you can always count on me the same way I count on you. Thanks to *Benci*, for bearing me for such a long time, I still have to understand how we ended up being such good friends while being probably the most different pair of people in the world, but I'm so thankful that it happened.

During the last two years I met (or met again) a bunch of amazing people, which became a second group of true friends in the Netherlands, as important and as loved as the ones in Bologna. I would have never thought that in just two years some friendships could become so strong and so intense. I feel like I have known you for way more than two years, maybe that's why I am already starting to get sick of you, *farabutti!*

Let's start with the family... Thanks to *Ciola*, I have known you since you were a kid, taller than the average, with huge flap years and unable to play football. Our destiny brought us here together, so after 20 years I have still the honour of sticking around with the kindest person on Earth. *Really? Amazing!* (cit.) By the way, you are still the same old kid, still taller than the average, still with huge flap years and still unable to play football. *Sei proprio babbeo!*

Thanks to *Coach*, you saved me through the first year in Delft, feeding me and giving me shelter whenever I was showing up at your place without any notice. How were you so stupid to decide to go and live together after that period? Jokes aside, I admire your value of friendship, honesty, loyalty and dedication to work, and I totally wait for you to sell your engineering soul and join me on the dark side of consultancy.

Thanks to *Julen*, for being the most grumpy *sugherone* of the whole Basque country. I will always be amazed by your love and knowledge for football, and I will always consider you my worst *fantacalcio* enemy, God how I hate you! Still, inside that hard scaly skin, a sensitive and golden hearth beats. I just hope that while reading this you will not still be wearing that ridiculous moustache...

Thanks to *Vitto*, you are one of the most clever and most stupid human beings I have ever met, basically a crazy clown suited up as a financial risk consultant. You are the first person I would call for a party and for a *carbonara*, or also for a carbonara party. While you are around, *ignoranza* and *tamarraggine* are assured. So, please, stay around here for as long as you can.

Thanks to *Marti*, your short presence in the house was a bunch of fresh air after many months of just *braga*. You are the most amazingly crazy and complicated mind I have ever met, and this makes you special. I am looking forward to your graduation and to visit you in Rome. *Daje!*.

Last, but definitely not the least, thanks to *Luca*. We shared this trip together on an extremely intense level, we helped each other, we insulted each other, our stories in the Netherlands have been entangled together and I truly hope they will still be in the future. It was fun, and if we did not kill each other so far I don't think this will happen soon, but in case it happens you will be the one *lasciato a terra sanguinante*. I am extremely proud of you and of what you achieved, and I am not only speaking of your biceps and pectorals.

Leaving Dirklangenstraat, I also have to thank the only *non-latino* person in Delft that managed to get a place into my heart, *Thomas*, you should be extremely proud of that. I would have liked to see you at my defence, but I'm sure you're having a better time in Australia. Your company and help during the Master was precious to me, almost as much as your *acqua speciale*. The window episode will always be one of my favourite memories in Delft. I hope you are proud of me, in the end I made it to finish my thesis.

The last person I want to thank from Delft is *Master peluchito sensei Calcabrini*, for his never-ending patience, for being the kindest, nicest and most heartily good person that I know. I don't know how you survived the stress of having me around the office and around your house, but I thank you for that. I would have not been able to finish my thesis without you, that's for sure. I am still waiting for the coffee count on your card. You are the most fake Italian Argentinian in the world, but I am extremely happy to have met you and that you will be in the Netherlands for some more years.

You guys have indeed been like a family for me, the moments we shared, both fun and hard ones, will always remain in my hearth and will probably make me cry every time I look back at them. *Non cambiate mai.*

I would also like to quickly thank all the flatmates of Arnhem, *Jop, Cynthia, Marieke, Robin, Zoey, Suman, Eva.* You welcomed me in your house and you made me feel at home since the beginning and I am looking forward to spend some more months with you guys.

Moving to my real family, there's no need to say anything that I have not already said to my *Nonna Piera*, thanks again for being the best grandma anybody could hope for and I am happy to make you (and *Nonno* from up there) proud.

But I am also proud of somebody, who is my *fratellino*, growing to become an adult, not always in an easy way. Never underestimate yourself, and remember that through easy and hard times, I will always be there to help you, even without physically being there with you.

C'é una persona rimasta da ringraziare, senza la quale nulla di tutto ció sarebbe stato possibile. Una persona che c'é sempre stata, dal primo giorno in cui, col mio gigante testone, mi sono affacciato su questo splendido mondo. Una persona che ha svolto il difficilissimo ruolo non di uno, ma di ben due genitori, riuscendoci alla grande nonostante tutto. Una persona che mi ha sempre dato tutto, senza mai chiedere nulla in cambio. Per questo io dedico questa tesi e questa intera laurea, totalmente a lei. GRAZIE MAMMA!

Michele Tagliapietra
Delft, 21 January 2019

Abstract

Even if renewable energy generation is improving and diffusing rapidly, reliable energy access is still a major issue for a consistent part of the global population, with more than 1 billion people still lacking energy access globally. The vast majority of this share of population is living in remote rural areas of developing countries, experiencing major issues in terms of living conditions. While consistent efforts have been done in the past decades to solve the problem, still a lot of work has to be done and novel approaches need to be implemented.

In the past, most of the new energy connections were achieved through national grid extension, which is proving to be a non-adequate short-term solution for a consistent share of the remaining part of the population living in rural areas. This is the reason why decentralised solutions, such as Solar Home Systems and DC micro-grids, are becoming more appealing as alternative ways to improve energy access in developing countries.

In this framework, this Master's thesis will focus on DC solar micro-grids as a solution to the energy access problem. More specifically, the aim will be to develop a methodology to gather, process and analyse data, for planning and evaluation of remote DC micro-grid networks in rural areas of developing countries. One of the main novelty aspects of this proposed methodology is the integrated implementation of Geographic Information Systems and concepts derived from the mathematical field of Graph Theory, together with an electrical analysis.

The methodology is clearly divided into three consecutive steps. The first step focuses on gathering and processing ground-level data using GIS, to compare different micro-grid layouts in term of geometrical length. The second step consists of a graph theory-based dual-objective optimisation algorithm to design meshed micro-grids from a set of starting topologies. The third step implements a DC power flow tool to analyse the operational behaviour of the optimised layouts. The proposed methodology is explained in detail throughout the report, with an example of its application to a sample of villages in different world-wide locations.

The results of this first application of the proposed methodology allow to draw some conclusions on the methodology itself and on the comparison of different micro-grid topologies. First of all, the huge potential of the combination of GIS tools and graph theory applied to micro-grid planning is shown. The results of the layout comparison show how typically implemented micro-grid layouts are generally outperformed by micro-grids designed using novel concepts and this integrated approach. Nonetheless, each specific case studies has peculiar characteristics and conditions that need to be taken carefully into account and can lead to totally different kinds of optimal solutions. It is hence of vital importance to have a methodology which is at the same time well-structured and flexible to adapt to changes and modification of parameters in order to perfectly reflect the specific needs and characteristics of each different rural electrification project.

Contents

List of Figures

List of Tables

1

Introduction

1.1. Energy Access Problem

In this historical period in which renewable energy generation is improving and diffusing rapidly, together with a sharp decrease in costs of technologies such as photovoltaic (PV) modules and energy storage systems (ESS), it is paradoxical that reliable energy access is still a major issue for a consistent part of the global population. According to the Energy Access Outlook 2017, more than 1.1 billion people globally still lack access to energy [1], representing approximately the 14% of global population. The situation is naturally unevenly distributed between different areas of the world, with higher electrification rates in already developed countries and a higher concentration of communities not reached by energy infrastructure in developing countries, especially in areas like South-East Asia, Sub-Saharan Africa and South America [1] [2]. Furthermore, around 84% of the population lacking energy access is, not surprisingly, living in remote rural areas rather than in urban settlements.

The United Nations in 2017 included solving the energy access issue as one of the 17 Sustainable Global Development Goals (SDGs) for 2030 (see SDG 7 in Figure 1.1), officially acknowledging the foremost importance of tackling this problem with an harmonised global effort [3]. Energy is one of the most important resources for any kind of economic activity and at the basis of everyday life. Most of the appliances we use on a daily basis to perform both the most trivial and the most complicated tasks need electricity to function. The

Figure 1.1: Sustainable Development Goals identified by United Nations [3]

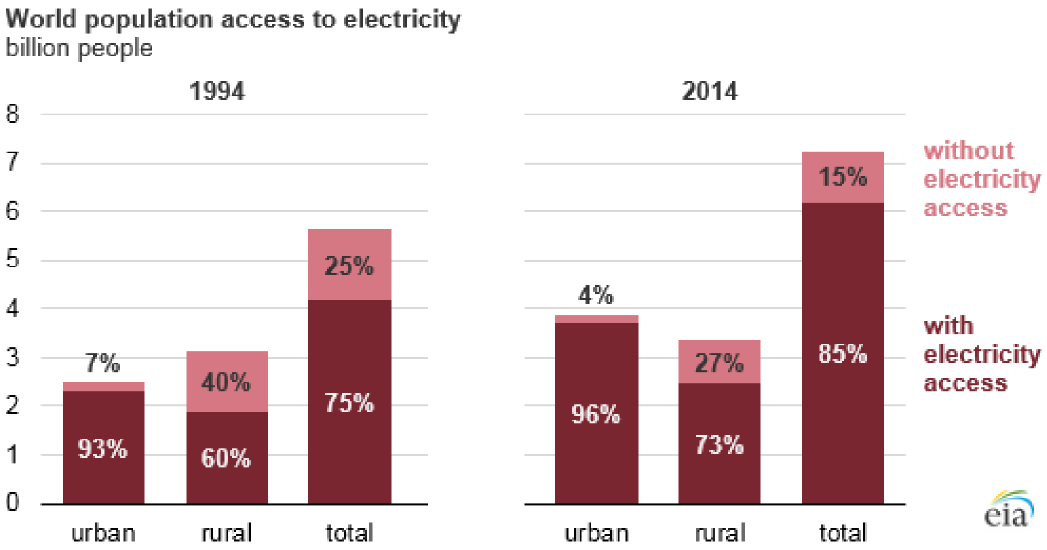


Figure 1.2: Electricity access of global population in 1994 and 2014. Source: eia [6].

benefits brought with rural electrification can then range from increased economic and productive activity, to better safety for communities, higher access to education and food, finally helping to escape poverty and to boost economic development, which is found to be directly related to electrification rate [4]. That is the reason why solving the energy issue in rural areas, especially using renewable energy to integrate and replace coal- and oil-based sources, can give a significant boost to the solutions of many other problems identified in the Sustainable Development Goals, such as SDG 1, 2, 3, 4, 8, 10 and 11, while reducing as much as possible the climatic impact of economic growth of developing countries (SDG 13). In the past, global economic growth has always been accompanied by an increased pressure on the environment by human activity. The awareness for climate threats nowadays has increased, especially in the developed world, but developing countries are still in the middle of a growth process, which cannot be stopped in name of climate impact issues, mainly caused by already developed countries. It is hence vital to achieve a decoupling between economic growth and growth of climate impact for developing countries [5]. The only way to achieve this goal is using mature sustainable technologies which are available as solutions for electrification.

1.2. Improving Energy Access

The urgency of the energy access issue is clear, but the steps to be taken in order to find a fast and effective solution are yet to be clarified. Even though the global electrification rate has been improving consistently in the past decades (49% in 1970, 75% in 1994, 86% in 2017 [1] [6] [7]), the majority of the improvements are benefiting only the urban population, while the rural electrification problem is still far from being solved (see Figure 1.2).

In the past, the main efforts and investments in the direction of rural electrification have been put on projects based on national grid extension [7] [8]. This solution, though, is facing major challenges and seems not to be technically and economically feasible for most of the remote rural areas which still need to be electrified [9] [10]. That is why, in the future, grid extension is not considered a feasible universal solution, expecting only 30% of new energy access improvements to be achieved with this solution [1], with the remaining 70% which needs to be achieved with alternative off-grid solutions. Off-grid solutions can be split into mainly two alternatives: stand-alone home systems and interconnected micro-grids. Traditional technologies which were typically applied for energy generation in micro-grids and in stand-alone systems are mainly diesel generators, because of the potentially continuous power supply, but their emissions, costs and safety risks are major drawbacks. Fortunately, the decrease in price of PV technology and energy storage systems led to a new market opportunity for rural electrification, both for stand-alone systems and micro-grids: Solar Home Systems (SHSs). A Solar Home System can simply be defined as a system composed by a generating unit, namely a PV module (usually between 10 and 350 W), an energy storage unit, usually a battery, and a Balance of System (BoS) to connect them and coordinate their operation, with the purpose of fulfilling the load

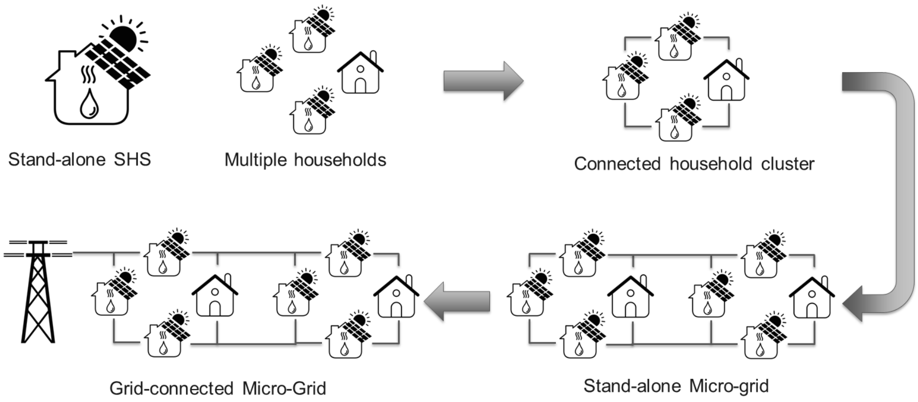


Figure 1.3: Bottom-up approach from SHS to micro-grid to connection to the main grid. Adapted from [13].

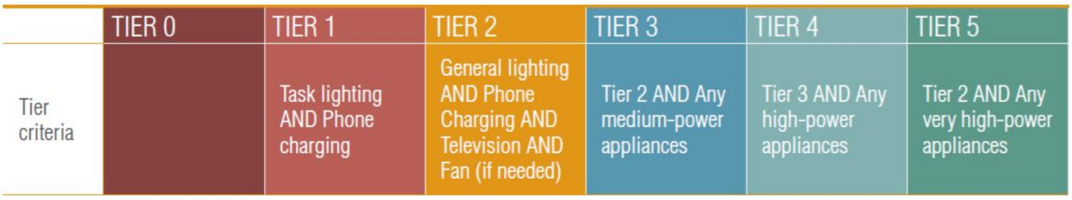
demand of a single household [11]. Since this study is not bound to be technology-neutral, other distributed generation technologies will not be considered. It is then possible to identify three distinct possible pathways to rural electrification:

1. Grid Extension
2. Single stand-alone Solar Home Systems
3. DC Solar Micro-grids on community level, interconnecting multiple SHSs

It is important to notice that these three options are not mutually exclusive, but can also be implemented on the same case study at different time steps, especially in the case of the last two options. Indeed, single Solar Home Systems of small size can satisfy initial energy demands of households, while, as the energy demand increases, it may be a cheaper option to interconnect multiple households rather than over-size the single SHSs. This reasoning would result in a bottom-up approach, as depicted in Figure 1.3 [12] [13].

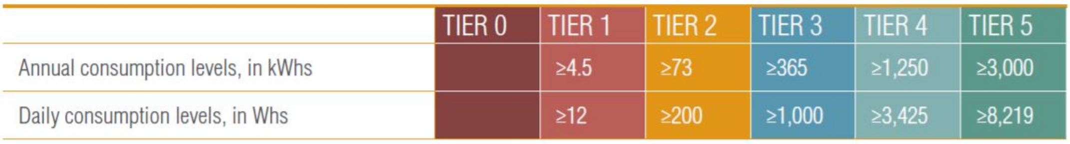
While mentioning energy demand increase and energy access, it is important to clarify how different levels of energy consumption are usually defined and distinguished. In order to harmonise global efforts towards energy access improvements and to have a clear framework, a Multi-Tier Energy Access Framework was developed in 2015 [14]. This framework distinguishes six different Tiers (0-5), using, among others, criteria of daily and yearly demand, energy availability and kind of appliances used, as seen in Figure 1.4 and Figure 1.5. Typically, for lower tiers it would be more convenient to adopt stand-alone SHSs, while, moving to the upper tiers, micro-grid connection and eventually main grid connection would become more appealing.

In this described framework, most of the studies which have been undertaken so far in rural electrification planning have focused primarily on the bigger picture, losing touch with ground-level data and with the



	TIER 0	TIER 1	TIER 2	TIER 3	TIER 4	TIER 5
Tier criteria		Task lighting AND Phone charging	General lighting AND Phone Charging AND Television AND Fan (if needed)	Tier 2 AND Any medium-power appliances	Tier 3 AND Any high-power appliances	Tier 2 AND Any very high-power appliances

Figure 1.4: Electrical appliances typical of each energy Tier, as defined by ESMAP [14].



	TIER 0	TIER 1	TIER 2	TIER 3	TIER 4	TIER 5
Annual consumption levels, in kWhs		≥4.5	≥73	≥365	≥1,250	≥3,000
Daily consumption levels, in Whs		≥12	≥200	≥1,000	≥3,425	≥8,219

Figure 1.5: Annual and daily energy consumption typical of eah energy Tier, as defined by ESMAP [14].

reality of single communities. There appears to be a lack of concrete and thorough analysis on how to physically connect in an optimal way the single households to each other, when the interconnection option results to be the more convenient.

1.3. Research Scope of Thesis Project

Having understood the foremost importance of improving energy access in rural areas, as well as its multi-faceted complexity, the scope of this thesis project was narrowed to a single specific aspect. Out of the many sub-problems related to rural electrification, it was then decided to focus on the micro-grid level, out of the three possible solutions listed in the previous section, evaluating technical feasibility and optimal layout of interconnection between multiple Solar Home Systems (SHSs). The main scope of the project was then identified as:

> *"To develop a well-structured Geographic Information Systems-based methodology to gather, process and analyse geographical data, for planning and evaluation of remote DC micro-grid networks to boost rural electrification."*

In order to reach this scope, the following research questions have been formulated and are going to be answered throughout the thesis work:

1. Which has been the approach used so far in addressing rural electrification projects and which are the micro-grid layouts suitable for Solar Home Systems interconnection in rural communities? Chapter 2 and Chapter 3

2. How can Geographic Information Systems be used to gather, analyse and process ground-level geographical data with the purpose of facilitating micro-grid planning? Chapter 4

3. How can graph theory and network analysis concepts be used to find an optimised layout for DC solar micro-grids? Chapter 5

4. What are the DC power flow operational parameters of different optimised micro-grid layouts and how do they compare to each other? Chapter 6

This thesis work, realised in the DCE&S (DC Systems, Energy Conversion & Storage) group of Delft University of Technology, is meant to contribute to a joint effort of various professors, researchers and Master students from TU Delft, in studying all the single aspects of the really complex and multi-faceted problem of rural electrification in developing countries. This falls within the framework of the Delft Global Initiative by TU Delft, which has the purpose to tackle various world-wide societal challenges through scientific research.

1.4. Contribution

In the context of scientific research towards an harmonised and well-structured effort to solve the energy access issue, this thesis work presents itself as unique in tackling some specific aspects of the rural electrification process and in using some innovative tools and a newly designed approach. The main novelty aspects of this work, which represent its contribution to academic research, are the following:

- **utilisation of modern Geographic Information Systems for gathering and analysing real ground-level data.** GIS offer the possibility to retrieve and analyse efficiently and rapidly a huge amount of high-resolution geographical data, without having to rely on large scale, aggregated macro-data.

- **A thorough geometrical analysis of alternative solutions for micro-grid layouts on a community or village level.** Many studies on rural electrification focus on the wider picture, relying on approximate assumptions for the layout and costs on the internal micro-grid level. This study is meant to make a first step into the direction of filling this gap.

- **Development of an algorithm for optimisation of meshed micro-grids layout taking into account at the same time cable length and operational parameters.** Geometrical and electrical aspects are often parted into micro-grid planning. This study aims to at least partially bridge these two separate but intertwined worlds.

The resulting methodology is supposed to be a flexible and customisable approach, both useful for large scale rural electrification planning and single case studies, using real ground-level information as fundamental starting point for the analysis, fitting in a fully bottom-up perspective. The approach is going to be structured using the lessons learned from past rural electrification projects and their flaws and gaps to be filled, identifying suitable network layouts and making use of concepts and algorithms from the mathematical and geometrical field, especially graph theory applied to network analysis. Rather than a self-standing methodology, it is meant to be integrated with other studies, focusing on aspects which have not been directly treated here, to ultimately form a complete and thorough process for rural electrification planning in developing countries.

1.5. Thesis Report Outline

After a preliminary literature review on the background information and concepts needed for the thesis project, a step-by-step methodology for DC micro-grid planning and evaluation is going to be designed and proposed. The report of this thesis project is going to be structured as follows.

In Chapter 2, a literature review is going to be performed and reported, examining more in detail the global energy context, the energy access issue and what kind of approach has been used so far to tackle it. The role of DC micro-grids and SHS as potential break-through technologies is also going to be described. Particular focus is going to be set on how Geographic Information Systems (GIS) have been used in the past and what can be their role in rural electrification planning.

In Chapter 3, a brief explanation of the most important mathematical concepts related to graph theory is going to be given. These concepts and their application to network analysis is going to be a vital part of the thesis work.

In Chapter 4, a thorough description of the application of Geographic Information Systems as a tool for micro-grid layouts comparison is going to be made. The rationale of the methodology used is going to be explained, together with the single steps of the process. The application of this methodology to selected case studies and its results are going to be shown and discussed.

In Chapter 5, the outputs of the GIS analysis are going to be used as an input for a dual-objective layout optimisation, considering not only the length-related network costs, but also the level of redundancy of the network. The rationale of the optimisation script is going to be clarified, with its assumptions and limitations. The algorithm is going to be applied to some of the case studies already used in Chapter 4 and the results are going to be analysed and discussed.

In Chapter 6, after the performed analysis based mainly on geometric and layout considerations, the electrical performance of the identified optimal layouts is going to be assessed, using a DC power flow tool developed previously at TU Delft, with some proper modifications. Aspects such as congestion, losses and voltage fluctuations are going to be examined for a restricted set of optimised layouts.

Finally, in Chapter 7, the overall methodology and results of the thesis work are going to be summarised and discussed. In addition, suggestions for future improvements to the methodology and further work needed are going to be listed.

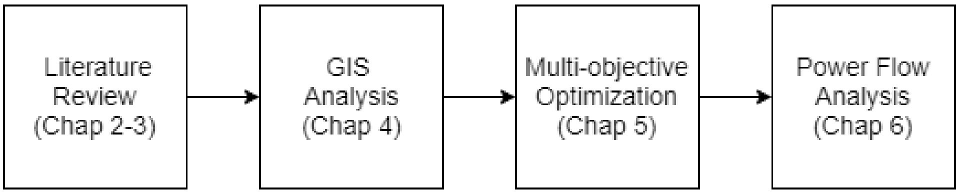


Figure 1.6: Schematic flowchart of methodology steps

2

Literature Review

In this Chapter, a literature review of the scientific research and effort in the direction of solving the problem of energy access is going to be undertaken. Possible solutions and approaches to the issues described in past papers are going to be listed and discussed, introducing the definition of energy access and the multi-tier framework. Lessons learnt from previous rural electrification projects are going to be discussed and potential gaps to fill are going to be identified.

After having described and identified the main focal issues of the rural electrification problem, the potential application of Geographic Information Systems (GIS) and its advantages and usefulness for the problem is going to be discussed, comparing its historical use in literature with the approach proposed in this thesis work.

2.1. Rural Electrification Issue

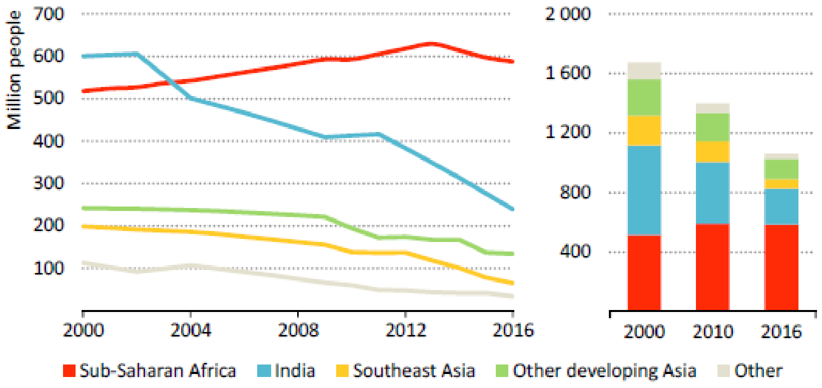


Figure 2.1: Evolution of the number of people lacking energy access in developing areas of the world [1].

Most of the activities we do every day and most of the appliances that we use require one primary input: energy. Now even more than in the past, energy is of crucial importance for economic activities, good life conditions and development. Nonetheless, a consistent percentage of global population, accounting to more than 1.1 billion people, is still lacking access to this basic resource on which our all life is relying [1]. Global energy access is unevenly distributed among different areas of the world, with higher electrification rates in already developed countries and a higher concentration of people not reached by energy infrastructure in developing countries, especially in areas like South-East Asia, Sub-Saharan Africa and South America. Furthermore, around 84 % of the population lacking energy access is, not surprisingly, living in remote rural areas rather than in urban settlements. In the last decades, awareness of the issue is increasing and improvements in energy technologies, especially renewable technologies such as photovoltaic systems and energy storage systems, are boosting the implementation of decentralised solutions for rural electrification [8]. Figure 2.1 shows the evolution of the amount of people still lacking energy access in some of the most problematic areas of the world in the past decades. Even though impressive results have already been reached in improving the global electrification rate (from 49% in 1970, to 75% in 1994, to 86% in 2017 [1] [6] [7]), the problem is far from being solved and a bigger effort is needed to implement efficient and wide-spread solutions, especially in rural areas. International aid and efforts from both the developing countries local government and foreign developed countries investors, together with the coordination and support of non-profit organisations, are needed to boost even more the solution to the energy access and rural electrification issue.

2.1.1. What does *energy access* stands for?

Before discussing potential solutions to tackle the issue, it is important to have a better understanding of what energy access actually means and how it has been treated and identified in the past by organisations and governmental bodies which have been involved in this field. A big step forward in the acknowledgement of the problem by the international community and a strong declaration of purpose in the direction of addressing and actively trying to solve it was represented, in 2015, by the inclusion of Sustainable Development Goal number 7 in the list of the 17 Sustainable Development Goals for 2030 set by the United Nations and signed by 193 countries. The statement of the SDG says:

> *"Ensure access to affordable, reliable, sustainable and modern energy for all"* [3].

On the same line of thoughts is is also possible to cite the definition of energy access given by the International Energy Agency (IEA) in their Energy Outlook 2017:

> *"[...] having reliable and affordable access to clean cooking facilities and electricity, which is enough to supply a basic bundle of energy services initially, and then an increasing level of electricity"* [1].

These statements already give a hint of how energy access definition is more complex than just a binary distinction between connected and unconnected households/communities. This dual, classical way in which energy access was defined in the past is not adequate to represent the multi-faceted nuances and grades of a complex step-wise process of moving from no energy access to a completely developed, safe and reliable furniture of electricity, which can be found for example in developing countries. Both statements mention two important aspects of energy access, which are reliability and affordability , while the statement from IEA mentions the concept of a non-binary vision of energy access and of a gradual process, sometimes called *climbing the energy ladder* [15], from zero energy consumption to appliances with higher power consumption (see Figure 2.2).

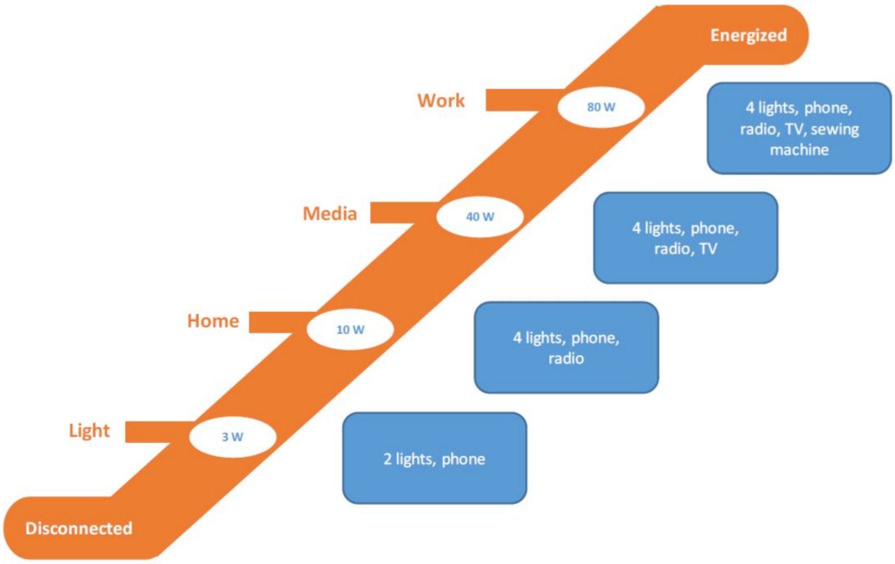
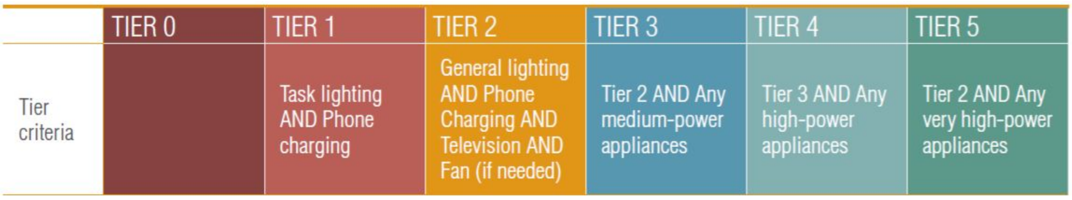
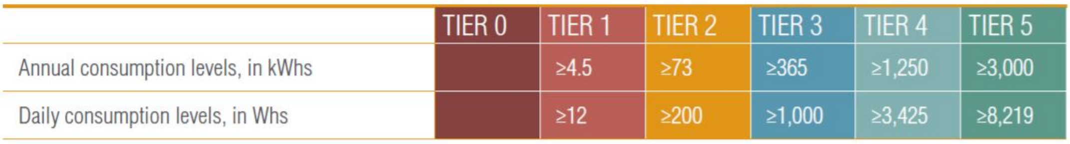


Figure 2.2: Schematic representation of the so-called *energy ladder*. Source: MWH [15].



	TIER 0	TIER 1	TIER 2	TIER 3	TIER 4	TIER 5
Tier criteria		Task lighting AND Phone charging	General lighting AND Phone Charging AND Television AND Fan (if needed)	Tier 2 AND Any medium-power appliances	Tier 3 AND Any high-power appliances	Tier 2 AND Any very high-power appliances

Figure 2.3: Electrical appliances typical of each energy Tier, as defined by ESMAP [14].



	TIER 0	TIER 1	TIER 2	TIER 3	TIER 4	TIER 5
Annual consumption levels, in kWhs		≥4.5	≥73	≥365	≥1,250	≥3,000
Daily consumption levels, in Whs		≥12	≥200	≥1,000	≥3,425	≥8,219

Figure 2.4: Annual and daily energy consumption typical of each energy Tier, as defined by ESMAP [14].

2.1.2. A Multi-tier framework

In order to overcome the inadequacy of the classical framework, a new Multi-Tier Energy Access Framework was developed in 2015 by the Energy Sector Management Assistance Program (ESMAP), with the additional purpose of harmonising the global efforts towards a solution to the energy access problem within a globally-recognised framework that takes properly into account the different levels of energy consumption [14] [16]. This framework clearly distinguishes six different Tiers (0-5), using as basic criteria the daily and yearly energy demand per household, as well as the kind of appliances which are being used, as shown in Figures 2.3 and 2.4. One of the major positive aspects of this framework is that the distinction is made not only according to these basic criteria, but also taking into consideration the continuity and availability of energy supply, the actual distribution of energy resources among different households in a community, the reliability and quality of the energy supply in terms of voltage and frequency fluctuations, the compliance to national safety and regulatory standards and also the affordability of the energy services for the community. Moreover, at the household level, distinctions can be made between energy use for different purposes, treating separately access to electricity, energy used for cooking and energy used for heating purposes, taking also into considerations which kind of sources are used for each of the mentioned energy use. This framework has received a positive response from the organisations and stakeholders involved in rural electrification projects and has been already widely accepted and applied by the community [17].

2.1.3. The importance of energy access

With this new framework, it is possible to better capture with specific and dedicated metrics all the aspects that are relevant for the evaluation of energy access of single households and whole communities, giving a more defined and accurate picture of what is their status of development. This is due to the fact that improvements in energy access can lead to increased productive and economic activity, as a direct correlation has been noticed between the electrification rate of a country and its income and economic development [4] [11]. A correlation has also been theorised between the energy consumption of a country and its Human Development Index (HDI), as shown in Figure 2.5, stating that the development of poorer countries is strictly linked to energy access [18]. The downside of this trend is that usually increased income and economic activity lead in turn to an increased environmental impact of the considered country. Given the already stressed climate situation of the planet, it is of foremost importance to find alternative ways to boost economic development in developing countries while at the same time decoupling their environmental impact [5].

apart from the correlation with economic development and income, lack of energy access can have negative consequences on multiple other aspects of everyday life of rural communities. As a first example, not having illumination infrastructure leads to higher safety risks, especially for women, while walking around the village during night time. Moreover, the possibility to have light for a longer period of time each day and night leads to potentially increased productivity, keeping industrial activities and shops open for longer periods during night time, and also for more chances for young people to receive some kind of education or to invest time in studying during the night [19]. Another factor not be forgotten is the use of electricity for cooking purposes: the most used solutions for cooking in rural communities if energy access is lacking is represented by coal, biomass and kerosene, which can have detrimental impact on people's health and safety,

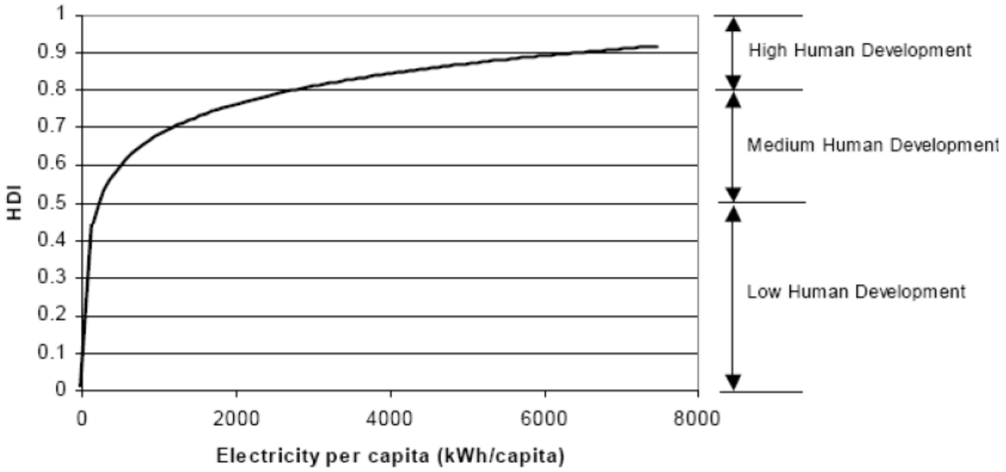


Figure 2.5: Correlation between electricity consumption per capita and Human Development Index (HDI) [18].

for their toxicity, and naturally have a higher environmental impact both on local and global level with respect to electric stoves, due to their emissions. An increased substitution of kerosene, coal and biomass with electric alternatives would lead to an important reduction in premature deaths due to use of dangerous and toxic fuels [14]. The same kind of issue is encountered also for lighting purposes, with kerosene lamps and candles being used to substitute electric lighting, potentially increasing also the risk of fires in households. Another improvement on the health side which is related to better energy access is the possibility to build and power health facilities such as small hospitals or clinics to provide sanitary and medical services to remote communities, which is extremely difficult without a reliable and constant energy access. Such activities, demanding an already relevant energy access, fall inside the category of *communal services and productive activities*, which include all the productive activities or services which are bringing benefits to the entire community and not only to the single households, such as hospitals, wells, irrigation systems, community-owned productive machinery, schools or infrastructure [20] [21]. Usually these kind of loads can be satisfied only at the higher steps of the energy ladder, after having satisfied the demand of single households, and represent a substantial improvement in the living conditions of the entire community.

Figure 2.6: Sustainable Development Goals identified by United Nations [3]

After all these considerations, it is clear how energy access is an issue which influences many different aspects of rural communities life, being of foremost importance for their development and for the development of poorer countries in general, not only on the economic point of view, but also considering safety, education, gender inequality, sustainability, access to food and employment. Referring back to the Sustainable Development Goals identified by the United Nations, it is possible to state that improving energy access would have beneficial effect not only on SDG 7, which is directly concerning energy access, but on many other identified goals, such as SDG 1, 2, 3, 4, 5, 8, 10 and 11 (see Figure 2.6). Clearly, the situation is much more complex than how it is described here and energy access would not solve at once all the problems of developing countries, but improvement on the energy side would surely lead to better living conditions and higher possibilities of development for rural communities in the short and long term.

Once energy access and its levels have been clearly defined, it is possible to move on to the actual implementation of solutions for the rural electrification issue, identifying which are the alternatives and how to evaluate which one is the best option for each specific case.

2.2. Solutions for rural electrification

The foremost importance of solving the energy access problem to boost development of poorer countries and to improve living conditions of rural communities is now clear. The following question is: how to solve the problem? What are the solutions available to grant energy access to remote communities and to disconnected communities in general?

2.2.1. The Grid vs Micro-grid vs Stand-alone trilemma

There are three possible solutions which have been discussed and can be implemented:

1. Extending the main transmission and distribution grid to reach the considered community;
2. Building a stand-alone micro-grid connecting between them multiple households of a single community/village/neighbourhood to satisfy their combined demand;
3. Installing stand-alone systems in single households, with sufficient power production (and eventually energy storage) to satisfy their demand, sometimes also referred to as *nano-grids* [20].

The second and third options can both be considered decentralised solutions, because they do not rely on existing national or regional infrastructures as in case of the first option, even though there are major differences between them, which are going to be treated in the following paragraphs. For a first quick evaluation of the potential solutions, it is also possible to group together decentralised solutions at the first stage of decision making, in order to preliminary compare them with the grid extension alternative and, only at a later stage, evaluate which option is more viable between single household systems and micro-grids.

2.2.2. Historical approach to the rural electrification issue

The trend in the past has been to heavily make use of grid extension to improve energy access, with only 3% of the total amount of newly connected households being reached with decentralised solutions, while the 97% is still relying purely on main grid extension [1] [7] [8]. This trend is efficient only for households and communities which are within reach of the already existing electricity infrastructure, while remote communities are left unconnected waiting for the main grid to reach them. The inadequacy of grid extension as a universal solution for electrification had already been stated earlier in the World Energy Outlook 2011 by IEA, which estimated that only 30% of global rural areas could be reached by simply extending the transmission and distribution grid [22]. In this context, the only way to grant energy access to extremely remote communities on the short term, without having to wait months (or years) for grid extension, is the implementation of decentralised solutions such as stand-alone micro-grids or single household systems (such as SHSs) [9] [10]. The reasons why grid extensions is not always a viable short term solution for many rural communities are manifold: excessive distance from the main grid; difficulty in reaching specific locations due to geographical barriers; poorly aggregated communities with high geographical dispersion; lack of economic resources for the government (or the community) to invest in expensive grid extension; risk of tampering or manumission of long transmission cables, especially in poor and politically unstable countries.

Positive trends have been registered in the last years, with an increased percentage of new energy connection being powered by renewable energy sources, with this trend being pushed by the consistent decrease in costs of renewable energy technologies such as photovoltaic systems and energy storage systems. In fact, between 2000 and 2012, a vast majority of the new connections to electricity were powered by fossil fuel-based generators (72%), primarily coal, while from 2012 onward the trend changed to move towards a higher share of renewable sources (35%), especially in sub-Saharan Africa [1]. Among renewable sources, solar energy has proven to be the most suitable solution for many off-grid application, given the high level of solar

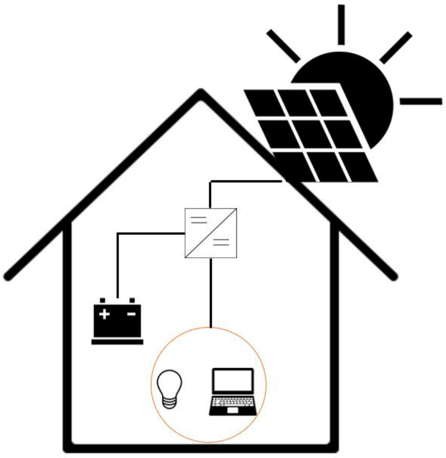


Figure 2.7: Schematic representation of a single Solar Home System [25].

irradiance typically found in most of the developing countries experiencing energy access problems [23]. PV systems are suitable both for applications in single household Solar Home Systems and in connected solar DC micro-grids, which are going to be discussed more in detail in the next section.

2.2.3. Decentralised solutions: Solar Home Systems and micro-grids

Moving the focus to the third option, namely stand-alone systems, in the scope of this thesis work the considered option is going to be narrowed down to a particular type of household system based on harnessing and storing photovoltaic energy, the already mentioned Solar Home Systems (SHSs).

The level of complexity of a Solar Home System is inherently lower compared to micro-grids. A Solar Home Systems can be broken down to a really simple set of components:

- A Photovoltaic system to generate power
- An Energy Storage System, usually electro-chemical batteries
- Loads to be powered, preferably DC appliances
- A Charge Controller (CC), usually integrated in a DC/DC converter, to control the different components and their operation
- (DC) Cables to connect all the parts between them
- An AC/DC converter in case of presence of AC loads

A schematic layout of a Solar Home Systems is depicted in Figure 2.7. SHSs can have different sizes and different layouts and can range from pico-solar applications to various kilowatts of production peak power, the most simple design being a single PV module, a single battery and a charge controller, meant to satisfy the power needs of basic DC home appliances, laying at the first steps of the electrification ladder [24].

Being isolated from any other energy source or network, a single Solar Home Systems is more sensitive to power fluctuations and power outages, not having the possibility to rely on a back up source once the energy stored in the battery is over. Hence, proper sizing (usually over-sizing) of the components has a much higher importance in SHSs than in micro-grids and requires additional care and more conservative assumptions [26]. A more and more considered solution is then the interconnection of single Solar Home Systems in a community micro-grid, in order to combine the demand and production of multiple households to face shortages and excess of production, to reduce costs needed for the over-sizing of components of single Solar Home Systems and to promote a community-based system which promote cooperation between different families or even different villages [27].

2.2.4. The comparative approach

In order to fully reach each single remote community and provide reliable energy access, the implementation of the three mentioned solution has to be evaluated carefully to find the most viable option for each singular case. A common mistake in the past was to directly plan grid extension without considering any off-grid alternative solution. Nowadays, fortunately, the technological improvements and decrease in price of the alternative options represented by micro-grids and Solar Home Systems are offering other viable solutions to take into account and to implement. Nonetheless, there is still an important bias towards one or another solution, depending on which kind of entity is taking care of rural electrification planning or what kind of financing is involved in the project: as an example, NGOs with low budget would prioritise off-grid solutions focusing on single villages, while governmental bodies with high financing opportunities would prioritise grid extension projects, making the whole decision making process biased and non-neutral [28]. This kind of single-way approach is detrimental for the research of a really optimal and most viable solution for each single community. It is then important to take into considerations multiple possible solutions at once and compare them with each other.

In the past decade the trend has changed, with many research papers being published on rural electrification making a comparison of different alternatives, sometimes just two of three mentioned ways, sometimes all three of them, trying to find the optimal mix of generation approaches to effectively and comprehensively tackle the energy access issue. Examples of really good publications in this direction are [4] [9] [28] [29] [30] [31] [32] [33] [34] [35] [36] [37] [38].

All the cited papers, though, have some downsides or some aspects which are not fully examined or taken into account, or present some needed assumptions which pose a threat to a full conformity with reality of their results. Discussing some of them into more details, as an example [9] gives a dualistic comparison of grid extension solution against an off-grid alternative, focusing on a really technology-specific case, namely hydrogen-based micro-grids, applied to one single village as case study. Despite using a well-structured cost model and a fair methodology for comparison of the two alternatives, the chosen technology unreadiness leads to a trade-off distance from the main grid which exaggeratedly favours grid extension rather than off-grid systems, making the whole comparison unbalanced. This is not the case in [31], which compares, even if on a purely qualitative way, different available generation sources for off-grid applications, trying to rank them on the basis of a set of a criteria and composite indices, including costs, emissions, acceptance by the community, easiness of implementation. Even though not making any quantitative analysis, this is a good overview of different available technologies for off-grid solutions, even if not making a comparison with grid-extension nor distinguishing stand-alone household systems from micro-grids. In [34], a clear distinction and comparison is made, on national scale, between the implementation of stand alone SHSs and grid extension as dual possible solutions for reaching unconnected communities. A really interesting aspects of this comparison is that it takes into consideration the latent energy demand of still unconnected communities, reaching the final conclusion that, for the case of Kenya, 17% of the population is more cost-effectively reached by off-grid PV system than with grid extension. The downside of this work is the lack of consideration for micro-grids as third option for electrification. A different approach is used for example in [38], with the comparison between grid extension and isolated micro-grids with two alternative sources of energy, taken singularly or combined: PV systems and diesel generators. The cost comparison is based on LCOE parameters and takes into consideration fuel price, solar irradiance, distance from grid and agglomeration of households, giving a good compendium of electrical and socio-economic data. It does not consider, though, the stand-alone SHSs option and, as in [9] and [31], it is extremely site specific and hard to generalise.

Some really comprehensive pieces of work have been carried out by a group of researchers in [28], [33] and [35]. In their work, the three mentioned alternative solutions for rural electrification issues are examined and compared at once. Along their work, many different parameters are taken into account for the cost comparison, including technology costs, available energy resources, population density, extrapolated household density and household number per agglomeration, distance from the grid, grid electricity cost and, remarkably, target level of electrification inside the multi-tier framework previously described. This last aspect is peculiar and extremely important, since the level of expected energy demand (and the desired target of electrification level) has a strong influence on the two cost metrics which are used to make the comparison in the cited papers: LCOE and Total Cost per Household. It can be seen in the results of [33] how the three alternatives compare to each other with varying electrification Tier assumed as target (see Figure 2.8). As expected, for lower electrification targets, grid extension is highly disfavoured compared to off-grid solutions, while higher electrification targets may justify the investments for grid extension. Another remarkable trend is the increasing convenience of micro-grids with increasing household density, which means more clustered and

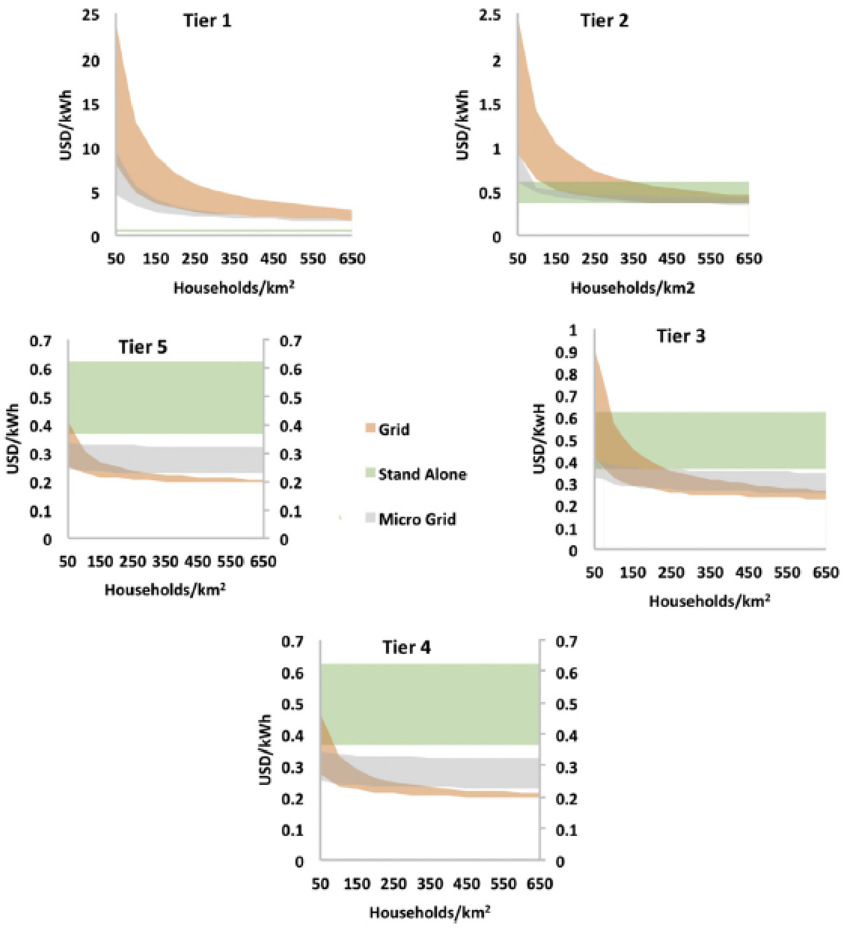


Figure 2.8: LCOEs of different energy access solutions for different Tiers of target electrification level as a function of household density [33].

condensed villages and communities and lower costs of connecting households to each other. The methodology presented in this line of papers is applied on national level for different developing countries, with similar results, suggesting a given energy access solution mix based on the mentioned macro-data and parameters. While being a really exhaustive, accurate and valuable work, there is still a gap that needs to be filled in their methodology, which is a deeper analysis of the internal layout at micro-grid levels, avoiding generalised assumptions on the costs of interconnection and possibly using ground-level detailed data at regional or even village/community-scale.

Some more detailed analysis of the actual costs of the micro-grid option is undertaken in [4], which compares the three mentioned options for a wide set of different areas of the world, taking as input macro-level data on electrification level, GDP per capita and population density. A rough estimation of micro-grid costs is given, assuming evenly distributed households spread over the considered areas. While still being a really simple approximation, it already introduces the concept that the internal layout of the mini-grids should be taken into account in cost comparisons with different alternative solutions. This is one of the aspects which is going to be taken explicitly into consideration in the work presented in this report.

2.2.5. The bottom-up approach

One of the main concepts that many of the cited papers fail to include in their analysis is the fact that the three alternative options are not meant to be exclusive with the respect to each other, but can also be considered different consecutive steps in time of a gradual bottom-up process. In fact, it is likely that communities having no energy access at all would also have an extremely limited access to finance and investment power for electricity infrastructure, making then more viable as a first electrification step the implementation of small solar home systems. This first electrification step would then lead to an empowerment of the single families and villages, granting higher incomes and more opportunities to invest in expansion of the single SHSs and eventually in a communal micro-grid connecting SHSs together to satisfy higher community demands and potentially communal loads devoted to productive use of energy. In this way, the transition from lack of energy access to climbing the energy ladders and moving from the lowest energy tiers to higher ones would be gradual and smooth, rather than a single one-time investment that could be not affordable for many poor remote communities. At an even later moment, the community may be then reached by grid extension, leading to the connection of an already strong and functioning micro-grid to the main electricity grid. This whole process is schematised in Figure 2.9.

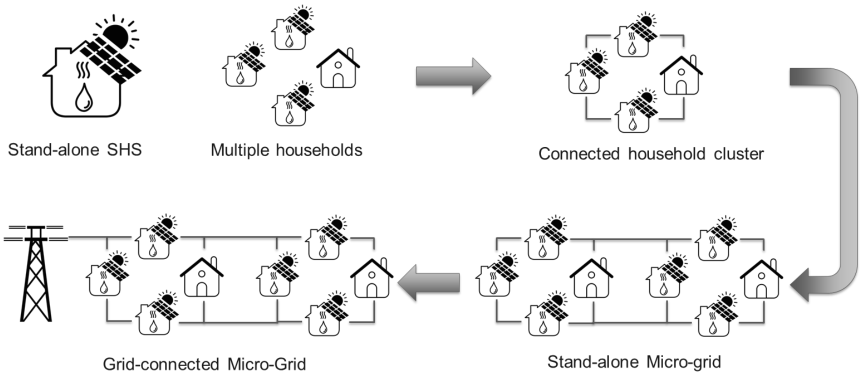


Figure 2.9: Bottom-up approach from SHS to micro-grid to connection to the main grid. Adapted from [13].

This bottom-up approach is well introduced and explained in [12], [13] and [21], also referred to with the term of *swarm electrification*. Such an approach leads to direct involvement of local people and increased responsibility for the owners of the Solar Home Systems and members of the community. Once the demand of the single household increases due to the positive feedback loop along the electricity ladder, micro-grid connection becomes feasible and more convenient with respect to over-sizing the solar home systems of the single households. In such a structure, energy exchange between members of the community is vital to grant a correct and efficient operation of the micro-grid, promoting the role of each single household as producer and consumer at the same time, thanks to the generation systems and storage systems which compose the Solar Home Systems. Another advantage of interconnecting multiple households in a micro-grid is a better utilisation of storage systems during operation in combination with other households demand and production profiles, leading to potentially improved battery lifetime [39]. Moreover, the modularity of Solar Home Systems and the easiness of connection of new storage or production units, as well as the easiness of connection of the SHS itself to a micro-grid infrastructure with a simple plug-and-play mechanism, make this approach at the same time easily scalable and less prone to major outages [12].

Inside this framework, it becomes crucial to properly design the interconnection layout in each community, identifying which is the most efficient and cost saving option to build up the micro-grid. Most of the papers which try to make a comparison between the three alternative solutions mentioned in this section, in addition to not considering this novel and useful bottom-up approach, tend to over-semplify the micro-grid connection option, making approximate assumptions on the costs and on the network layout at village level, losing then contact with the real implementation issues that may be encountered while physically installing a micro-grid. In this context, this thesis work aims to fill the gap found during this literature review, examining in-depth the possible micro-grid layouts which have been implemented in the past, their drawbacks and advantages and how they compare to each other.

In the next section, a more detailed analysis and description of solar DC micro-grids is going to be given, mentioning best practices and the most usual topologies which can be implemented in rural electrification micro-grid projects.

2.3. Solar DC Micro-grids

After some preliminary considerations and a short overview of what have been the trends in the process of rural electrification in the previous years, the main focus of this thesis work was set to the second option of the trilemma described in the previous section: stand-alone micro-grids.

2.3.1. What is a *micro-grid*?

The definition itself of the term *micro-grid* can be quite confusing and misleading, as it can have different nuances and different sources define it in slightly incoherent ways. In the context of this thesis work, the definition which is going to be used to define micro-grids along the whole work and report is the following:

> *"A micro-grid is a localised grouping of distributed energy resources, loads and energy storage devices that have the capability to operate in islanded and in grid-connected mode" [8]*

More precisely, micro-grids operating in islanded mode are going to be the main focus of this work.

The term micro-grid in recent years has often been related to, or even confused with, an other term belonging to the field of electricity networks, which is *smart grid*. It is important to clarify that a smart grid is an electricity network which is based on two levels of infrastructure: the physical power infrastructure composed of cables, converters, generators, loads, etc. and the communication infrastructure which manages and controls the operation of the network in a smart way [40]. Naturally, in order for a *micro-grid* to be fully efficient, being at the same time a *smart grid* is of foremost importance, but also transmission and distribution grids are meant to be *smart*. Hence, micro-grids are usually smart grids (or at least they should be), while not all smart grids are forcedly micro-grids. Even though micro-grids are a relatively modern technological concept, the amount of installations is already reaching interesting figures and there is enough experience on-field, testified by already well structured manuals and publications on many aspects of micro-grids applied in particular to rural electrification, for example [11] [41] [42].

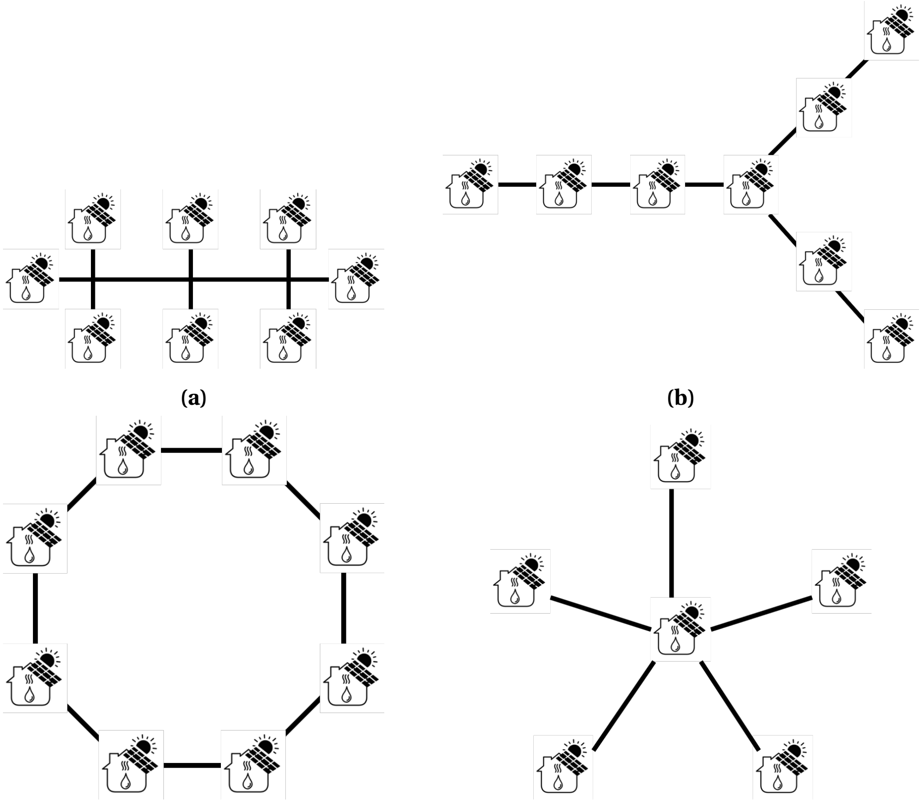
It is also important to make a distinction between AC and DC micro-grids (and the so-far most common option of hybrid AC/DC micro-grids). As trivial as it may seem, the difference between these types of micro-grids mainly consists in the type of generators and appliances which are connected to the grids, them being AC or DC or a mix of the two options. Even though AC micro-grids are a more mature technology and AC cables in general have been used for already more than a century, DC micro-grids present some advantages, especially in islanded applications and in combination with Solar Home Systems. First of all, knowing that batteries usually have a DC input and output and PV systems produce a DC energy output, DC networks are the most natural and obvious solution to transfer energy without any need of conversion steps between AC and DC and viceversa [43]. This fact lead to increased efficiency in energy exchanges, in addition the already lower cable losses that are typical of DC cables compared to AC cables, due to the absence of reactive power, which is a major cause of losses in AC networks [43] [44]. One of the possible downsides of DC micro-grids is an increased level of fault protection from short circuits [20].

It has already been described in the previous section what can be the advantages of implementing an interconnected micro-grid in rural communities, so in the next paragraphs the focus is going to be set on how to actually interconnect in an optimal way the households in a village, community or neighbourhood. This decision naturally depends on the spatial location and distribution of the households, but some typical examples of layouts can be identified in literature and in previous rural electrification projects.

2.3.2. Micro-grid topologies

In the past, before the great rise and diffusion of distributed renewable energy technologies, the easiest way to provide energy to remote communities, where grid extension was not possible or economically viable, was to install centralised generators (such as diesel generators) to which all households were connected via two possible layouts: spider topology, sometimes also called *"hub-and-spoke"* layout [11], meaning that each household is directly connected to the generator by a dedicated cable segment, or radial topology, in which multiple branches span from the central generator, reaching each household in a consecutive way along the branch, with the possibility of sub-branches being implemented [45] [46]. These two topologies present a major drawback in terms of fault tolerance, since if one of the cable segments gets disconnected because of a failure or of manumission, the household(s) which is(are) connected to that branch, will be totally cut out from energy supply and will suffer major power outages [20].

In the context of solar DC micro-grids, in which energy is produced potentially by every single household in the network, the two just mentioned network topologies are less prone to fault isolation then their centralised counterparts, due to the presence of generators (PV systems) in multiple locations along the branches. A representation of how a radial layout and a spider layout would look like in a micro-grid with decentralised energy generation is given in Figure 2.10b and Figure 2.10d respectively. Even though those layouts are typical of centralised generators, they are going to be anyway included into the comparative analysis between decentralised micro-grid layout which is going to be carried out in Chapter 4.



(a) (b) (c) (d)

Figure 2.10: Examples of possible micro-grid layouts: (a) Bus, (b) Radial, (c) Ring, (d) Spider.

Another kind of topology, which is actually applicable both to centralised and decentralised micro-grids, is called bus topology, or *"hub-and-trunk"* [11], and consists of a single (DC) bus to which all households are singularly connected (see Figure 2.10a). In [45], a comparative analysis of this DC bus layout and the already mentioned centralised radial topology is carried out, highlighting how the distributed bus solution is more easily modularised and scalable, fitting in the framework of a bottom-up approach to electrification.

Another possible layout option, which is mentioned and described, among others, in [46] and [47], is called ring topology, consisting of a consecutive connection of all the involved households, forming a closed loop which spans through and around the village/neighbourhood, as depicted in Figure 2.10c.

The last kind of topology which has been encountered in literature and which is becoming more and more common in micro-grids for rural electrification projects is the meshed grid topology. Meshed grids have the peculiar characteristics of having redundant cable segments, providing alternative paths for energy exchange. A schematic example of a meshed micro-grid can be seen in Figure 2.11. One major advantage of meshed micro-grids is an highly increased fault tolerance, since even if one or multiple cable segments fail

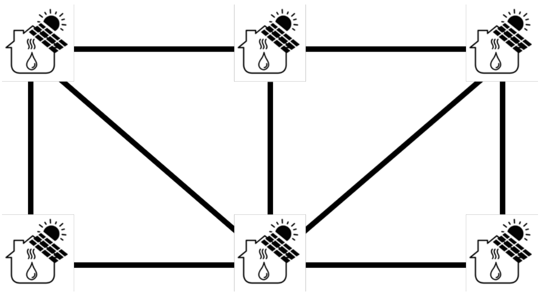


Figure 2.11: Example of meshed micro-grid layout.

or are tampered, the alternative paths manage to keep the whole micro-grid functioning properly. Another advantage is given by a more even voltage level distribution, which leads to lower fluctuations and differences in voltage between different sections of the grid [46] [48]. The obvious downside of this kind of topology is a more complex layout and higher costs due to more cable length needed to connect households with redundant paths.

2.3.3. Comparison of topologies in literature

It is peculiar to notice how most of the publications encountered during this literature review tend to stick to the most traditional topologies, namely spider, radial and bus topologies, without considering as a valuable alternative the promising meshed topology. There are still some good examples of papers taking into consideration also this last type of layout, which are going to be quickly described in the next paragraphs.

Having mentioned some of the most common micro-grid topologies, it is important to analyse and compare them to see which can be the best option for rural electrification projects. Of course, there is no straight answer to such a problem, because each single village, neighbourhood or community has different characteristics and household spatial distributions. In literature the most common approach is to directly select and apply a pre-defined topology, usually the radial or bus topology, without making a full preliminary comparison of different potential layouts on a quantitative basis. Nonetheless, some generalised comparison and analysis of the mentioned topologies have been found in literature. [46] and [48] are two examples of detailed research papers that compare between them three of the mentioned topologies, namely ring, radial and meshed layouts, trying to assess their advantages and disadvantages compared to each other. It must be mentioned that, even if the layouts analysed are considered also in this thesis work, both papers perform an AC power flow analysis, not a DC power flow, so their results are not directly comparable with the results of this thesis project, even if still relevant in terms of methodology and parameters used. Some of the parameters and characteristics which are used as meter of comparison are: cost of the network, fault tolerance, voltage fluctuations, adequateness to renewable energy penetration and reliability. In order to assess some of these parameters, a power flow simulation of some schematised theoretic samples of network are run, to compare operational behaviours of the considered topologies. The results of this qualitative comparison show how meshed networks perform better than the other alternatives in terms of voltage fluctuations and reliability, thanks to their redundant lines. A brief summary of some of the results of their analysis is shown in Table 2.1.

What seems to be missing the most, in these examples of generalised comparison between hypothetical sample networks, is the consistence with reality, an application to real cases, using ground-level data to evaluate which can be the best option for each particular village and community, performing at the same time a comparison on costs, represented by cable length needed to connect the households under certain assumptions, and on operation of the suggested micro-grids, to be assessed, as already done in [46] as an example, with (DC) power flow simulations. Another aspects which is not examined carefully enough is the comparison itself between alternative micro-grid layouts in term of geometrical design, since almost only radial, bus, ring and (in some cases) meshed topologies are considered to be classical and typically optimal solution, without properly considering other kinds of layout, which are going to be included instead in this work. The aim of this thesis work is to fill this gap making use of Geographic Information Systems to obtain and process ground-level data, as it is going to be explained in the next section and in Chapter 4, before smartly designing optimised meshed grids, as it is going to be described in Chapter 5, and finally running power flow simulations on the optimised networks in Chapter 6.

Table 2.1: Summary of the comparison of different network layouts performed in [46].

Network Layout	Generation	Stability	Reliability	Capital Costs	Maintenance Costs	Fault Tolerance	Protection Required	Renewable Penetration
Radial	centralised	low	low	low	high	low	high	problematic
Ring	distributed	high	medium	high	low	medium	high	accepted
Meshed	distributed	high	high	high	high	high	very high	moderate

2.4. Application of Geographic Information Systems

An extremely important aspect of rural electrification is the consistency with ground-level data which can be retrieved or processed through dedicated databases with the help of modern Geographic Information

Systems (GIS), such as ArcGIS or QGIS. In general terms, a Geographic Information System can be defined as any system which has the purpose to collect, analyse, process, represent and communicate spatial and geographic data.

GIS software has an incredible potential for manifold applications, giving the possibility to process incredibly huge amounts of data in a fast and automatised way, identifying trends, connecting variables, while keeping touch with real, geographically located information. GIS has already been used in a wide variety of fields, with different purposes and different methodologies: agriculture, astronomy, archaeology, geology, hydrology, infrastructure, aviation, nautical/marine applications, social sciences, real estate, environmental studies, military, disasters forecast and management, macro-economics, telecommunications, energy planning, statistics, history, IT applications, urban planning, land use planning, routing and many more [49]. This list already gives an idea of the versatility and potential of this kind of tools, which is becoming more and more widely spread and which bases its success on the availability of field-specific online databases, created, managed and used by a world-wide community of users. In this section, the focus is going to be set on how GIS tools have been used in the past only in the fields of energy studies and network planning, which are the most relevant for the purpose of this thesis project.

Why is GIS so important for rural electrification planning? First of all, one of the main hurdles for rural electrification projects is the lack of reliable, exhaustive ground-level data on location and distribution of settlements, as well as their (potential) energy demand, existing infrastructure and available resources. Without reliable ground-level data, any project is bound to be inaccurate and extremely sensitive to underlying assumptions. GIS software can fill this gap by using openly available databases online or by retrieving information through remote sensing from satellite imagery and, in addition, by analysing location-based energy information to extrapolate trends and forecasts for missing data. The more information is retrieved and used through GIS tools, the higher the accuracy and the conformity with reality of the analysis. Another useful function of GIS tools, which is not going to be used extensively in this piece of work, but which may prove useful in the communication of results and scenarios for electrification, is the great easiness of visualisation of data through map drawing which can be easily implemented and shared using Geographic Information Systems means.

The application of GIS tools to projects addressing the rural electrification and energy access issues is not something completely new. Some examples can be retrieved in literature, some of them already mentioned in Section 2.2, using to some extent GIS tools for tasks which are related to energy assessments and electrification planning, such as solar and wind resources evaluation [29] [50], evaluation of demographic parameters [38], cost estimation and comparison [28] [35], transmission infrastructure planning [30] [51] and decision making in location of generators [29]. Some of the most interesting examples and aspects of these research papers are going to be quickly described in the next paragraphs.

Once again, an interesting set of papers derives from the group of researchers which collaborated to the realisation of [28] and [35], in which they make use of GIS systems in various steps of their analysis to find the best electrification mix for specific countries. The main tasks for which GIS is used are the aggregation and processing of macro-level data on population density, energy resources such as wind and solar and infrastructure availability, in order to use them as parameters in a cost assessment of different electrification solutions and resulting in an optimal split between off-grid solutions and grid extension, mainly based on LCOE comparison. GIS aggregated data are also used to forecast the expected energy demand for each individuated geographical section of the country. Each step of the data analysis is also accompanied by a visualisation of the geo-spatially located information on dedicated maps, including the final results of the suggested electrification mix, as shown in Figure 2.12 for the case study of Nigeria examined in [28]. Another, more schematised way of geo-spatially representing the results was used in [35], including also the effect of different energy tier targets, as mentioned in Section 2.2, lowering the resolution of the representation, as seen in Figure 2.13.

Another good example of how GIS systems can be used to process large amount of energy-related data in order to plan and evaluate rural electrification projects is given by the online Network Planner tool used in [30] and [51], which is completely open source and can be potentially applied to any location in the world, as long as a sufficient amount of data is available. While incorporating a really well-thought algorithm, taking into consideration a wide set of geo-spatial input parameters, this tool relies heavily on the quality, resolution and availability of data, making use of default assumptions to substitute missing data, especially on the micro-grid dimension. This approach is bound to lead to great approximations and inaccuracies, hence the tool is still useful to make high-level generalised trend analysis, but for detailed planning alternative ways have to be found and implemented.

An interesting piece of work using GIS is also found in [37], which partially builds up on the work done in

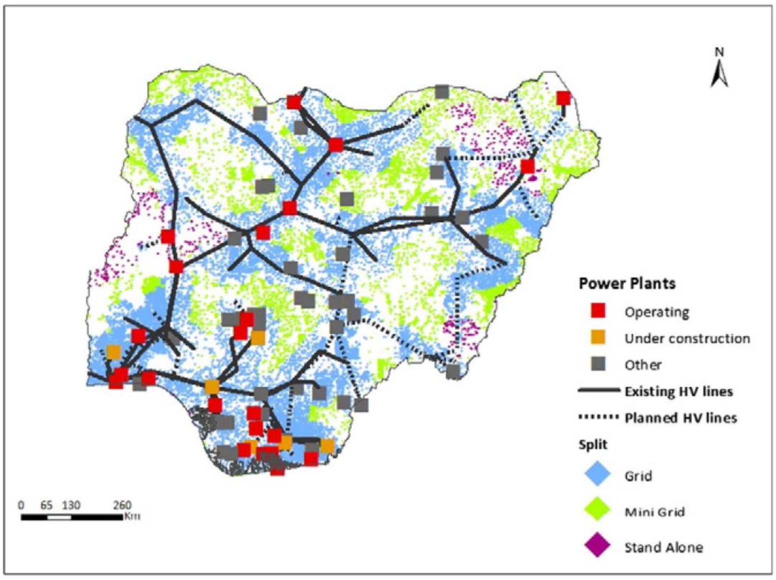


Figure 2.12: Electrification mix solution for Nigeria [28].

previously cited papers ([28] and [35]), adding the aspect of access to finance and of *willingness to pay* of local communities for electrification projects, if planned using a bottom-up locally-based approach as explained previously in this chapter. The combination of macro-level data on energy resources and infrastructure with finance aspects is a novel aspect of this work, which also includes a time dimension using GIS tools to subsequently evaluate at different steps in time which option would be the most convenient for each examined region.

After having analysed in the previous section what is the state of art of rural electrification projects and how GIS tools have been used in this context, some gaps to fill have been identified, which are going to be the basis of the work described in this report. The scope of this thesis project will then be to make use of GIS systems to analyse and process real ground level data, keeping contact with the dimension of the single village/community while evaluating the possible solutions for micro-grid planning. Rather than focusing on macro-level data and trends, which has been already done in a comprehensive and useful way in the cited papers, the focus here is going to be set on the village micro-grid level, filling the gaps that have been identified in this literature review. The methodology which will result from this effort is supposed to be applicable to any possible location worldwide without major changes, making it a useful, flexible, universal tool in rural electrification projects in the future.

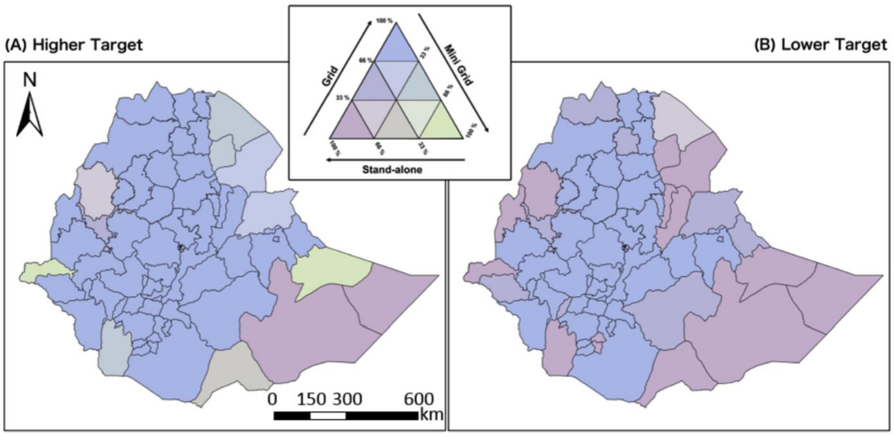


Figure 2.13: Electrification mix solution for Ethiopia with different electrification targets [35].

3

Graph Theory Applied to Network Analysis

A field of science which is too often underestimated by electrical or power engineers is applied mathematics. The manifold applications that can be found for mathematical concepts, theorems and algorithms span from infrastructure, to society, economics, chemistry and many more [52]. For the purpose of this thesis project, a review of the main basic concepts of Graph Theory is going to be made, focusing in particular to the potential applications for micro-grid topology design and more in general to network analysis. A deeper focus is going to be set on the Spanning Tree and Steiner Tree problems and how they relate to the case of micro-grid planning.

While an extensive analysis of Graph Theory in all of its aspects and concepts could be interesting, in this part of the report the focus will be mainly on the specific concepts and definitions that can prove useful for the purpose of the thesis project. In that way, the level of complexity of quantity of needed concepts is reduced and a purpose-centred approach in introducing this field of applied mathematics is used. It is important to note that is not a comprehensive summary of Graph Theory, so many concepts will be left out for sake of simplicity.

3.1. Basic Graph Theory Concepts and Definitions

As the term itself indicates, Graph Theory is that branch of mathematics which studies mathematical elements called *Graphs*. A Graph $G = (V, E)$ is a pair of subsets, in which V is a set of n vertices (nodes) and E is a set of m edges (lines). Vertices and edges are the fundamental elements composing any kind of graph and are the ones that define it. Each edge e in the set E is an object that connects between them two objects (nodes) i and j of the set V (vertices). Edges E are mathematically represented as 2-element unordered subsets of vertices V, namely the vertices which are connected to each other by the edge. An edge is said to be *incident* to the two vertices it connects, while each of the two vertices can be said to be *adjacent* to the other vertex to which it is connected through the edge. It is important to mention that edges can also be created connecting a single vertex to itself: in this case the edge is named a *self-loop*. During our study, anyway, treating electricity networks, *self-loops* will not be considered. Edges are usually assigned a weight, which is an important parameter for the analysis of the network characteristics. Weight can represent any kind of edge characteristic, depending on the kind of network which is being considered and on the purpose of the analysis. In the specific case of micro-grid networks, as an example, the length of each edge (line) in meters can be used as weight.

An important distinction has to be made between directed and undirected graphs. In undirected graph, each edge is only defined by the unordered couple of vertices it connects and, eventually, its weight. In directed graphs, also called *DiGraphs*, an additional information has to be provided to identify and define a single edge: its direction. This means that a single edge is not defined anymore by just an unordered couple of vertices (i, j), but by an ordered couple composed by a start node and an end node, making (i, j) different from (j, i). In practical terms, a single edge in directed graphs represents a two-way relationship between two elements of the set V, while in directed graphs it represents a one-way relationship from one single vertex to another one (or to the same one in case of a *self-loop*), meaning that in order to represent a two-way relationship between two vertices a pair of separate edges with opposite direction would be needed. Another category of graphs worth mentioning, even if not encountered during this thesis work, is *multi-graphs*. In *multi-graphs*, regardless of them being directed or undirected graphs, more than one single direct edge is available to connect two given vertices i and j.

Focusing on the characterisation of vertices, a concept that will prove useful during this thesis project is the *degree of connectivity* of a node. This can be defined as the number of edges which are incident to the considered vertex, including (and counting twice) self-loops. Some particular vertices can consequently be classified on the basis of their *degree*: vertices with degree equal to 0 are called *isolated* vertices, because they are completely disconnected from the rest of the graph and do not have any edge reaching them; vertices with degree equal to 1 can be called *end* vertices, because they constitute a dead-end of a branch of the graph; vertices which are connected directly through a single edge to any other vertex in the graph, so having a degree equal to $n-1$, are called *universal* vertices or *dominant* vertices. An example of dominant vertex can be the central node of the spider diagram mentioned in Chapter 2.

Some other useful definitions about graphs are the concepts of *connected* and *unconnected*. To better define these two concepts, first of all the concept of *path* has to be clarified. A *path* is a sequence of edges connecting an ordered sequence of vertices in a graph. A path from a node i to a node j, then, is a sequence of edges starting from node i and finishing in node j, possibly going through a finite or infinite number of other vertices on the way. It can be said that an edge (i, j) is a possible path from i to j, but it may not be the only one available in a graph. As can be deduced from the word itself, a *connected* graph is a graph in which there is at least one existing path connecting any pair of vertices in set V through a number of edges in set E. On the contrary, if it exists at least one pair of vertices which are not connected by any path in the graph, then the graph is considered *unconnected* or *disconnected*. In other words, if any single vertex, or any single subset of vertices, is not reachable from all the other vertices in the graph, then the graph is *unconnected*. This thesis work will focus mainly on connected graphs, namely single connected micro-grids, reaching all the households (vertices) of a single village. The graphical difference between a unconnected and a connected graph can be seen in the example represented in Figures 3.1b and 3.1c respectively.

In addition, the concepts of *empty* and *complete* graphs can be introduced. An *empty* graph is a graph $G = (V, E)$ in which the set E is empty. In other words, there are no edges in the graph, which is then only composed by an unconnected set of isolated vertices V having degree equal to 0, without any possible existing path between any pair of vertices, as shown in the example in Figure 3.1a. Applying this definition to the cases examined in this thesis work, an empty graph is the representation of the village households before connecting them through a micro-grid. On the contrary, a *complete* graph is a graph in which each vertex i in the set V is directly connected to any other vertex j by a direct edge (without self-loops), as shown in the example in Figure 3.1d. As a side consideration, it can be said that every vertex in a complete graph is a universal vertex. Complete graphs can also be described specifying that every pair of vertices (i, j) can be connected by a path composed by a sequence of one single edge. Through simple combinatorics it can be calculated the number of edges in a graph having a set of n vertices. This is different, though, for directed and undirected graphs. For directed graphs, the number of edges m can be calculated as:

$$m = n \cdot (n-1) \tag{3.1}$$

while for undirected graph the formula becomes:

$$m = \frac{n \cdot (n-1)}{2} \tag{3.2}$$

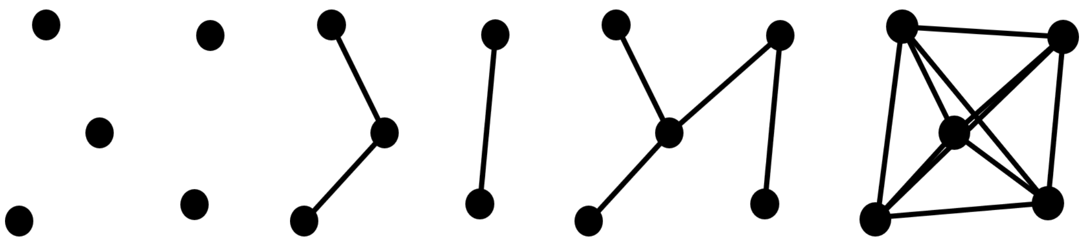

(a) **(b)** **(c)** **(d)**

Figure 3.1: Example of representations of different graphs for the same set of vertices V: (a) Empty graph (b) Unconnected graph (c) Connected graph (d) Complete graph.

After having defined these terms and concepts typical of graph theory, it is important to understand how graphs can be mathematically represented with a computational purpose. While it is simple to understand how nodes, edges and their weight can be simply represented to a network, it is less straight-forward to represent in a non-graphical way the correlation between nodes and edges, in order to easily store it and process it. This can be done using matrix structures such as the *adjacency matrix* and the *incidence matrix*.

An *adjacency matrix* A is a square $n \times n$ matrix, representing the structure of the graph in terms of adjacency between its vertices. The adjacency matrix is built as follows:

$$A_{ij} = 0 \text{ if } (i,j) \notin E \text{ (if i and j are not adjacent)}$$
$$A_{ij} = 1 \text{ if } (i,j) \in E \text{ (if i and j are adjacent)}$$

From adjacency matrices information about the layout of a graph can be easily retrieved and processed. To give some examples, the adjacency matrix of an empty graph would have all elements equal to 0, while the adjacency matrix of a complete graph would have all elements equal to 1, except for the diagonal elements, which represent self-loops. It can be easily deduced that adjacency matrices are always symmetric for undirected graph, while they can be not-symmetric for directed graphs.

Defined differently, *incidence matrices* I are $n \times m$ non-square matrices which represent the structure of the graph in term of incidence of each edge e with each vertex v of the graph. The incidence matrix is built as follows (for an undirected graph):

$$I_{ve} = 0 \text{ (if e is not incident with v)}$$
$$I_{ve} = 1 \text{ (if e is incident with v)}$$

Each edge is hence identified in the matrix by means of the two vertices it connects. A little different structure is used for incidence matrices of directed graph: in each edge column e, the start node is identified with a value of -1 and the end node is identified with a value of 1, giving then information on the direction of the edge as well.

After this brief general introduction on Graph Theory and its concepts and definitions, in the next section a deeper explanation of some more advanced concepts and some algorithms is going to be given, keeping in mind the specific purpose of this thesis work.

3.2. Trees

Another graph theory which is extremely important in network analysis, and in the context of this thesis project in particular, is the concept of *trees*. A really thorough and exhaustive overview and description of all the graph theory concepts related to trees can be found in [53]. A simple definition of a tree in graph theory is "an acyclic connected graph", meaning an undirected graph in which no loops (cycles) are present, so that each pair of nodes in the tree is connected by means of a single unique path. It can be easily deduced that, for any given set of n nodes, a tree is composed of $m = n - 1$ edges. It is important to stress the fact that a tree is a sub-graph of the given graph $G = (V, E)$ which connects at the same time all the vertices included in a given subset of V, using edges included in set E. This defines the difference between a *tree* and a *forest*, which is an unconnected union of multiple trees inside a given Graph G. If all the vertices in set V are connected by one single tree, that tree is called a *spanning tree*.

3.2.1. Minimum Spanning Tree

The concept of *trees* is extremely important in network analysis, because it can be applied to many practical problems in which the aim is to connect between each other a subset of elements (nodes, vertices, people) using the lowest possible number of relations (edges, lines, connections). This kind of problems usually have minimisation of costs as a main criteria for connecting the desired elements. That is the reason why one of the main geometrical problems which have been addressed by graph theory researchers is a way to obtain the shortest (or least costly) spanning tree for a given graph G, or a given subset of the vertices V in a graph G, which is called *Minimum Spanning Tree* (MST). Naturally, if edges have no weights, the problem itself is trivial since every spanning tree would have the same weight (null), regardless of the layout of the tree, so this kind of problem is only applied to weighted edges. An example of a Minimum Spanning Tree of a weighted graph is shown in Figure 3.2. The Minimum Spanning Tree problem has been discussed for decades, with many

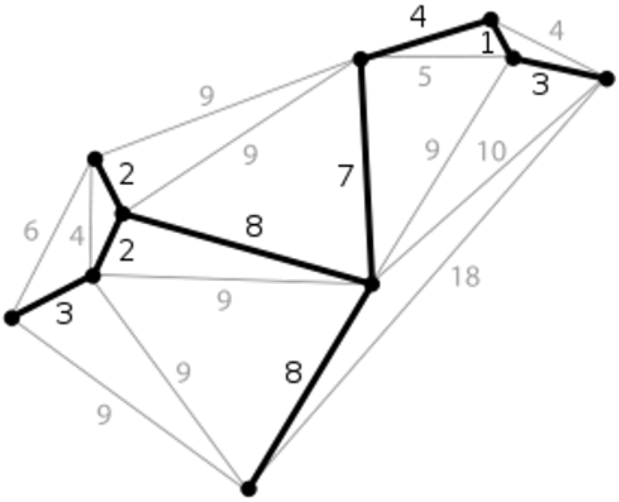


Figure 3.2: Example of a Minimum Spanning Tree for a Weighted Graph [52].

different algorithms found in literature to solve it, the most known and efficient solutions being Kruskal's [54] and Prim's [55] algorithms, developed and published in the 1950s [52] [56].

To briefly describe the rationale of Kruskal's heuristic, the steps of the algorithm can be summarised as follows [54]:

1. Take the starting graph $G = (V,S)$ for the set of vertices V, knowing the weight of each single edge contained in S;
2. Start creating a tree $T = (V,E)$, E being an empty set of edges;
3. Remove from S and add to E the edge in S having the lowest weight which does not create a loop in graph G;
4. Repeat step 3 until a spanning tree is formed in T;
5. The obtained spanning tree T is a Minimum Spanning Tree (MST) for the set of vertices V and the set of edges S. If G is a forest, then T will be a Minimum Spanning Forest instead.

This algorithm can be also applied to connect an undefined number n of vertices with the shortest possible spanning tree by setting the starting graph G to be the complete graph for the set of vertices V, using as weight of the edge the euclidean distance between one vertex and the other. In that way, S would be the set of all possible edges connecting any pair of vertices in set V. The example in Figure 3.2 can be used as a reference to better understand the principle of the algorithm.

A slightly different approach is used in Prim's algorithm, while still using the weight of each edge as a criterion for the choices of the edges to be part of the Minimum Spanning Tree. Before stating the steps, it is needed to quickly define the concept of *nearest neighbour*, which means the element (vertex) which has the shortest distance (weight) in the graph G from a given element (vertex or sub-tree). The steps of the algorithm can be rewritten as follows [55]:

1. Take the starting graph $G = (V,S)$ for the set of vertices V, knowing the weight of each single edge contained in S;
2. Start creating a tree $T = (V,E)$, E being an empty set of edges;
3. Perform one of the following actions:
   (a) Add to set E the edge in S connecting any vertex in V to its nearest neighbour;
   (b) Add to set E the edge in S connecting any sub-tree in T to its nearest neighbour;
4. Repeat step 3 until a spanning tree is formed in T;

5. The obtained spanning tree T is a Minimum Spanning Tree (MST) for the set of vertices V and the set of edges S. If G is a forest, then T will be a Minimum Spanning Forest instead.

It is easy to see how the working principle of the two algorithms is similar, while having a different approach in the consequential way with which the Minimum Spanning Tree is created from scratch.

Both algorithms are particularly useful for euclidean applications to connect points in a geometrical plane using Euclidean distance as weight, because the weight of the potential edges does not need to be known as an input of the algorithm, as long as the geometrical location of the points is known and can be used to calculate the distance between each pair of points.

3.2.2. The Steiner tree problem

Another well-known problem dealing with trees on the euclidean point of view is the Steiner Tree problem, which, in its Graph-applied version, aims to find a minimal weight tree connecting a subset of vertices (terminals) in an undirected graph with non-negative weights, eventually including additional vertices (called in this context *Steiner points* S, which are not included in the subset [52]. It must be noted that, in case the subset of vertices contains all the vertices contained in the starting graph, then the problem is reduced to the already mentioned minimum spanning tree problem. During this thesis project, a particular version of the Steiner tree problem results particularly useful: the Euclidean Steiner Tree problem in the plane. In this particular version of the problem, the weight of each edge is the euclidean distance between the two terminals point which it is connecting [57]. The set of terminals (geometrical points which given x-y coordinates) is considered a subset of a greater set of potential terminals which contains all the points in the euclidean plane. This means that any point in the plane can be used in the drawing of the minimum Steiner tree for the given subset of vertices. To better understand the concept, the case for three vertices is going to be briefly shown. Before doing that, the concept of Fermat point for a triangle has to explained. The Fermat point of a triangle is the point in which three segments touching respectively the three vertices of the triangle meet each other forming three 120° angles. The Fermat point falls inside the triangle area if all the angles of the triangle are smaller than 120° and outside of it if one of the angles is bigger than 120°. One important characteristic of the Fermat point (when it falls inside the triangle), is that it minimises the sum of the lengths of the three segments incidents to it and to the triangle vertices. Any other set of three segments incidents to the triangle vertices and another point inside the triangle area will have a higher sum of their lengths. It follows that the solution for the minimal Steiner in the euclidean plane for three vertices can be of two natures:

- If the triangle formed by the three vertices has every angle smaller than 120°, then the shortest Steiner tree is composed by the 3 edges connecting each vertex to the Fermat point of the triangle (see Figure 3.3a).

- If the triangle formed by the three vertices has one angle bigger than 120°, then the shortest Steiner tree is composed by the 2 edges adjacent to that vertex. In this case the Steiner tree does not need the addition of any Steiner point S, so it reduces to be equivalent to the Minimum Spanning Tree (see Figure 3.3b).

As the number of vertices increases, the complexity increases to be a NP-hard problem, but any Steiner point S added to the tree will still be a Fermat point of the triangle composed by three of the vertices in the subset that needs to be connected.

After this short description of the concept of *trees* and some particularly relevant example of tree problems, in the next section some examples of potential applications of graph theory to different fields in network analysis are going to be shown. Following, it is going to be explained in more detail the approach used in this thesis work to apply graph theory to network analysis on the micro-grid planning level.

3.3. Graph Theory Applied to Network Analysis

The potential applications of graph theory are manifold and customisable on an almost infinite scale. Basically any kind of natural, physical, tangible or intangible phenomenon that can be schematised in terms of elements and relations or connections between those elements is potentially keen to be studied and analysed using graph theory. The most straight-forward kind of application is related to physical networks connecting physical objects in the euclidean space. Typical examples can be planning (or maintenance) of road networks, water supply networks, electricity transmission grid, GPS navigation systems and many more. Sticking to real

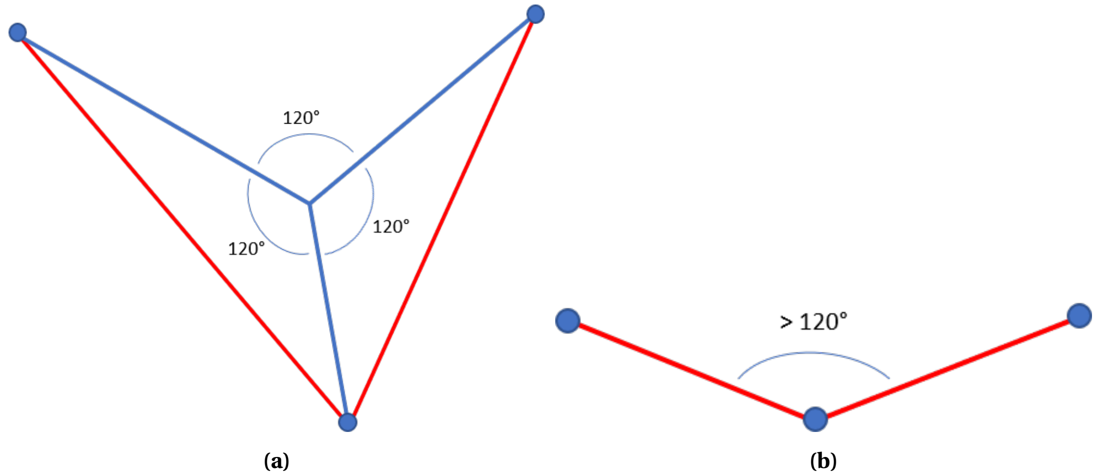


Figure 3.3: Solution of the Euclidean Steiner Problem (blue) compared to Minimun Spanning Tree (red) for three vertices: (a) Fermat point inside the triangle (b) Fermat point outside the triangle.

physical objects, even if on a different scale, molecular structure of chemical substances can also be represented with the same concepts, as well as living beings' neural or circulatory systems. Another field in which graph theory is widely used is communications and computer science, in which data and computational flows can be easily represented as graphs, for example describing a web-page link structure. Communications on social media can also be tracked and analysed using graph theory, for example identifying people as nodes and connections (Facebook friendship, LinkedIn connection, etc.) as edges. Even if the representation as nodes and edges is not as straight-forward as for physical networks, business relations or commercial transactions can also be represented and analysed using graph theory, as well as, reaching an even higher level of abstraction, sociological phenomena and literature studies.

By this really varied and apparently totally disconnected list of potential applications, it is possible to understand how flexible and useful such a set of concepts can be. In this particular case, graph theory is going to be used as a tool to evaluate, analyse and compare electrical networks, namely micro-grids. In all steps of this analysis, the vertices in the graphs would be the single nodes of the micro-grid, which are the single households inside a village that need to be connected to each other. Trivially, the edges of the graph would be the lines (cables) of the electrical network physically connecting the households to each other. One major advantage of using these concepts is the great flexibility they offer, making it viable to study different parameters and aspects of micro-grids with just minor modifications. As an example, different aspects of the micro-grids can be analysed considering different characteristics of the lines as weight of the edges: using the physical euclidean length of the cables as weight is a good way to analyse and compare cable costs for different layouts, while using cable resistances as weight is useful for power flow analysis and for losses calculations.

Examples of graph theory applied to electrical network problems are already found in the past, on different levels and with different scopes and purposes. A really extensive overview of how electrical networks can be analysed and represented using graph theory, focusing mainly on the circuit modelling point of view, is given in [58]. The specific aim of this thesis work is the generation, analysis and comparison of different micro-grid topologies, on the scale of intra-village interconnection, under a geometrical (relatable to cost considerations) and electrical point of view at once, starting from real ground-level data.

As mentioned, the most direct parameter which is going to be used to compare between them different topologies for micro-grids is the length of the network itself, more precisely the distance that needs to be covered with cables per household in each considered solution. No complex explanation is needed to understand the importance in terms of costs of keeping the length of the network as small as possible. In order to actually derive the costs of a network, other aspects need to be taken into considerations, such as the type of cables used and their thickness as well as the cost of other components of the network such as poles, but this falls outside of the exact scope of the thesis project, so the length is going to be used directly to represent the cost of the network and to compare alternative solutions.

Another useful parameter to be analysed is the *average shortest path length* of the network. In order to understand the usefulness of this parameter, it is necessary to take a step back and define the shortest path problem in general. Having already defined previously what a *path* is, the shortest path problem aims to find, trivially, the path between two vertices i and j in a weighted graph G which has the lowest possible sum of the weights of the edges included in the path. Many different versions of the problem can be formulated, depending on the graph being directed or undirected and on which vertices are taken into consideration. The simplest form is the single-pair shortest path problem, which consider only one pair of vertices i and j, i being the starting point of the path and j the end point of the path. This version is typical for example of routing problems, such as finding the fastest route to drive from a location to another on a road network. More complex versions can include the computation of the shortest paths from one single vertex i to all other vertices in the graph or even for all pairs of vertices (i, j) at once. There are many examples of algorithms providing solutions to the different versions of the problem, including among the others the Dijkstra algorithm [59] and the Floyd-Warshall algorithm [60], which addresses specifically the all-pairs shortest path length problem and is particularly useful for the calculation of the *average shortest path length* of a graph, by simply averaging the length of all the shortest paths found by the algorithm. As will be further explained in Chapter 5, the average shortest path length can be used as a representation of some characteristics of the network, such as redundancy, congestion avoidance and losses reduction. This is the reason why this parameter is going to be used for network comparison and optimisation.

While these Graph Theory concepts are extremely useful for the analysis of proposed network solutions, the true value of graph theory for this thesis work and for tackling rural electrification problems lays in using its algorithms and structure actively in the phase of network layout choice. Once the households of the considered village are identified, they can be represented as vertices and connected between them in many different ways, which can be designed using imposed criteria and layouts, such as the Spider Diagram and the Bus Topology mentioned in Chapter 2, or can be generated directly using those algorithms or taking inspiration from them. That is the reason why, in addition to the typical micro-grid topologies encountered and described in literature, some more graph theory-based solutions are going to be considered as well.

First of all, the Steiner tree, defined in the previous section, is going to be used as a benchmark for comparison, being the universally shortest possible solution to connect a given set of vertices in the euclidean plane without any constraint, solely considering network length. The Steiner tree can be found using known algorithms and does not require any further modification nor assumption. Another layout solution which can be directly adopted from graph theory using known algorithms is the Minimum Spanning Tree, which differs from the Steiner tree by not considering the use of additional nodes (Steiner points, see Figure 3.3). Two other typical topologies found in literature, the Ring topology and the Radial topology, can be implemented and designed using concepts adopted from Graph Theory. The Ring topology can be seen as an *Hamiltonian circuit*, which is a closed path in the given graph that touches exactly once every vertex. In an Hamiltonian circuit, every vertex is adjacent to two other vertices. Concerning the Radial topology, which consists of a central root with radial branches reaching out to the other vertices, the design can be inspired by the Esau-Williams algorithm for capacitated spanning trees, which starts from a spider diagram with a central root and subsequently merges together edges to create branches with cost (weight) lower than the sum of the merged edges [61]. The Radial topology considered in Chapter 4 is generated with a similar rationale of weight-saving, even if not using the Esau-Williams algorithm directly. As mentioned in Section 2.3, meshed grids are also a valuable solution for micro-grid layout. Even though meshed topology is not clearly identified as a known graph type, Graph theory concepts and parameters are anyway going to be used to generate and analyse optimised meshed grids solutions.

Some examples of application of graph theory to network planning at the micro-grid level are already found in literature, usually focusing on just one algorithm or one suggested starting topology, without usually making an extensive comparison of different possible solutions [62], while a more consistent amount of research papers and developed tools focus on transmission and distribution grids rather than micro-grids level [30] [51]. Once again, this thesis work has the purpose to fill the gap in literature and focus on micro-grid planning, evaluating the just mentioned different layout solutions under both a geometrical and electrical point of view.

After this brief introduction to Graph Theory and to how its concepts and algorithms are going to be used in this thesis work, the following chapters will describe the actual methodology designed and applied during the thesis work, showing step by step results. To sum up the concepts and definition contained in this chapter which are going to be used in the following parts of this report, Table 3.1 is shown in the next page.

Table 3.1: Glossary of the most relevant terms and concepts of Graph Theory.

Graph	Fundamental object of graph theory, composed by a set of vertices V connected in pairs by a set of edges E.
(Un)Directed Graph	Graph with (un)oriented edges.
Node (vertex)	Basic unit of Graphs. Punctual objects without internal structure.
Edge	Basic unit of Graphs. It connects two nodes (endpoints) to each other. Can be directed or undirected.
Incident	An edge is incident to a vertex (and viceversa) if the vertex is one of its endpoints.
Incidence Matrix	n x m Matrix representing the incidence relationship between edges and vertices.
Adjacent	A vertex is adjacent to another vertex if they are endpoints of the same edge.
Adjacency Matrix	n x n Matrix representing the adjacency relationship between vertices.
Weight	A numerical value assigned to an edge (or a vertex).
Degree (of connectivity)	Number of edges incident to a vertex in a graph.
Isolated Node	Vertex with degree = 0.
End Node	Vertex with degree = 1.
Dominant Node	Vertex adjacent to all other vertices (degree = n - 1).
Path	Sequence of consecutive edges to reach a vertex starting from another vertex.
Loop	Path starting and ending in the same vertex.
Empty Graph	A Graph with no edges.
Complete Graph	A Graph in which all pairs of edges are adjacent.
Tree	Connected, acyclic Graph in which each pair of vertices is connected by a unique path.
Spanning Tree	A tree including all the vertices in the set V of a Graph.
MST (Minimum Spanning Tree)	The shortest(lowest weight) possible spanning tree in a given Graph.
Steiner Tree	A tree connecting a subset of vertices, eventually using additional vertices not included into the subset, but included into the set V of graph G.
Shortest Path Length	Length of the shortest (lowest weight) possible path between two vertices.
Hamiltonian Circuit	A loop touching exactly once every vertex in the Graph.

4

Geographic Information Systems Analysis

As already described in Chapter 2, Geographic Information Systems are a powerful tool for site analysis, remote sensing, trend identification, infrastructure planning and many other kind of applications. For the purpose of this thesis project, GIS were used for different subsequent tasks.

First of all, locations for off-grid rural electrification case studies were identified in different areas of the world. After having localised a set of villages which were suitable for the purpose, the layout and distribution of the houses, together with the village boundaries, GIS was used to derive parameters and characteristics of those villages. The following step was to develop a methodology to design different layouts for micro-grids for each village and compare them. In order to validate the accuracy of the optimal layout found with this methodology, an external algorithm on Maple developed by Dr. Ir. Petra Heijnen (TU Delft - faculty of Technology, Policy & Management) was used. Another kind of analysis that was performed using GIS software was the calculation of the actual distance of some of the identified villages from the existing electricity grid, as a basis for comparison between grid extension and micro-grid as rural electrification solutions.

In this chapter, each of these steps (shown as an integrated flowchart in Figure 4.1) is going to be described, with its rationale and its procedure, and results are going to be presented and discussed. As many details on the procedure as possible are going to be given, so that the reader would be able without major efforts to replicate the methodology at a later moment. The outputs of this section are going to be used in the layout optimisation phase which is going to be described in Chapter 5.

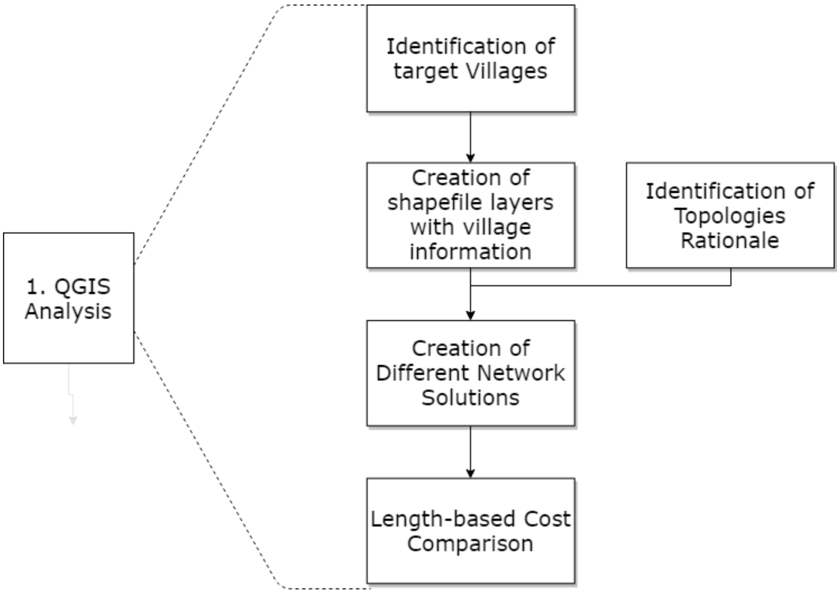
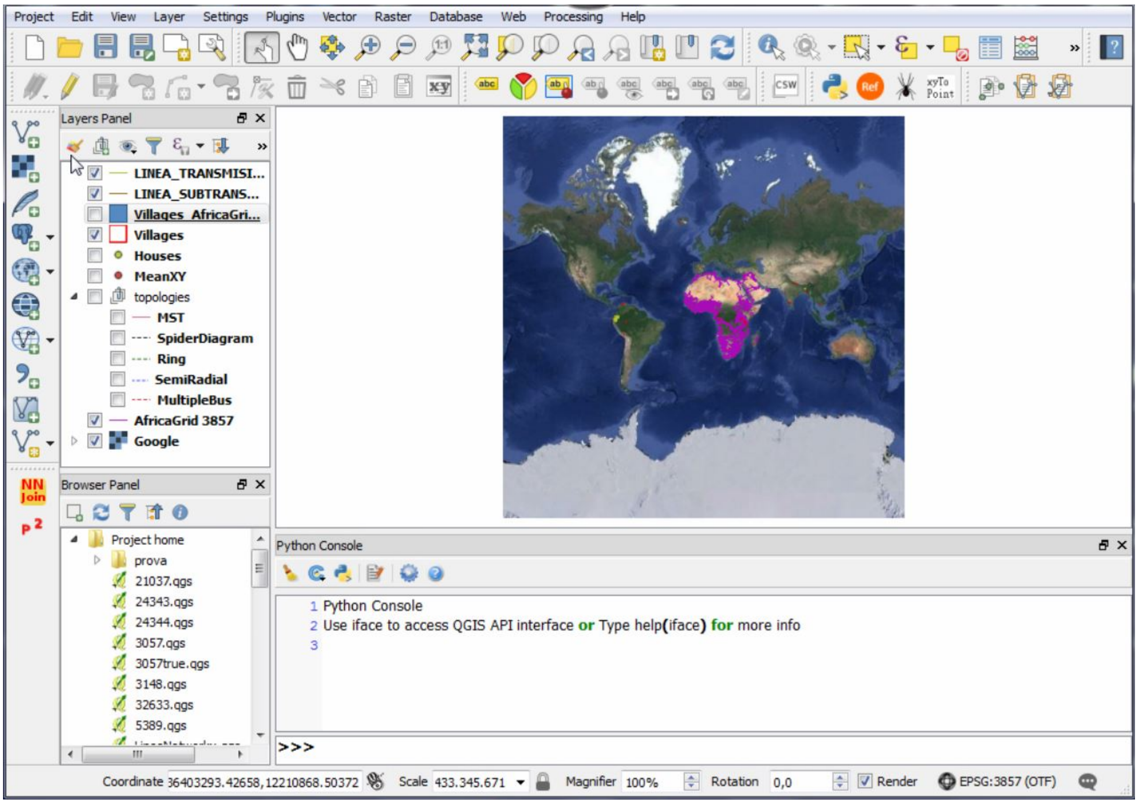


Figure 4.1: Flowchart of QGIS analysis methodology steps.

4.1. GIS software

In general terms, a Geographic Information System can be defined as any system which has the purpose to collect, analyse, process, represent and communicate spatial and geographic data. Google Maps and Google Earth are examples of very simple Geographic Information Systems which are commonly used by the wide public, probably without even knowing what a GIS is.

As a first step into this GIS-based approach, it is important to evaluate and choose which is the most appropriate and adequate software to use, among the possibilities that the GIS world offers. This choice is not as trivial as it may seem, because of the huge variety of GIS software available on the internet, ranging from proprietary to open-source alternatives, differing between each other in terms of coding language, user interface, built-in tools available, potential applications and easiness of integration with other complementary tools. During the beginning of the thesis period, ArcGIS for Desktop by ESRI (Environmental Systems Research Institute) [63] was chosen, for his long-term presence in the market and the extremely wide set of already implemented tools for geo-spatial analysis, network analysis, mapping and many other GIS-based functions. ArcGIS was one of the first complete commercially released Geographic Information Systems, launched in 1999, so that the most used data format for storing vectorial geospatial data takes its name from ESRI itself (ESRI shapefiles, .shp). The main drawback of ArcGIS is the multiple license levels available, limiting access to the most advanced built-in tools to the higher levels of permission, which are not available for free.



Python Console
Use iface to access QGIS API interface or Type help(iface) for more info

```
>>>
```

Coordinate 36403293.42658,12210868.50372 Scale 433.345.671 Magnifier 100% Rotation 0,0 ☑ Render EPSG:3857 (OTF)

Figure 4.2: QGIS user interface

That is the reason why, for the full scope of this study, it was chosen to use another GIS, which is fully open source: Quantum GIS (QGIS) [64]. QGIS has an interface which is quite similar to the ArcGIS interface, providing a visualisation window for geographical features, a customised toolbar, a browser panel, a layer panel and a useful built-in Python console, including an internal editor, as shown in Figure 4.2. The whole software is written in Python and C++, and it is possible to use Python language through the command window to manually execute functions, both internally and externally to the software itself. Even though the amount of built-in tools and functions is limited compared to ArcGIS, this gap is filled thanks to a well-furnished internal repository of plug-ins, developed by users or external parties, which can be easily activated and used in QGIS. QGIS can also be integrated with other GIS based applications which provide additional functions,

Figure 4.3: Location of identified case study villages, as red dots on the world map.

such as PostGIS, GDAL, GRASS GIS, PostgreSQL and SQLite. Similarly to ArcGIS, all the data is stored in layer files, having the typical format of ESRI shapefiles (.shp), which is a composite ensemble of different files storing geometrical features (points, lines or polygons), their geographical location and a given set of attributes. This advantage was fundamental in the choice of migrating the project from ArcGIS to QGIS for two main reasons: the possibility of migrating the files from one system to the other without major modifications; the possibility of using data from layer files and maps stored in the online ArcGIS repository, which proved useful particularly in the phase of data gathering which is going to be described in the following section.

The most updated version of the software that was released at the time when the thesis work was started, was QGIS 3.0 Girona, while the latest update (QGIS 3.4 Madeira) was released in October 2018. For this thesis work, nonetheless, it was chosen to use a previous version (QGIS 2.18 las Palmas de Gran Canaria), because some of the plug-ins needed for this study were only updated to be compatible to this older version.

4.2. Data gathering and processing in GIS

The first step for the GIS analysis of this thesis work was naturally the identification and localisation of a sample of villages to be used as case studies, from different rural areas of the world. It was chosen to look for villages in different countries and different environments, to give a varied sample, and to focus mainly on small villages, with a number of households up to around 40, to take into consideration the extremely early steps of settlement development. In the process of choosing the single villages, it was also chosen to select settlements showing different geometrical household distribution patterns, once again for sake of variety and diversity of the sample. Part of the sample was selected starting from available databases for real rural electrification projects, while other villages were chosen from specific targeted countries with energy access issues and based on their distance from available infrastructure. For example, the set of Tanzanian villages were retrieved from the REDP (Rural Electrification Densification Programme) [65] database available on the online ArcGIS repository. Similarly, for villages in Bangladesh an available database on "hard-to-reach" communities was used to pick case studies [66]. The majority of the other case studies were chosen on the basis of country data on access to electricity for rural population published by the World Bank Group [2].

Table 4.1: Ranges of parameters for selected sample of villages.

	Min	Max
Number of Households [hh]	5	40
Density [hh/km^2]	93.58	5016.03
Village Area [m^2]	2851	145324

The result of this identification process was a shortlist of forty-two (42) villages, spread over twelve (12) different countries, as shown in Figure 4.3. Each of the village was assigned a unique identification number from 1 to 42 in order to distinguish them. The sample of villages, as shown in Table 4.1, ranges from 5 to 40 households per village, from a density of around 100 households per square kilometre to around 5000 households per square kilometre and from a village surface of around 3000 square meters to around 150000 square meters. The countries which were included into the analysis are Tanzania, Gabon, Cameroon, Madagascar, Bangladesh, India, Nepal, Cambodia, Colombia, Peru, Brazil and Iceland. Iceland was included just to give an example of an already developed country and to show how the methodology can also be applied to isolated areas of rich countries. All the other selected countries have low rural electrification rates and the chosen villages are currently disconnected from the main grid.

Once the sample of villages had been selected, the next step was to derive, from the geographical map, the geometrical features, in terms of point, lines and polygons, of each of the villages. It is important to mention that the world map which have been used in this project is an Hybrid Map, showing geographical natural features as well as political borders and basic infrastructure [67]. The map, which is in raster format (so not readable as a shapefile layer by QGIS), was imported using the built-in "Tile Server connection tool" in QGIS. Another base map was considered for the project, from the OpenStreetMap database, which included a rasterised version of buildings represented as polygons in the map, not using satellite imagery. This polygonal rasterisation was only available for urban areas, while data were missing for isolated rural areas, making it not fitting with the purpose of the thesis work.

While for the localisation of the sample villages one single reference system was sufficient, to go deeper into the geographical and geometrical analysis, every single area of the world must be analysed with the correct Coordinate Reference System (CRS). The identification of the correct CRS for each single villages is

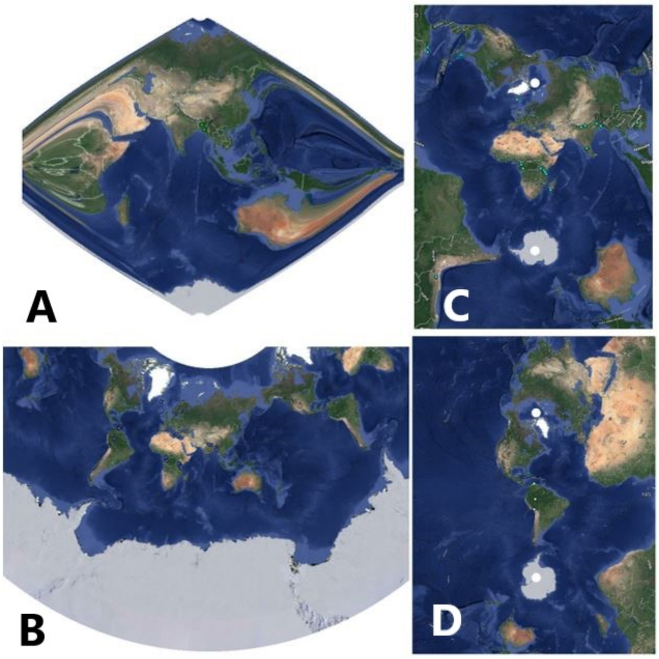


Figure 4.4: Different world map projections with different CRS. A: EPSG 3106. B: EPSG 24344. C: EPSG 21036. D: EPSG 31979.

Table 4.2: List of Coordinate systems used for the selected villages.

Projected Coordinate System	Villages
EPSG:5389 Peru96/UTM zone 19S	1,12 (Colombia), 4,5 (Peru), 10,11 (Brazil)
EPSG:3057 ISN93/Lambert 1993	2,3 (Iceland)
EPSG:31972 SIRGAS 2000/UTM zone 18N	6,7,8,9 (Colombia)
EPSG:21036 Arc 1960/UTM zone 36S	26,28 (Tanzania)
EPSG:21037 Arc 1960/UTM zone 37S	13,15,16,17,18,19,20,21,22,23,24 (Tanzania)
EPSG:24343 Kalianpur 1975/UTM zone 43N	30,31,32,33,34,35 (India)
EPSG:24344 Kalianpur 1975/UTM zone 44N	25 (Nepal), 40 (India)
EPSG:29738 Tananarive/UTM zone 38S	14 (Madagascar)
EPSG:3148 Indian 1960/UTM zone 48N	27,29 (Cambodia)
EPSG:3106 Gulshan 303/Bangladesh Transverse Mercator	36,37,38,39 (Bangladesh)
EPSG:32633 WGS 84/UTM zone 33N	41 (Cameroon), 42 (Gabon)

extremely important, because the correlation between geographical coordinates (latitude and longitude) and physical measures, like lengths and surfaces, changes consistently, depending on which location of the globe we are considering.

A CRS can be distinguished first of all by its origin, in which the x and y coordinates are set to 0. Trivially, the closer the origin is to the considered location, the more accurate are the measurements performed by the GIS. Secondly, another important feature of each CRS is the kind of projection it uses to reflect on a planar map the non-planar geometrical shape of the terrestrial globe. There are many different possible ways to do that, but the four main types of projections are conic, cylindrical, planar and polar. Examples of different types of map projections are shown in Figure 4.4. Each CRS has a preferred area of application, in which it can grant optimal accuracy. Each CRS has a unique code, registered in the EPSG (European Petroleum Survey Group) Geodetic Parameter Dataset, which is a world-wide recognised register of all the available CRS [68]. ESRI published on his website a complete updated list of all the CRS which can be used into GIS software, specifying type of projection, origin and suggested areas of application [69]. The default CRS of commonly used maps for global visualisation (including the Google Hybrid Map used for village localisation) is WGS 84 / Pseudo-Mercator (EPSG code: 3857), which is centred at the intersection between the Equator and the Greenwich meridian, spanning from -180° to 180° longitude and from -85.06° to 85.06° latitude. This CRS is not suitable for accurate local analysis in the locations of the selected villages, so single project files had to be created per each different area, using different *ad hoc* CRS. A list of the CRS which have been used can be seen in Table 4.2. Both the project and each of the shapefile layers must be set to the correct CRS in order for the measurements to be consistent and accurate. It is important to notice that, while doing the geometrical analysis, the "On The Fly" (OTF) dynamic CRS transformation option of the software must be toggled on, in order to properly take into account the deformation of the globe in the given CRS.

Once the villages have been identified and the proper CRS set, the next step was to create shapefiles containing information about the villages. Two type of shapefile layers were created:

- Point shapefiles, containing point features representing households of the villages
- Polygon shapefiles, containing the external perimeter of the villages

Given the limited number of samples, this procedure was done manually. For larger samples, automatised ways of feature recognition could be implemented to directly translate from raster to shapefile, but the accuracy of this process is highly dependent on the quality of the rasterised image. As briefly mentioned before, geometrical features in shapefiles are stored with their geographical location in the given CRS and they can be given an unlimited number of attributes, in form of a number, a text string, or a date, using a tool called "field calculator". All the attributes are then easily retrievable for visualisation into the attribute table of each layer. Concerning the point shapefiles, each geometrical point (each household) feature was given the following attributes to begin with: x-coordinate and y-coordinate (calculated automatically using the commands `$x` and `$y` in the field calculator), number of the village and country. Concerning the polygon shapefile, each geometrical feature (each village) was given the following attributes: village number, country, area of the village (calculated with `$area`), perimeter of the village (`$perimeter`).

After having created the shapefiles of the village sample and having assigned these first simple attributes, it was possible to start the real analysis of the gathered information, using the geographical correlation, based on coordinate proximity, between the separate shapefiles. It was possible for example to count the number of point features contained in each of the polygon features and assign it as an attribute to each village, using the command `intersecting_geom_count('Houses.shp')`, obtaining then the number of households of each village and deriving from that, with simple calculations, the house density in [hh/km^2].

Once the two shapefiles and the mentioned attributes had been created, the following step of the analysis was to create different micro-grid layouts to connect the households between each other.

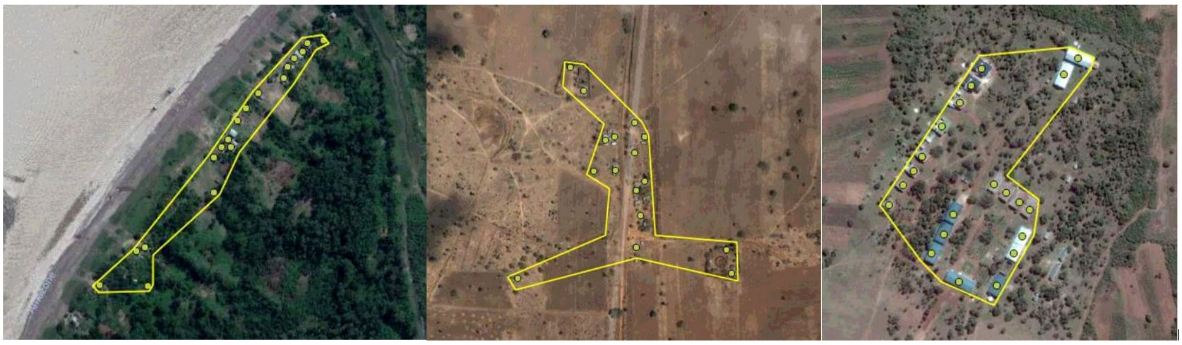


Figure 4.5: Samples of villages with identified households and external perimeter. From left to right: villages 10, 15, 26.

4.3. Micro-grid topologies creation

As previously described in Chapter 2 and Chapter 3, five different kind of micro-grid topologies were identified and chosen to be compared in this study:

- Spider diagram
- Ring topology
- Bus topology
- Radial topology
- Minimum Spanning Tree (MST)

Before moving on directly to the creation of these topologies for each village, one additional feature had to be identified: the centre of the village, which is the point having the minimum average distance from the households. The (x,y) coordinates of this central point were found using the following formulas:

$$x_c = \frac{\sum_{k=1}^{hh} x_k}{hh} \tag{4.1}$$

$$y_c = \frac{\sum_{k=1}^{hh} y_k}{hh} \tag{4.2}$$

in which hh is the number of households of the given village, x_c and y_c are the coordinates of the centre of the village and x_k and y_k are the coordinates of each of the households inside a given village. It must be noted that in this case the same relevance (weight) was assigned to each household, assuming homogeneous starting energy demand and energy production. It is anyway possible to multiply each coordinate by a relative weight if different profiles have to be taken into account.

These calculations can be done automatically by QGIS internal tools, using `Vector -> Analysis Tools -> Mean coordinates`, which requires as an input a set of points grouped by means of a common attribute, which can be for example the village number that we previously defined. If each household has a weight, identified as an attribute field, the tool can take it as an additional input for the calculations. The output of this tool is a new point shapefile (`MeanXY.shp`) containing the central points of each village, having already as an attribute the desired village number. This is going to be needed for some of the desired topologies. It was then possible to relate each of the point features to its reference village centre point using the plug-in "refFunctions" and the commands `geomnearest(MeanXY.shp,X)` and `geomnearest(MeanXY.shp,Y)`.

Spider Diagram

Starting from the Spider Diagram, which is a kind of topology in which every single household is connected directly to a central point, in this case the just identified centre of the village, a tool from the QGIS repository called "RT QSpider" can be used. This tool creates a line shapefile layer containing multiple line geometries, taking as inputs starting and ending (x,y) coordinates from a defined input layer. In this case, the starting coordinates are the single household coordinates, and the end coordinates are the central point coordinates which had just been calculated. It is possible then to join these lines together to form a single poly-line geometry, representing the final Spider Diagram for each village, using `Vector -> Geometry Tools -> Singleparts to Multiparts`.

Ring Topology

Moving on to the Ring topology, it consists of a single line, connecting the households one by one forming a closed loop. Noticeable characteristics of this kind of topology is that every node in the network has a degree equal to 2, meaning that it is directly connected to exactly two other nodes, and that, if any of the edges of the network is removed, the network is still a connected graph (see definition in Chapter 3). Following this criteria, households are connected to each other with single lines, which can then be joined together with the same procedure used for the Spider Diagram.

Bus Topology

Concerning the Bus topology, the design is slightly more complicated. This topology consists of one (or more) straight connection lines (bus lines) running through the village, to which the households are directly connected via a (ideally) perpendicular line. To implement that in QGIS, the bus(es) has to be identified and drawn manually. Then, using a short Python script in the console, it is possible to identify, for each household, the closest point on the bus line and connect, with an additional, perpendicular line, the household to the bus line.

Radial Topology

The Radial topology uses the household closest to the centre of the village (x_c, y_c) as a central starting point for the micro-grid. If there is not any household sufficiently central, then the centre of the village itself is used as origin of the network. This duality implies that some of the villages will have a radial network with a number of nodes equal to the number of households, while other villages will have one extra hub node as origin which does not correspond to an actual household, exactly as in the Spider Diagram. From this origin, the network expands radially towards the most peripheral households. Once each radial branch is created, the following households are either connected to the closest branch, or a new branch is created connecting them directly to the origin, depending on the distance. For bigger villages, sub-branches can also be added.

Minimum Spanning Tree

Finally, the Minimum Spanning Tree layout, as defined in Chapter 3, consists of the shortest possible combination of edges that leads to a completely connected graph. The identification of this particular spanning tree is relatively easy in the QGIS environment using a plug-in available in the QGIS repository: "ReconstructLine". This useful plug-in accepts as input a set of points, which have to belong to the same shapefile layer, and gives as output a set of single lines connecting one point to the other in a pre-defined line shapefile layer, so that the resulting overall network is the Minimum Spanning Tree for that specific set of points, using as working principle the Prim's algorithm [55]. As per the other topologies, the single lines can be merged into a poly-line representing the whole network using the already mentioned process.

After this process of network creation, each of the identified villages should have five different solutions to connect its households with each other, without any household being left disconnected. An example of the graphical visualisation of the resulting networks is shown in Fig 4.6. The geometrical length of each of the

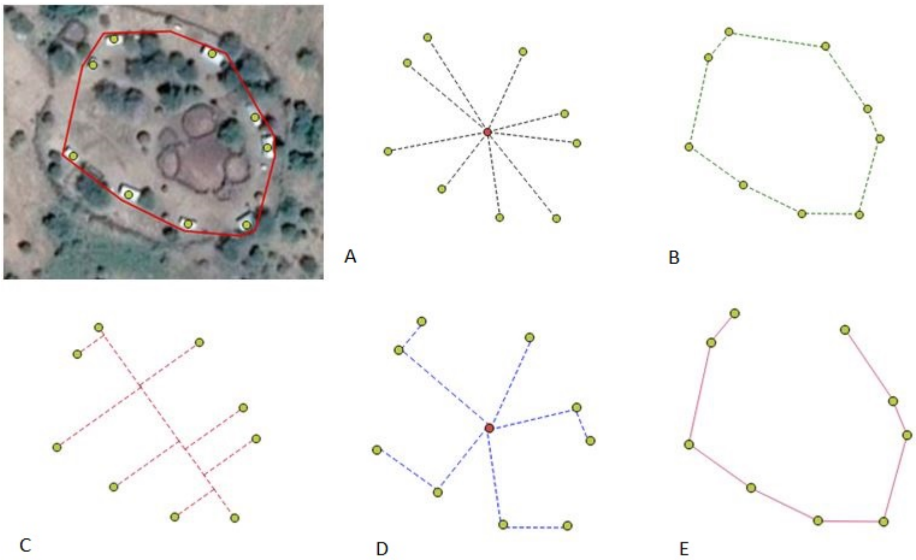


Figure 4.6: Map view and resulting network topologies for village 23. A: Spider Diagram. B: Ring. C: Bus. D: Radial. E: MST.

created network can be easily derived in the attribute field calculator of the line shapefile, using the command `$length`. Spatially correlating each of the network to the original village, it is then possible to collect in one single shapefile, namely the original polygon layer in which all the villages were stored, all the information gathered so far: village number, country, area, perimeter, number of houses, house density and now length of each of the identified topology. It is then trivial to calculate for each village the average needed length of the network, in [m/hh], directly in the attribute table. A summary of all these parameters and characteristics, for the considered sample of villages, is shown extensively in Table 4.3.

Table 4.3: Attribute table of the Villages shapefile, containing gathered, measured and calculated data on the 42 sample villages.

Village Number	Country	Area [m^2]	Perimeter [m]	Houses [hh]	Density hh/km^2]	MST [m]	Spider [m]	Ring [m]	Radial [m]	Bus [m]	MST [m/hh]	Spider [m/hh]	Ring [m/hh]	Radial [m/hh]	Bus [m/hh]
1	Colombia	11043,0	678,0	14	1267,8	362,5	878,7	564,4	368,0	418,6	25,9	62,8	40,3	26,3	29,9
2	Iceland	11823,1	618,0	6	507,5	304,0	410,5	468,0	309,8	348,3	50,7	68,4	78,0	51,6	58,1
3	Iceland	85486,2	1679,5	8	93,6	820,7	1720,3	1413,4	892,2	863,9	102,6	215,0	176,7	111,5	108,0
4	Peru	15914,7	747,7	16	1005,4	444,6	1176,4	692,1	456,5	546,3	27,8	73,5	43,3	28,5	34,1
5	Peru	79484,6	1270,4	40	503,2	1305,8	4403,3	1594,0	1456,1	1962,1	32,6	110,1	39,9	36,4	49,1
6	Colombia	6406,2	598,8	10	1561,0	283,9	549,5	506,0	283,9	300,5	28,4	54,9	50,6	28,4	30,1
7	Colombia	11407,5	521,0	8	701,3	300,3	520,4	437,9	324,8	319,4	37,5	65,0	54,7	40,6	39,9
8	Colombia	2851,0	225,7	5	1753,8	113,9	141,0	164,1	106,9	118,3	22,8	28,2	32,8	21,4	23,7
9	Colombia	30192,6	1059,5	12	397,4	644,7	1328,0	934,6	717,4	761,7	53,7	110,7	77,9	59,8	63,5
11	Brazil	11440,3	645,8	9	786,7	309,8	723,1	524,8	326,7	416,7	34,4	80,3	58,3	36,3	46,3
10	Brazil	30608,2	1361,9	18	588,1	727,4	2806,2	1285,6	728,8	901,5	40,4	155,9	71,4	40,5	50,1
12	Colombia	7870,6	496,4	13	1651,7	281,5	546,7	414,3	296,4	398,7	21,7	42,1	31,9	22,8	30,7
13	Tanzania	24714,9	1099,5	29	1173,4	869,4	2362,8	1154,5	943,1	1116,5	30,0	81,5	39,8	32,5	38,5
14	Madagascar	13455,5	520,3	19	1412,1	424,7	995,2	525,9	461,5	707,9	22,4	52,4	27,7	24,3	37,3
15	Tanzania	34537,6	1391,8	16	463,3	832,9	1611,2	1250,6	861,2	923,1	52,1	100,7	78,2	53,8	57,7
16	Tanzania	18130,1	947,4	10	551,6	489,2	1242,5	805,1	522,4	547,5	48,9	124,2	80,5	52,2	54,7
17	Tanzania	17974,1	604,9	8	445,1	365,9	601,6	523,6	365,9	423,7	45,7	75,2	65,5	45,7	53,0
18	Tanzania	52032,9	1475,2	15	288,3	863,8	2501,0	1408,4	921,4	969,4	57,6	166,7	93,9	61,4	64,6
19	Tanzania	7555,5	556,2	14	1853,0	363,7	896,5	515,2	379,9	421,5	26,0	64,0	36,8	27,1	30,1
20	Tanzania	15026,5	636,9	11	732,0	357,6	771,1	566,9	370,9	521,1	32,5	70,1	51,5	33,7	47,4
21	Tanzania	8643,9	466,5	8	925,5	270,2	453,7	387,3	270,2	280,7	33,8	56,7	48,4	33,8	35,1
22	Tanzania	41539,1	1051,6	17	409,3	842,9	1650,8	1091,1	890,0	1039,0	49,6	97,1	64,2	52,4	61,1
23	Tanzania	6608,0	301,0	9	1362,0	229,5	393,7	274,7	288,3	312,3	25,5	43,7	30,5	32,0	34,7
24	Tanzania	14606,7	645,3	14	958,5	426,8	1075,0	652,6	455,5	487,4	30,5	76,8	46,6	32,5	34,8
25	Nepal	45111,1	1680,3	38	842,4	1145,0	6159,5	1688,5	1160,8	1415,8	30,1	162,1	44,4	30,5	37,3
26	Tanzania	25229,2	713,2	21	832,4	544,1	1553,4	687,1	572,8	586,6	25,9	74,0	32,7	27,3	27,9
27	Cambodia	145324,5	2385,7	36	247,7	1979,4	7683,6	2458,3	2083,2	2250,4	55,0	213,4	68,3	57,9	62,5
28	Tanzania	46461,8	1234,6	20	430,5	936,2	2520,0	1267,4	959,3	1167,3	46,8	126,0	63,4	48,0	58,4
29	Cambodia	16860,5	784,0	23	1364,1	503,8	1897,3	693,8	510,5	790,2	21,9	82,5	30,2	22,2	34,4
30	India	57058,7	1442,2	29	508,2	957,1	3961,4	1339,0	999,9	1352,6	33,0	136,6	46,2	34,5	46,6
31	India	11606,2	589,9	24	2067,9	454,2	1074,7	606,3	471,5	637,8	18,9	44,8	25,3	19,6	26,6
32	India	11628,8	629,4	27	2321,8	461,0	1749,8	561,9	469,5	587,0	17,1	64,8	20,8	17,4	21,7
33	India	7201,9	575,4	26	3610,2	400,7	1050,3	547,4	411,0	487,1	15,4	40,4	21,1	15,8	18,7
34	India	6877,8	499,9	12	1744,7	287,5	563,6	392,6	287,5	328,8	24,0	47,0	32,7	24,0	27,4
35	India	5130,6	378,7	20	3898,2	284,2	653,6	384,5	293,9	334,3	14,2	32,7	19,2	14,7	16,7
36	Bangladesh	3189,8	310,4	16	5016,0	204,5	437,7	230,2	212,6	234,0	12,8	27,4	14,4	13,3	14,6
37	Bangladesh	6808,9	370,4	20	2937,3	326,4	633,8	418,7	335,5	440,6	16,3	31,7	20,9	16,8	22,0
38	Bangladesh	3312,7	385,9	10	3018,7	173,6	367,6	251,4	173,6	206,3	17,4	36,8	25,1	17,4	20,6
39	Bangladesh	6076,7	417,5	24	3949,5	310,1	835,3	425,2	329,3	372,6	12,9	34,8	17,7	13,7	15,5
40	India	42668,7	1157,7	29	679,7	990,0	3553,5	1290,5	1034,3	1595,5	34,1	122,5	44,5	35,7	55,0
41	Cameroon	31263,5	1026,1	31	991,6	849,1	3381,9	1016,3	878,8	1225,0	27,4	109,1	32,8	28,3	39,5
42	Gabon	4527,3	290,1	9	1988,0	198,6	289,6	237,6	198,9	207,3	22,1	32,2	26,4	22,1	23,0

4.4. Validation of methodology and comparison with Steiner tree

Before moving on to the comparison of the different topologies and to the discussion of the results, it is necessary to perform a check on the methodology used for the design of the network topologies. To this purpose, an algorithm developed in Maple by Dr. Ir. Petra Heijnen, from the faculty of Technology Policy and Management of TU Delft, was used [70]. The algorithm, which is an heuristic based on Gilbert-Melzak method [71], aims to find a minimal cost layout for networks connecting multiple sinks and sources, having as known information the energy production and consumption of each node, starting from specifically-chosen networks [72]. The inputs needed for the algorithm are the x and y (relative) coordinates of the nodes, the values for expected supply or demand of each node, the number of sources and sinks, and a capacity-cost-exponent β. Different initial tree can be chosen for running the model: Minimum Spanning Tree, Star Network, Minimal Cost Spanning Tree and a Random Spanning Tree. The cost function which is minimised is the following [72]:

$$C(G) = \sum_{e \in E(G)} l_e q_e^{\beta} \tag{4.3}$$

in which $E(G)$ is the set of all the edges e composing the graph G, l is the length of each edge and q is its capacity. For the desired purpose, since considerations about the demand and supply of each node are going to be implemented only at a later stage, only the length had to be taken into consideration. In order to do so, the coefficient β was set to a value of 0, so that the only cost component of the target function is the length of the micro-grid network, the Minimal Spanning Tree automatically becomes also the Minimal Cost Spanning Tree and the output can directly be compared with the measured length on QGIS.

In the process of validating the methodology, since three of the five created topologies are only semi-automated and are not comparable with the networks identified by Dr. Ir. Heijnen's algorithm, the validation of accuracy of the measurements can be performed only on the Spider Diagram and the Minimum Spanning Tree. Since the Spider Diagram, as it will be shown later in this chapter, is in most of the cases the most costly option and it is more suitable for non-distributed generation, it was chosen to perform the validation on MST exclusively. In order to do so, MST was chosen as initial tree for the algorithm. The algorithm gave then three useful outputs: the cost(length) of the initial MST tree, which was used as comparison to the MST tree length found via QGIS, the cost of the minimal cost Steiner tree found, and its graphical representation.

Treating the first of these outputs, the cost (length) of the starting MST identified by the algorithm, using the (x,y) coordinates given as input, was directly compared with the MST length measured on QGIS, identified following the process described in Section 4.3 for all the villages of the sample. The result of the comparison is shown in the graph in Figure 4.7. The average difference between the two values is around 0.035%, with a maximal error value of 0.087% among all the forty-two villages examined. Such a small of error can also be considered to be due to the maximum amount of significant figures which can be inserted as inputs in the algorithm compared to the significant figures used by default in QGIS calculations. It is then possible to state

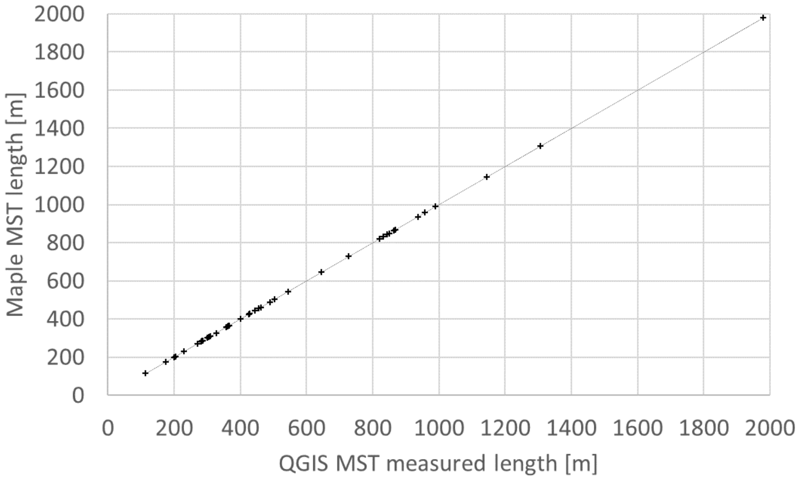


Figure 4.7: Comparison of the length of MST measured using Heijnen's algorithm and using GIS methodology.

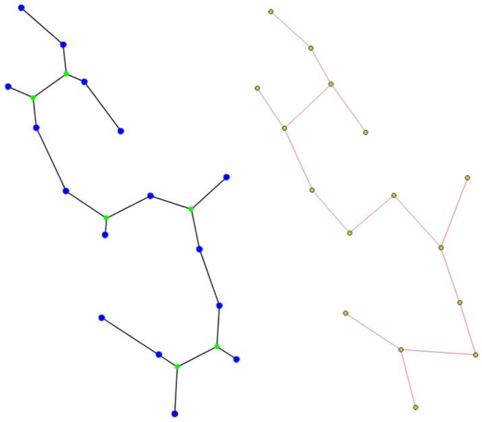


Figure 4.8: Comparison of Steiner Tree (left) and Minimum Spanning Tree (right) for sample village 36.

that the measurements done through QGIS are reliable and it is possible to proceed further with the analysis.

An additional consideration that may be discussed alongside with the outputs from the algorithm is the possibility of including in the topology comparison also the Steiner Tree optimised by the algorithm. As explained in Chapter 3, the difference between a Steiner tree and a Spanning tree is that Steiner tree can include points (nodes) which are not part of the starting set of nodes to be connected. Potentially any point in the plane can be added to the graph. The addition of these "passive" Steiner nodes can lead to cost savings in terms of network length in this particular application. Figure 4.8 shows an example of a Steiner Tree and a Minimum Spanning Tree for the same set of points (households). The six green points are the Steiner points which have been added by the algorithm to minimise the costs. It must be noted that, even if the addition of Steiner points lead to savings in terms of length (5.7%, or 11 m, in the example from the Figure), the additional costs of having extra passive connection nodes (eg. installation of additional poles) may offset the savings. In Figure 4.9 the statistic results of the comparison between the two options are shown, considering only the savings in length. The only case in which the Steiner tree is consistently better is village 8, which is one of the villages in which the origin of the radial topology is not an actual household. This means that actually the radial topology is better than the MST topology, as it will be discussed in the following section. In fact, comparing the Steiner tree with the Radial topology instead of the MST, the savings decrease from 7.9% to 1.9%. On average over the forty-two cases, the addition of each Steiner point saves around 2.5 m of network length. It was then decided to proceed with the analysis of the five previously identified topologies, without adding an extra unnecessary layer of complexity, a part from including the Steiner tree in simple length comparisons

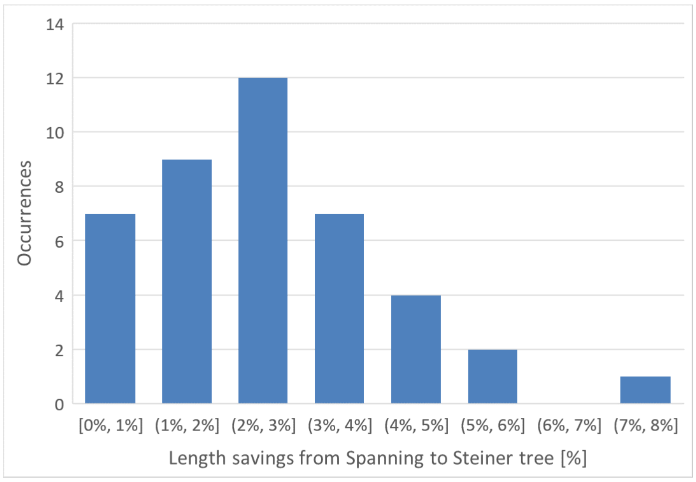


Figure 4.9: Distribution of occurrences of length savings from Minimal Spanning to Steiner tree for the 42 selected villages.

as a benchmark for the other topologies.

4.5. Results and discussion

In Table 4.3, the forty-two analysed villages and their characteristics are listed. That is the basis for the comparison of different topologies in terms of costs and for the identification of trends and patterns useful for an optimal design of micro-grids. In this section, these results are going to be analysed and discussed.

4.5.1. Topologies comparison

A first consideration to be made concerns the correlation which is present, as expected, between the household density of the village and the average network length needed per household to completely connect them in a micro-grid. Trivially, the more clustered the houses are, the smaller the distance between them. This trend is confirmed by the measured length of all the five kinds of topology (six counting the Steiner tree), as shown in Figure 4.10. In the figure, the single results for each village and each topology are shown, together with topology-specific trend lines. The average lengths range from a minimum value of 12 m/hh, for Steiner topology in really clustered villages, up to a maximum of around 215 m/hh for a spider topology in really sparse villages. It is clear, then, how the feasibility of micro-grid is highly affected by the household density and the household distribution of each village.

At a first glance, it is also possible to notice how the different identified topologies compare to each other and to the benchmark represented by the Steiner Tree, found using the algorithm described in Section 4.4. In terms of general trend, the topologies described can be ranked in the following order, from the best to the worst in terms of average network length:

Minimum Spanning Tree < Radial Topology < Bus Topology < Ring Topology < Spider Diagram

It must be noted, though, that this ranking and the ratios between different topologies is not the same for each of the identified villages. In fact, it is dependent on various different factors, such as the number of houses, the shape of the village and, most importantly, the geometrical distribution of the households inside the village, which make one topology more convenient for specific cases. Some interesting examples

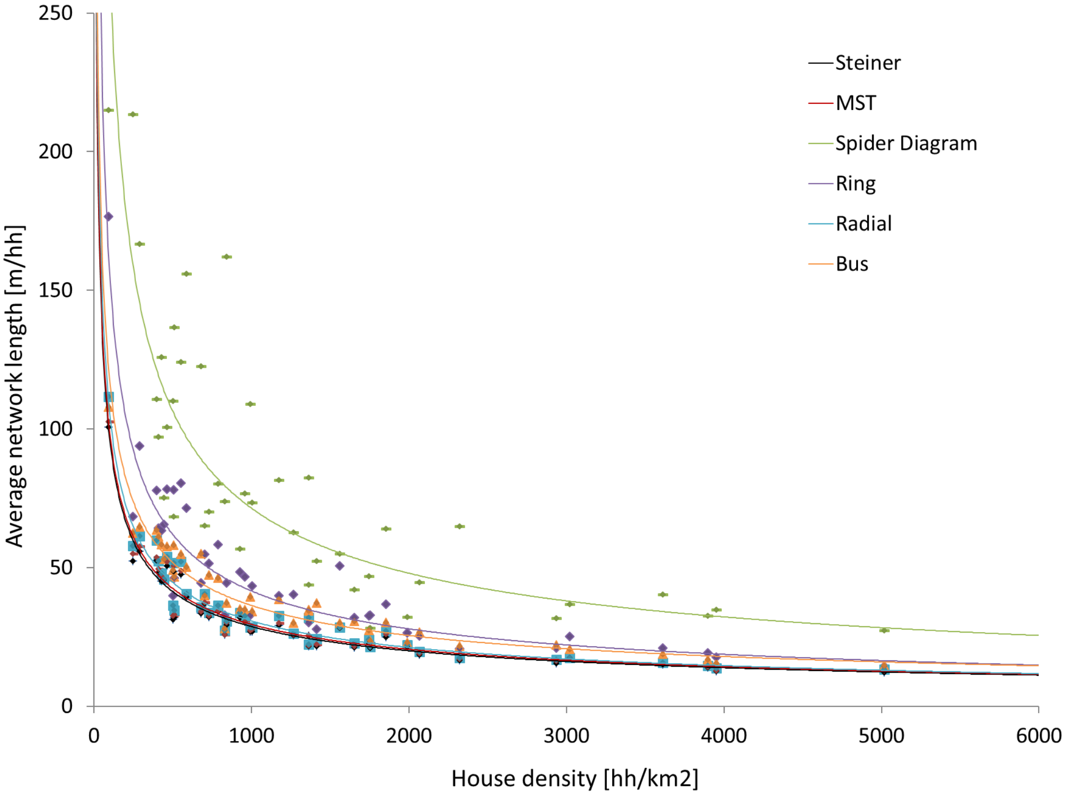


Figure 4.10: Average length for the different topologies as a function of the house density.

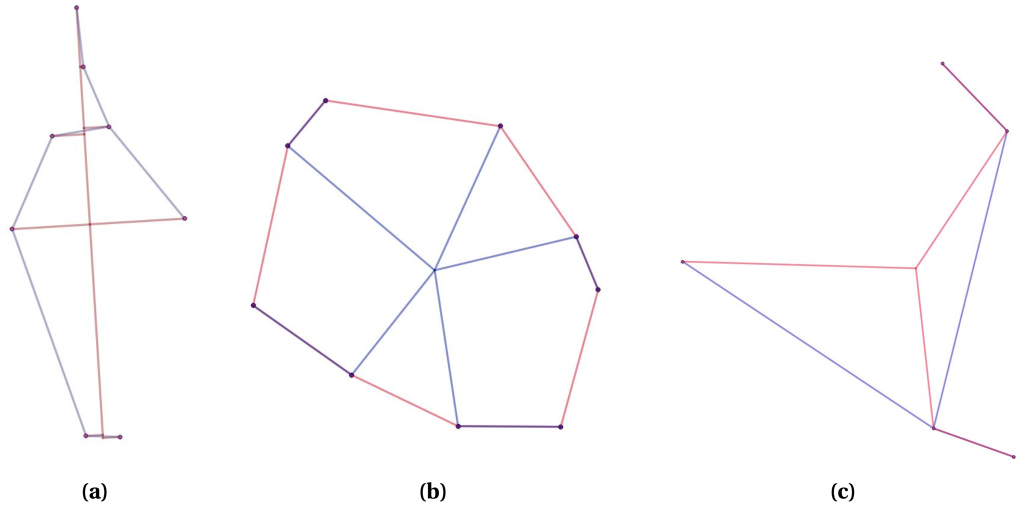


Figure 4.11: Specific cases of network comparison: (a) Village 7, red = Bus, blue = Radial, (b) Village 23, red = Ring, blue = Radial, (c) Village 8, red = radial, blue = MST.

of specific villages are the following, shown also in Figure 4.11:

- Village 7, which has a clear, long shape, it is symmetrically distributed on the sides of a straight road and it does not have any household near the cent er of the village, is more suitable for a Bus topology (108 m) than for a Radial topology (111 m). See Figure 4.11a
- Village 23, which has a very peculiar circular layout for household distribution. In this small village, a Ring topology (30.5 m) is slightly more convenient than a Radial topology with the origin in the cent er of the village (32 m). See Figure 4.11b
- Village 8, which is the smallest village of the sample. Its shape and the scattered distribution of the houses make in this case a Radial topology with the origin in the centre of the village (21 m) even more convenient than a Minimum Spanning Tree (23 m). See Figure 4.11c

These are just some of the cases that could be analysed singularly, but it already gives a hint of how the optimal layout topology is highly case-specific.

Trying to analyse the general trend for different topologies, a direct correlation doesn't appear to be found between number of houses and most convenient topology. In fact, as can be seen in Figure 4.12, the ratio between the length of each network and the Steiner tree benchmark does not seem to be related to the number of households, except for the Spider Diagram case, shown separately in the left graph. For this particular topology, as the number of houses in the village increases, the comparison with the benchmark becomes worse and worse, reaching networks up to more than 5 times longer than the Steiner Tree. This could have been expected and can be easily explained geometrically by the fact that in the Spider Topology each household gets connected directly to the centre of the village, creating a number of "branches" which are inefficiently close to each other. This realisation, together with the fact that, in 40 out of 42 cases in the identified sample, the Spider Diagram is the least convenient topology, makes it possible to state that this particular layout should not be taken into consideration for micro-grid planning. The only case in which it would still make sense to consider a Spider topology would be a fully centralised generator with no distributed generation systems, which is not the situation considered in this thesis work.

Leaving out the Spider Diagram option, the other topologies lead to extra network length needed up to around 80% compared to the Steiner benchmark. As already mentioned, on the contrary with respect to the Spider Diagram, there is not any dependence on the household number for the extra length needed in comparison with the Steiner benchmark. It can be noticed both in Figure 4.10 and in Figure 4.12 how the Minimum Spanning Tree is deviating only slightly from the benchmark, 2.7 % on average and in some cases it is actually identical to the Steiner tree. This could be expected, given the fact that both the Steiner Tree and the Minimum Spanning Tree are generated with mathematical algorithms, while the four other topologies have superimposed criteria and rationales which leads to constraints that cause deviations from the geometrically

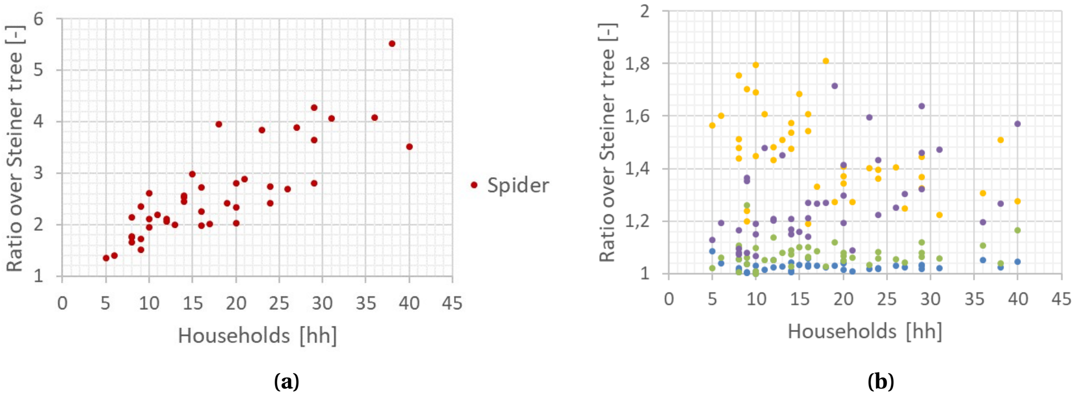


Figure 4.12: Comparison of (a) Spider diagram length and (b) other network topologies length with the Steiner tree benchmark

optimal solutions. Nonetheless, in the case of the Radial topology, the results do not show a big difference from the first two mentioned topologies. In fact, the average relative additional length needed for the Radial topology in comparison with the Steiner benchmark is around 7% and only in 8 cases out of the 42 examined it exceeds 10%. In case of more strict superimposed constraints for the construction of the topology, such as in the case of Bus and Ring topologies, the deviation from the optimal solution become larger. In the case of the Bus topology, the average increase in length is approximately 27%, while for the Ring topology the average is even higher, around 46%. Nonetheless, there are particular cases, such as the ones described earlier, in which also these structured superimposed topologies are comparable to the algorithmically found layouts, with length increases lower than 10%. This usually happens only in case of peculiar shapes and households distribution.

4.5.2. Conclusions

After this analysis of the five different identified network topologies, it is possible to draw a first set of conclusions regarding micro-grid planning, keeping in mind that this first step only considers geometrical distances and network length, without making any consideration on technical aspects, supply and demand of each household, regulations and peculiar geographical features:

- There is a clear correlation between household density and average network length needed, for all of the topologies considered. Denser settlements will have to face lower investments per household in terms of micro-grid length.
- Keeping the Steiner Tree as optimal benchmark, layouts such as MST and Radial topology do not show major increases in network length, settling around 2.7% and 7% average, respectively.
- More structured superimposed topologies have significantly higher average increase in network length, due to higher constraints. Nonetheless, Ring and Bus topologies can still be considered valuable options in case of specific village shapes.
- Spider Diagram is in most cases not suitable for micro-grids with distributed generation, being substantially longer than any other considered topology, especially for villages with higher number of households.
- It is of foremost importance to analyse the geometrical layout, the shape, the households distribution and other characteristics of each single case study in order to identify the most convenient option. The methodology explained here uses available tools and a newly proposed rationale to help in decision making during the micro-grid layout process.

5

Layout optimisation

In the previous chapter, a purely geometrical analysis has been performed, comparing different kinds of micro-grid layout in terms of needed length. The networks created with the methodology described in Chapter 4 are going to be utilised in the following step as starting topologies for a dual-objective optimisation, which is going to be performed using an algorithm written in Python.

First of all, the rationale and the methodology of the algorithm are going to be described in detail, in order to understand the steps and the assumptions made in the process. After this description, the algorithm is going to be applied to some of the villages of the identified samples and the results of the optimisation are going to be shown and discussed. As a conclusion, since the algorithm can be subject to future improvements and modifications, some potential flaws and ideas to further develop it are going to be suggested.

5.1. Rationale and Assumptions

In Chapter 4, the analysis and comparison of different solutions for micro-grid was carried out, using as sole parameter of comparison the network length (per household). That is the reason why only topologies which are spanning trees have been selected as initial networks, meaning that, given two random households (nodes) A and B in the network, there is only one possible path to go from A and B. In other words, there are no unnecessary redundant edges in the network. The only exception in the set of starting topologies is represented by the Ring topology, which has one extra edge closing the ring loop. This fact assured a minimisation of the network length, within the constraints imposed by each given topology definition. Length is directly correlated to costs of the network, so minimising the length is the most direct way of minimising the costs of interconnection.

Nonetheless, cost saving is not the only important aspect while designing a DC micro-grid network. Additional network lines, which lead by definition to increased costs, can bring benefits for other aspects of the micro-grid operation. Three of those aspects are going to be discussed in the following paragraphs: redundancy, congestion avoidance and losses reduction.

5.1.1. Redundancy

Redundancy can be interpreted as safety or reliability of the network in case of damages to single lines or in case of manumission. Each spanning tree, if one of the edges is cut, becomes separated into two different sub-networks, isolated from each other. This means that potential exchanges of energy from one sub-network to the other are not possible anymore. Furthermore, if in one of the sub-networks the energy production and energy reserve is not enough to satisfy the demand of the same sub-network at a given moment, the risk of energy outages and the loss of load probability dramatically increases. With additional edges, increasing the redundancy of the network, additional alternative paths can be found for at least some sub-sections of the network. The most peripheral nodes of the network would always suffer higher risks of isolation then more central, better connected households, but, as more edges are added to the network, the overall risk of isolation of sub-networks is reduced. An example of that can be seen in Figure 5.2.

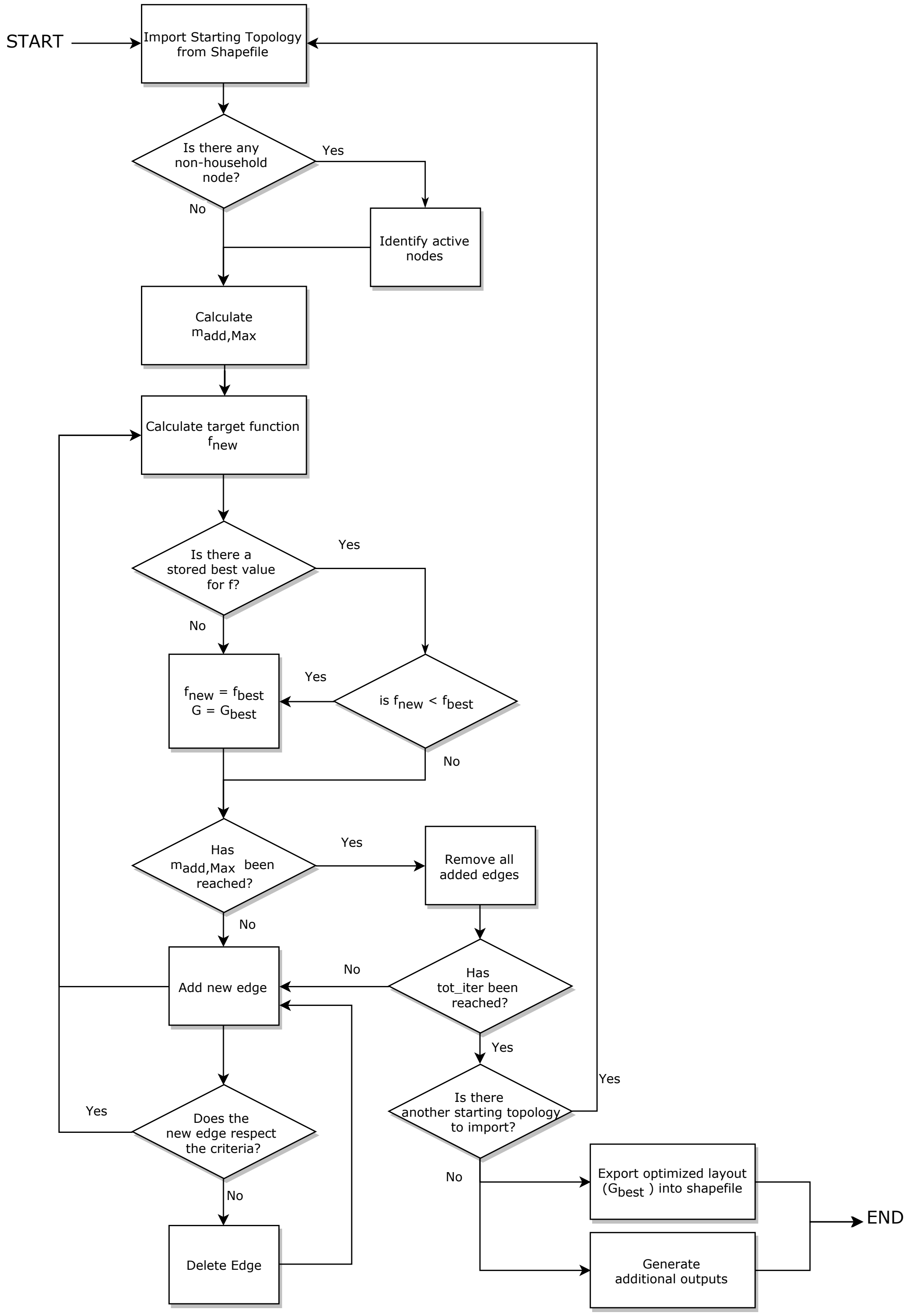


Figure 5.1: Flowchart of the proposed algorithm. A starting network topology is imported in .shp format to be processed by the algorithm. Passive nodes are identified, if present. Edge addition iterations are performed until tot_iter is reached, re-initialising the edge addition process from the starting topology every $m_{\text{add,Max}}$ edges. Each new edge is checked under a set of criteria before being added. A target function is calculated at every iteration and the best overall graph is stored and saved as optimised output. More than one starting topologies can be used at once, in a consecutive way.

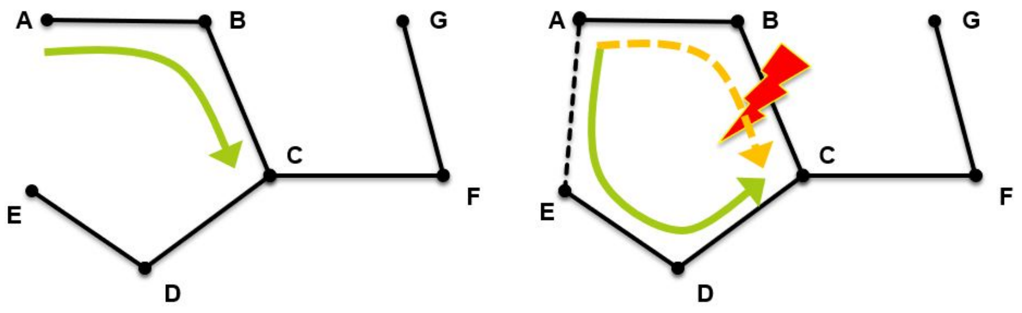


Figure 5.2: Example of increased redundancy by edge addition.

5.1.2. Congestion avoidance

Congestion in one or multiple lines is detrimental for the performance of micro-grid networks. If the power flow exceeds the technical limits of a cable, the consequence can be its failure and hence the complete disconnection of the line. Moreover, the power losses in a cable are proportional to the square of the current flowing in it, as per the basic formula:

$$P_{\text{loss}} = I^2 \cdot R \tag{5.1}$$

in which I is the current in [A] and R is the cable resistance in [Ω]. High power losses are also directly reflected in overheating of the cable. Hence, if the current flow is too high, there is a double negative effect, in terms of increased losses and in terms of risk of overheating. As in the case of redundancy, adding new edges may offer a second alternative path for power flow from a given household A and a given household B, reducing risks of congestion. An example can be seen in Figure 5.3.

5.1.3. Losses Reduction

Additionally to the reduction of losses due to the avoided congestion, there is another reason for which adding edges leads to further reduction in losses. To understand that, it is important to define, as already mentioned in Chapter 3, the concept of average shortest path length. The shortest path length in a network for a given pair of nodes is the length of the shortest route to connect those two nodes to each other following the edges of the graph. The average shortest path length is simply the average of this distance over all the possible pairs of nodes in the network. That parameter is directly linked to power losses calculated in 5.1, because the cable resistance R can be calculated with the following formula:

$$R = \rho \cdot \frac{l}{A} \tag{5.2}$$

in which ρ is the resistivity of the cable measured in [Ω/m], l is the length of the cable in [m] and A is the cross-section area of the cable in [m^2]. It is clear then how the power lost in a power exchange between two

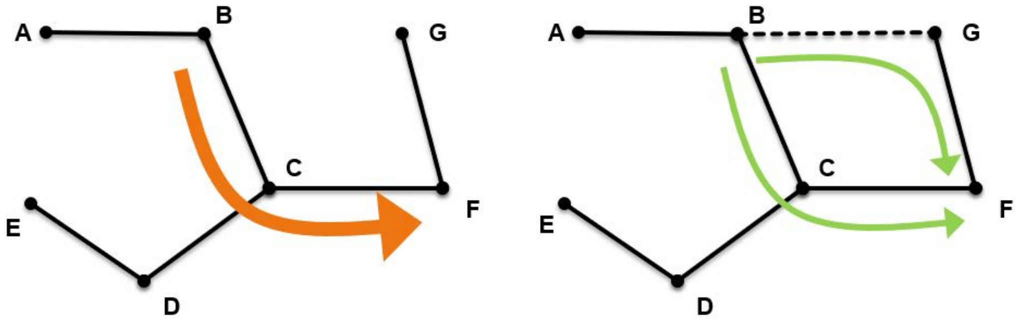


Figure 5.3: Example of congestion avoidance by edge addition.

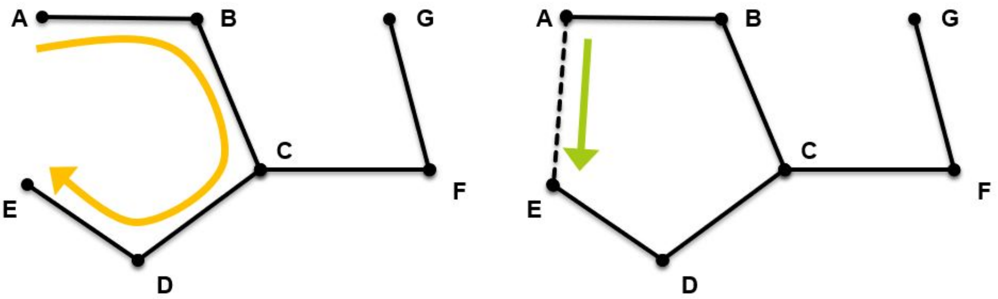


Figure 5.4: Example of decreased shortest path length by edge addition.

households will be higher if the shortest path length separating them is bigger. The addition of new edges can create new shortest paths between pair of households, reducing the average shortest path length of the network and potentially the losses in power exchanges. An example of that can be seen in Figure 5.4.

5.1.4. Dual-objective optimisation

These benefits, potentially brought by the addition of redundant edges, lead to the need of an optimisation which has not the sole aim of minimising cable lengths and cable costs, but also tries to find a trade-off between costs reduction and the just mentioned benefits. While it is complicated to quantify these benefits in order to compare them with cable length costs, especially in this first phase of network planning, without any specifications on households demand, production and power exchanges, it can still be assumed that they are proportional to each other. With that purpose, the topologies designed in Chapter 4 are going to be used as starting networks for addition of new edges, hence forming so called "meshed" micro-grids, which can be simply defined as micro-grids in which one or more loops can be identified.

In order to compare these benefits, it was decided to use the average shortest path length of the network as a second parameter to optimise, taking it as a general indicator of the level of redundancy, congestion avoidance and losses reduction, as opposed to average network length, which is taken as an indicator of direct cable costs. To perform this dual-objective layout optimisation, an algorithm in Python language was written, following the flowchart shown in Figure 5.1. The algorithm takes as inputs one or more of the starting topologies identified and drawn using the methodology explained in Chapter 4, and it processes it adding new edges for a user-defined number of iterations, trying to minimise a target function which takes into account both average length and average shortest path length. This target function is a tentative to represent with one single value the trade off between costs and redundancy and it is written in order to be highly costumisable by the user according to the kind of prioritisation and relevance given to each of the single parameters. The proposed function in the algorithm is the following:

$$f = \sqrt{wf \cdot (\bar{l})^2 + (\overline{spl})^2} \tag{5.3}$$

in which $\bar{l}$ is the average network length in [m], $\overline{spl}$ is the average shortest path length of the network in [m] and wf is a user-defined weight factor. This weight factor is tuned to prioritise one or the other parameter which is considered more important for each single application.

In this thesis work different weight factors were used during the analysis to check the sensitivity of the result to weight factor variations. For real case studies, a deeper analysis is needed to identify a precise case-specific weight factor to be applied, depending on a combination of technical, economical and socio-political aspects. Examples of such aspects can be availability of materials, expected energy demand, planned size of SHS, risk of tampering, energy security requirements, climate conditions and many more. In this work weight factors in the range from 0.1 to 2 are suggested as reasonable values for the optimisation algorithm. The way 5.3 is written, higher weight factors will prioritise cable costs over the mentioned benefits, so minimisation of the average network length will be more relevant than the minimisation of average shortest path length, hence resulting in less edges added. Lower weight factors, on the contrary, will prioritise redundancy, congestion avoidance and losses reduction, minimising the average shortest path length as much as possible, tending hence to add more edges to the initial topology.

5.1.5. Algorithm Methodology

In order to perform this algorithm, a Python library called NetworkX can be used, containing some useful already implemented functions to calculate parameters which are needed for network analysis and optimisation. For example, with the function `average_shortest_path_length()` it is possible to directly calculate one of the two parameters of the defined target function, namely $\overline{spl}$, and with the functions `read_shp()` and `write_shp()` it is possible to work with files in the ESRI shapefile format.

The first step of the algorithm is indeed importing the ESRI shapefiles as graphs into the algorithm to process them as mathematical elements. The default NetworkX function import line shapefile as directed graphs (DiGraphs), in which a direction is assigned to each edge, which can lead to mistakes into the shortest path length calculation because of negative values. It is then important to convert directed graphs into undirected graphs (Graph) before starting to process them. The weight of each edge for calculation purposes is assigned taking the "length" attribute from the attribute table of the shapefile, as it had been previously created and assigned during the methodology described in Chapter 4. To set a first benchmark, the target function f is calculated on the starting network and stored as temporary best value.

For post-processing analysis purposes, the starting number of edges in each imported network is stored, simply using the internal NetworkX function `number_of_edges`. After having calculated it, the number of edges which have to be added to reach a complete graph are calculated using the formula:

$$m_{\text{add,Max}} = \frac{(n \cdot (n-1))}{2} - m_{\text{start}} \tag{5.4}$$

in which n is the number of active nodes (households) in the network and m_{start} is the number of edges in the starting topology. It must be noted that in some of the starting topologies there are passive nodes which are not corresponding to households, hence it is needed to find criteria to identify which nodes are actually to be considered as active households. The criteria change according to the topology, for example using the degree of connectivity of a node is a good way to distinguish the origin of a spider diagram (degree > 1) from the peripheral nodes which are real households (degree = 1). It is important to calculate this $m_{\text{add,Max}}$ because once the complete graph has been reached, the iteration process has to be re-initiated from the starting topology.

It must be noted that the number of possible edges rapidly increases with the number of households in the network. Using simple combinatorics, it is possible to derive the general formula for number of possible edges to connect a given set of nodes, without the constraint of the initial edges present in Equation 5.4:

$$m_{\text{Max}} = \frac{n \cdot (n-1)}{2} \tag{5.5}$$

As can be seen in the graph in Figure 5.5, already for the restricted sample of villages which have been selected in this study (maximum 40 households per village), the number of possible edges almost reaches the

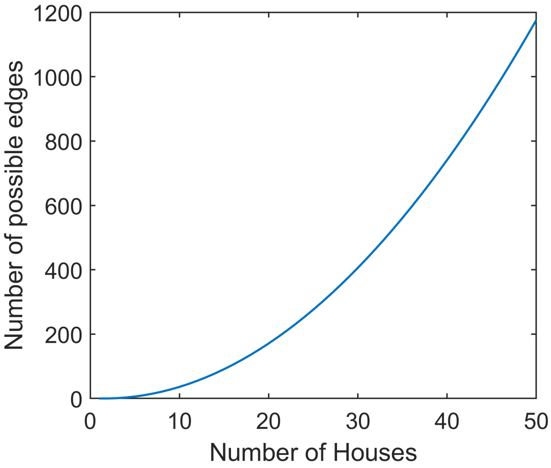


Figure 5.5: Number of possible edges as a function of number of nodes

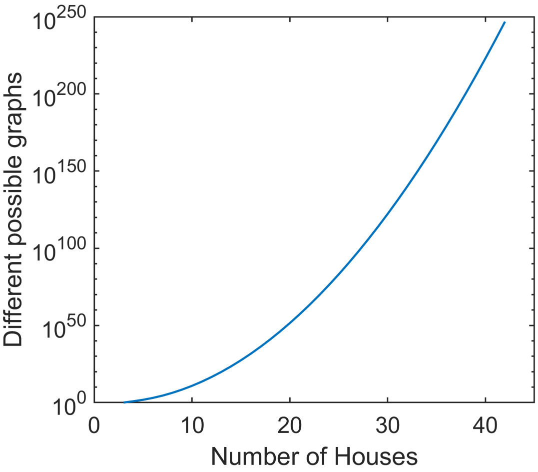


Figure 5.6: Number of possible graphs as a function of number of nodes

order of thousands. The figures increase even more rapidly if we calculate the number of potential graphs containing a number of edges from 1 to m_{Max} for a given number of nodes n, which can simply be calculated as:

$$G_{\text{Max}} = 2^n - 1 \tag{5.6}$$

As can be seen in the graph in Figure 5.6, the order of magnitude of the potential combinations of edges is far too high to consider an optimisation algorithm which scans through all of them to find the absolute optimal solution. Such an algorithm would require computational times far beyond the reach and purpose of our study, so a simplified heuristic to reach a quasi-optimal solution was designed.

The initial approach chosen for this heuristic optimisation algorithm, in terms of choice of edges to add to the starting network, is a randomised process. At each iteration, one single random edge is added to the graph and the target function is calculated again for the resulting graph and compared with the stored function value. If the new value is lower, then the stored best function value is updated and the new Graph is stored, substituting the former. Since NetworkX, as already mentioned, does not recognise active nodes from passive nodes, for topologies such as Bus and Spider, it is needed to write an extra piece of code to force NetworkX to iterate and average the calculation of the shortest path length only between pair of active nodes.

If no additional constraints are set, the edge addition is iterated until the complete graph for the given set of nodes is reached, after the addition of $m_{\text{add,Max}}$ edges. Once the complete graph is reached, all the added edges are removed and the iteration is reinitialised from the starting topology, as the order of edge addition is also influencing the optimisation. This reiteration can be repeated as many times as desired by the user, which has to give as an input the number of total iterations (one iteration means one edge added) that are to be performed on each starting topology (input parameter `tot_iter`). Since the computational complexity and the number of edge combinations increases drastically with increasing number of houses and since it is reasonably expected to find the optimised graph with a number of added edges consistently lower than $m_{\text{add,Max}}$ as calculated in 5.4, the algorithm can be forced to reinitialise the edge addition process before reaching the complete graph, simply modifying the formula (or the value directly) to reduce $m_{\text{add,Max}}$. The combination of the choice of the values for `tot_iter` and $m_{\text{add,Max}}$ has a strong influence on the optimality of the results of the algorithm.

Since in Graph objects it is possible to have multiple edges connecting the same pair of nodes, it is needed to verify for each iteration that the newly added edge is not already present in the Graph. Furthermore, on the basis of the geographical (x, y) coordinates of the nodes, the physical length (weight) of the new edge has also to be calculated and added as an attribute to the edge for further calculations. At this step of the algorithm, further modifications could be made and additional criteria for edge selection could be applied, as will be discussed later on in Section 5.3. Once the new edge has been checked, added and its length calculated, the value of the target function is calculated for the newly created network and compared to the temporary stored best value. If the new target function value is lower than the stored best value, this value gets updated. It is important to notice that the algorithm stores not only the target function value, but also all the information about the (temporarily) optimised network, in form of NetworkX Graph element, so it can later be retrieved and displayed at the end of the algorithm. It would be too time (and memory) demanding to store the newly created network for every iteration, so for the non-optimal graphs only the coordinate couples ($\overline{spl}$, $\bar{l}$) are stored to be finally shown in a scatter graph.

Table 5.1: Summary of algorithm Inputs and Outputs.

Inputs		Outputs	
	Format		Format
Starting networks	.shp	optimised network	.shp
Weight factor	float	Nodes	.mat
Desired Iterations	integer	Edge Lengths	.mat
		Incidence Matrix	.mat
		Scatter Graph	.pdf
		Layout graphical representation	.pdf
		Data points	.xlsx
		Log	.txt

This process can be repeated for each starting topology separately, hence obtaining the (quasi-)optimised solution for each given imposed starting layout, or in an integrated way, obtaining a single final solution which already compares internally the different starting topologies. In the second case, the algorithm also stores information on which was the starting layout from which the optimal solution was found. At the end of the algorithm, the NetworkX Graph object identified as optimal network solution is reconverted into an ESRI shapefile for further analysis, processing or import in other software.

Overall, the algorithm needs as input the shapefile(s) of the starting layouts, the number of iterations desired for each of the starting topology and the weight factor for the target function. The primary output of the algorithm is then the new shapefile of the optimised graph, together with a spreadsheet of the average length and average shortest path length values for all the graphs evaluated during the iterations, a scatter graph with those values, a log report with information about the used parameters for post-processing analysis purposes and a graphical representation of the network layout. An additional and useful output which was included in the algorithm is a set of .mat files containing information on the nodes and edges of the optimised network, which is going to be used as input for the DC power flow analysis described in Chapter 6. Inputs and outputs can be seen in a summarised way in Table 5.1.

In the next section, results of the application of the just described algorithm for some of the networks identified in Chapter 4 are going to be shown, described and discussed,

5.2. Results and Discussion

The algorithm just described in the previous section of this report has been applied to a limited number of villages chosen among the forty-two composing the full GIS analysis sample. The results of this micro-grids layout-optimisation are going to be described and discussed in this section, together with some considerations on the methodology itself.

For most of the examined villages, the algorithm was run separately for each different starting topology, obtaining topology-specific optimised meshed grids as well as overall best solutions. For other villages, especially with higher number of houses, the algorithm was run starting only from some of the potential starting topologies, which were considered more suitable, for time reasons.

For reasons of simplicity and clarity, the first results which are going to be discussed are for the smallest village of the sample, counting five households. it is a good example to see how the algorithm works and how the different weight factors used influence the outcome of the optimisation. For this small village, the number of possible edges and possible graphs are relatively small, as can be calculated using equations 5.5 and 5.6 respectively. The couples of value $(\overline{spl}, \overline{l})$ for all the potential network layouts are shown in the scatter graph in Figure 5.7, resulting as an output of the algorithm using all five different starting topologies. The bigger points with different colours on the top-left corner of the graph are representing the starting networks.

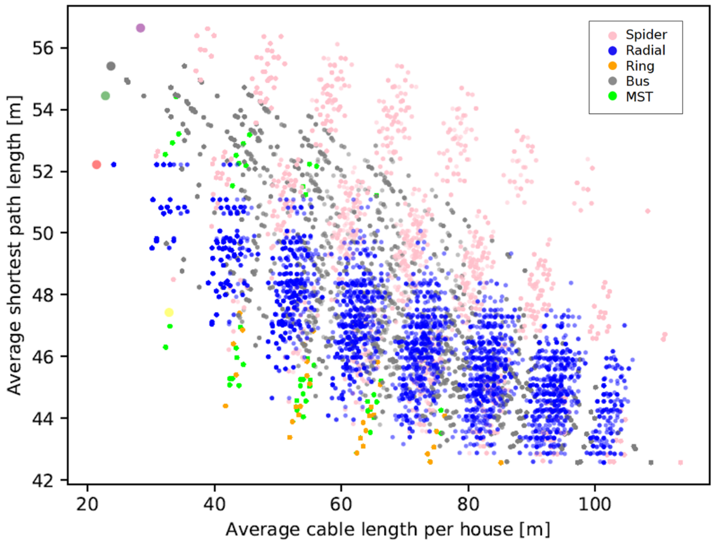


Figure 5.7: Average network length and average shortest path length of network layouts generated and tested by the algorithm for village 8. Different colours represent different starting topologies.

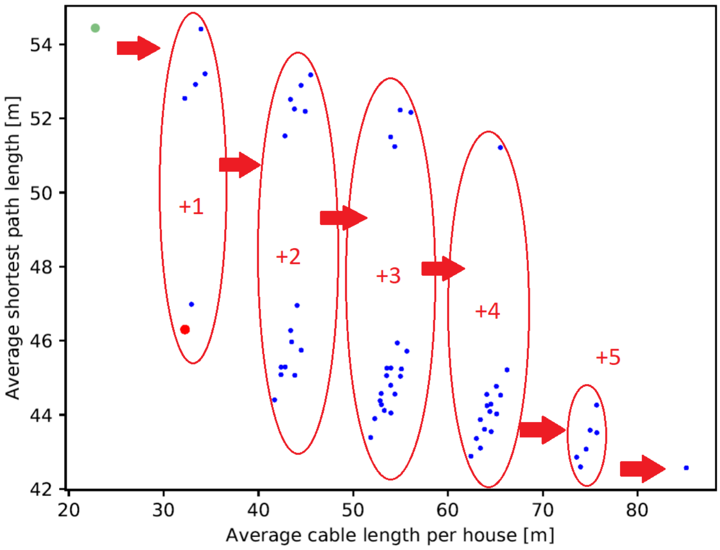


Figure 5.8: Average network length and average shortest path length of network layouts for village 8, starting from MST topology. The green dot indicates the starting network, while the red dot indicates the optimal solution, found after one edge addition while using a weight factor of 0.3.

A few things can be noticed from this graph. First of all, it appears clear and confirmed that as more edges are added, thus increasing the average cable length, the average shortest path length decreases, as expected, until it reaches a minimum threshold. This threshold is reached when the network has become a complete graph, having then the maximum possible average cable length and the minimum average shortest path length, since every pair of nodes is directly connected via an edge. It is important to notice that the complete form of the network changes according to the starting topology, for example because of non-household nodes and bus lines, hence changing the maximum average cable length, but not the minimum average shortest path length, which is reached with spanning trees. This explains why in Figure 5.7 the complete graph starting from the spider topology (the farthest on the right pink dot) is not the same as the complete graph starting

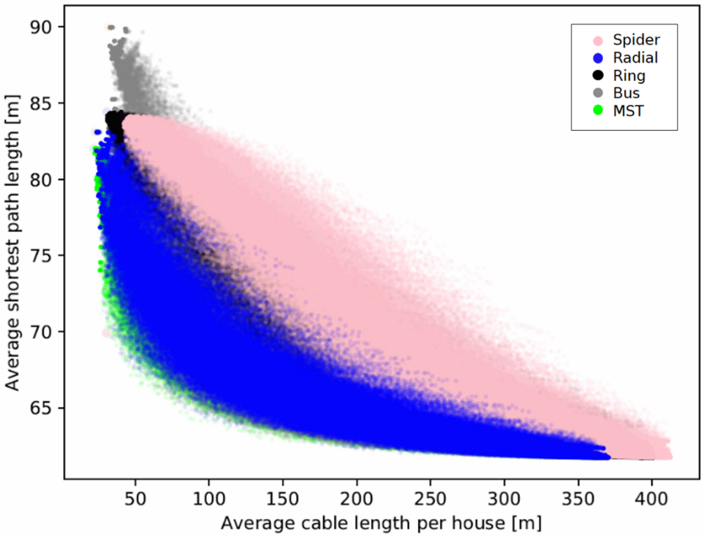


Figure 5.9: Average network length and average shortest path length of network layouts for village 12, which has 13 households. Different colours mean different starting topologies.

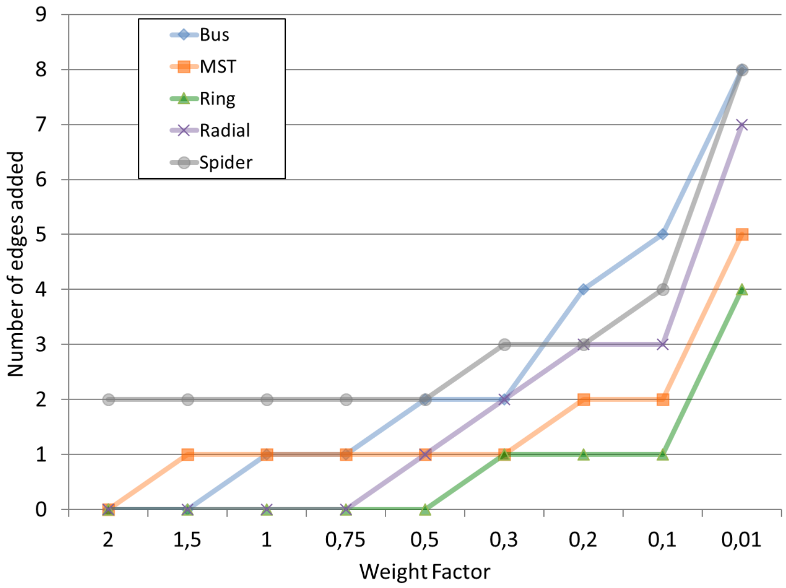


Figure 5.10: Edges added to reach the optimised layout for village 8 with different weight factors and starting topologies.

from the bus topology (the farthest on the right grey dot). It is also interesting to see how the spider diagram, which was the least convenient topology in terms of pure cable length (see Figure 4.10), is also performing worse than the other layouts in terms of average shortest path length, reinforcing the statement made in Chapter 4 about its inadequacy to this kind of application.

Analysing the scatter graph resulting from running the algorithm with only the Minimum Spanning Tree as starting topology, shown in Figure 5.8 it becomes even clearer how the addition of each edge leads to a gradual change in the parameters, from the starting graph to the complete graph. It is also possible to graphically see six "stages" of solutions, each representing one consecutive edge addition.

As the number of nodes increases, the complexity increases as well, as already mentioned previously, leading to scatter graphs which are less clear and distinct compared to the ones just shown, but the trends can still be identified. As an example, in Figure 5.9 the scatter graph resulting from running the algorithm for village 12, which has thirteen households, is shown.

5.2.1. Weight factor and starting topology influence

Going back to village number 8, the algorithm was performed using different weight factors in the range between 0.1 and 2. As expected this leads to different optimised layouts, especially in terms of number of added edges. The lower was the weight factor, the lower was the importance of network length and the more edges were added to the initial network in order to reach the optimised topology, as shown in Figure 5.10. It can be noticed that for such a small village, it can happen that no edges are added at all for some topologies at high weight factors, because the benefit would be considered too small to justify the addition of a single edge, hence the optimised layout ends up being the starting topology itself. That is for example the case for a

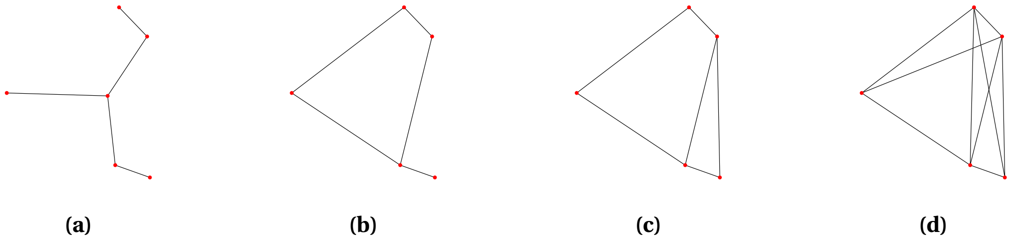
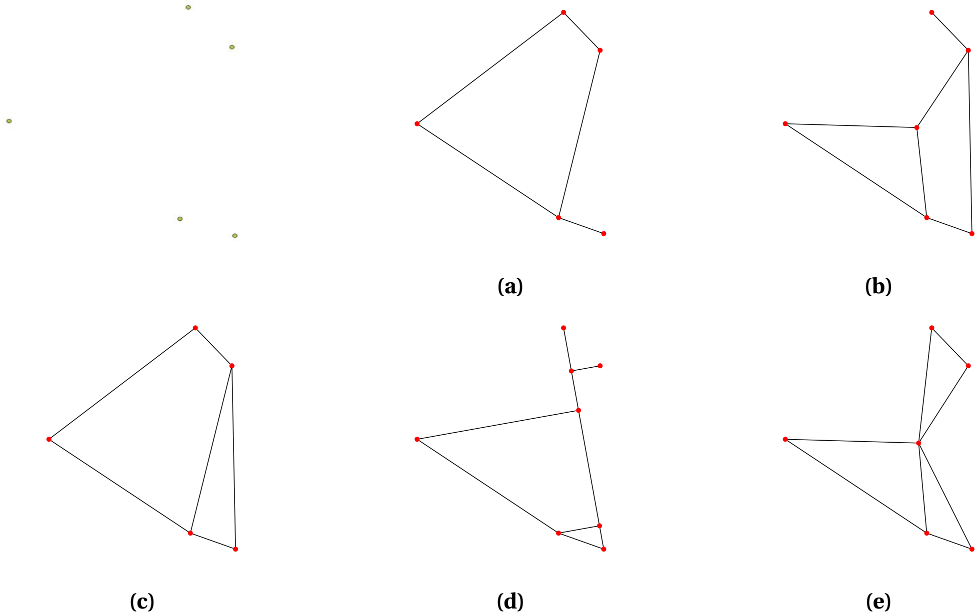


Figure 5.11: Global optimised layouts for village 8 with different weight factors: (a) 2 (b) 0.3 (c) 0.1 (d) 0.01.



(a) (b) (c) (d) (e)

Figure 5.12: optimised layouts for village 8 with weight factor 0.3 and different starting topologies: (a) MST (b) Radial (c) Ring (d) Bus (e) Spider.

weight factor of 2, in which 4 out of 5 starting topologies do not require any edge addition.

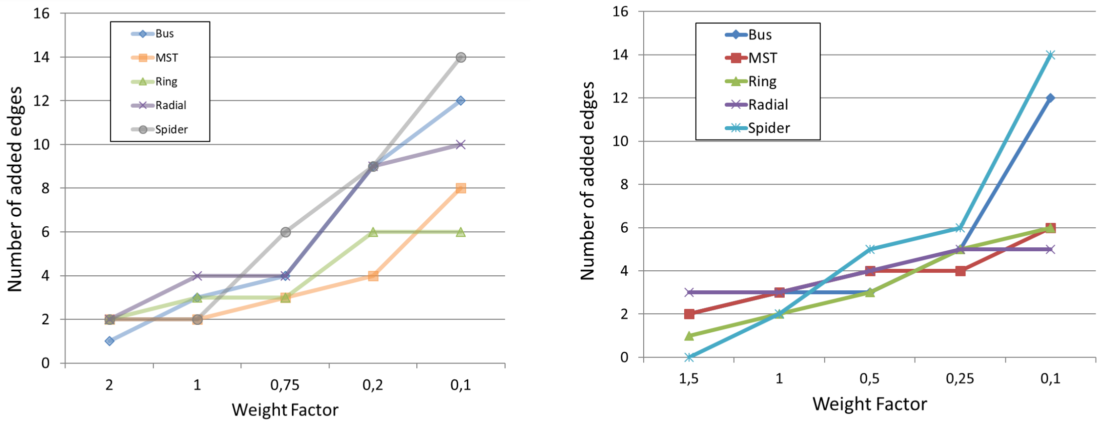
Apart from the single starting topologies, the influence of the weight factor can also be seen in the final optimised layouts. In Figure 5.11, the final layout for village 8 is shown for different weight factors. The weight factor 0.01 is willingly exaggerated on the side of deprioritisation of cable length costs and it was applied to stress edge addition to its excess. It can be noticed how the layout changes with the weight factor, but also how different weight factors can lead to different starting topologies to be optimal. As an example, while for a weight factor of 2 the radial topology was a better starting topology, for a weight factor of 0.3 an MST was optimal. It must be also noted that in some cases the optimal layout can also be reached by starting with different topologies. For example, the optimal layouts for weight factors 0.1 and 0.01 could be reached both starting from a Ring topology and from a MST topology.

In Figure 5.12, the optimised layouts of village 8 are shown for different starting topologies. It was chosen to show the topology-specific results for a weight factor of 0.3 because all the results had at least one edge added to the starting network. The partial optimised layouts were then compared to each other to find the global optimal, which in this particular case resulted to be the one obtained from a MST starting topology.

The algorithm was also applied using all the five starting topologies to other villages, of different shapes and sizes, using different weight factors. The list of the performed full optimisations is summarised in Table 5.2, with the different weight factors used and with records of which starting topology resulted in the final optimised layout. Consistently with the findings from Chapter 4, the Minimum Spanning Tree topology resulted to be the best one in most of the cases, while the Radial topology resulted to be the optimal starting network only in a few cases. The Ring topology could be also included into this list, but it resulted to be optimal only

Table 5.2: Optimisations performed with full set of starting topologies. The starting topology which lead to the optimised layout is also recorded per each weight factor.

Village	Houses	Weight factors	Optimal starting topology
6	10	0.2, 1	MST, MST
7	8	0.2, 1	MST, MST
8	5	0.01, 0.1, 0.2, 0.3, 0.5, 0.75, 1, 1.5, 2	Ring/MST, Ring/MST, Ring/MST, MST, MST , MST, Radial, Radial, Radial
9	12	0.1, 0.2, 0.5, 0.75, 1, 2	MST, MST, MST, MST, MST
12	13	0.2, 1	MST, MST
42	9	0.1, 0.25, 0.5, 1, 1.5	Radial, MST, Radial, MST, MST



(a) **(b)**

Figure 5.13: Edges added to reach the optimised layout for villages 9 (a) and 42 (b) with different weight factors and starting topologies.

in those cases in which the optimal final layout could be possibly derived both from a starting Ring and a starting MST. To give two more examples of how the weight factor influences the number of added edges, the same kind of results shown for village 8 in Figure 5.10 are also reported similarly for villages 9 and 42 in Figure 5.13. Looking at this set of results on the number of added edges, it can be noticed that the optimised layout is always reached with the addition of a number of edges consistently lower than $m_{\text{add,Max}}$ (not considering the exaggerated case of weight factor 0.01).

On the basis of these results, it was decided to perform additional optimisations starting only from the MST and Radial topologies, in order to give just a few more examples of optimised layouts, without using all the different starting layouts and weight factors as done in the first set of trials. Moreover, the formula for calculating $m_{\text{add,Max}}$ in the algorithm was modified, dividing it by a factor which was tuned for each particular village and topology, to tune it down and to reinitialise the iteration before reaching the complete graph. Both of these modifications lead to a decreased number of total iterations needed to reach the optimal solution. The outputs of these additional trials of optimisation are not going to be described in detail, but some examples of optimised layouts are going to be shown at the end of this Chapter in Figure 5.14. Some of those examples were used for other purposes, external to this thesis project, to assess energy sharing mechanisms in micro-grids, in another thesis project inside the DCE&S group [25].

5.2.2. Conclusions

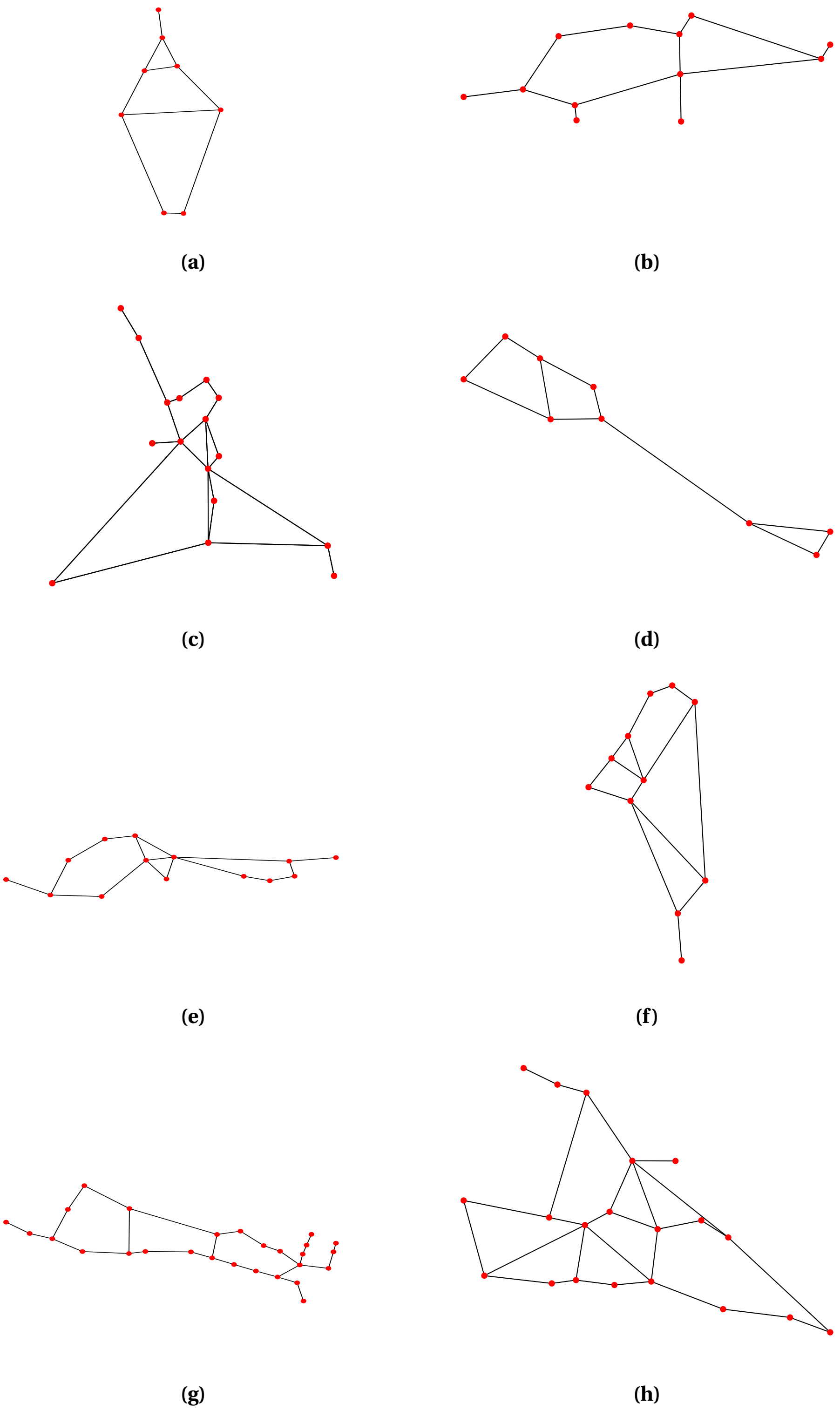
After this analysis of the application of the first version of the algorithm, some preliminary conclusions can be drawn, before moving on to the DC power flow analysis to be performed on some of the identified optimised layouts.

- Lower weight factors in the target function (lower priority for cable costs savings) lead to the an increased edges addition to reach the optimal layout, independently from the starting topology.
- As per the single-objective cable length optimisation, the Minimum Spanning Tree and the Radial topology seem to be the most adequate for micro-grid planning, independently from the weight factor used.
- The edges that are more likely to be added to reach an optimal layout are usually short edges closing new loops into the network or cutting through already existing loops.
- The algorithm still needs improvements in order to be applied more efficiently, especially to villages with higher number of households. Some of these improvements are going to be suggested in the next section.

5.3. Future Improvements

Even though it is already giving quite satisfactory results, especially when applied to networks with a limited number of nodes, this algorithm is just a first, very generalised version of what could be an universally applicable methodology for multi-objective layout optimisation for rural DC micro-grids. While performing this thesis work, many potential ideas for future improvements were thought of, but not implemented because they were out of the purpose of the project or because of time issues. A short list of some of the most interesting changes that could be applied are going to be listed here.

- Rearranging the target function to optimise, trying to better quantify the benefits of meshed topologies in terms of redundancy, congestion avoidance and reduction in power losses. It would be ideal to translate the benefit into real money savings, in order to be able to compare them directly with cable costs.
- Quantifying the weight factor to reflect real case-specific conditions in order to more consciously prioritise the right parameters in micro-grid planning.
- Including considerations on the demand and production profiles of each household (node), in order to calibrate the layout based not only on the geographical distribution of households, but also on the energy profiles (eg. more redundancy around households with higher production).
- Adding ulterior criteria for checking adequateness of edges to add to the initial layout, rather than just the uniqueness check which is used in this first version. Examples could be adding only edges below a determined length, or adding only edges which are adjacent to nodes with a low degree of connectivity.
- Substituting the random selection of edges with a prioritisation or with a more complex rationale, for example taking inspiration from Prim or Kruskal algorithms.
- Reinitisialing the addition of edges already before reaching $m_{\text{add,Max}}$ added edges if the target function is already deviating too much from the best recorded value, in order to avoid waste of computational time adding ulterior edges to a far-from-optimal layout.
- From the coding point of view, there are still a lot of improvements to do in terms of computational speed, code cleanness and clarity.
- Mechanisms of edge removal could be implemented into the algorithm, for example to remove a newly added edge if the negative impact on the function value is considered too high, without having to reinitialise the process or to wait to reach $m_{\text{add,Max}}$.



(a) **(b)** **(c)** **(d)** **(e)** **(f)** **(g)** **(h)**

Figure 5.14: Examples of optimised meshed layouts: (a) Village 7, wf = 0.2 (b) Village 9, wf = 0.8 (c) Village 15, wf = 1.2 (d) Village 16, wf = 0.8 (e) Village 19, wf = 0.8 (f) Village 20, wf = 0.3 (g) Village 32, wf = 1.2 (h) Village 35, wf = 0.8.

6
DC Power Flow Analysis

In this Chapter, after having performed a GIS analysis on a purely geometrical length basis in Chapter4 and a qualitative layout optimisation taking into account redundancy, congestion avoidance and losses in Chapter 5, DC power flow simulations are going to be run on a sample of optimised layouts, in order to assess the electrical performance of those layouts. To do so, a DC power flow tool designed by a TU Delft Master thesis student, Dario Chaifouroosh, is going to be used [73]. The rationale and methodology of the used tool are going to be explained in this chapter, together with some modifications which were applied during this thesis work to improve the applicability of the tool to the specific purpose of this project. The DC power flow calculations are then performed on some of the resulting layouts found with the optimisation algorithm for the simplest village (village 8, with 5 households), given a set of assumptions on the production and consumption of each household. The results of the power flow simulations are going to be discussed and analysed in detail.

6.1. DC Power Flow tool

During a Master thesis project in the DCE&S group of Delft University of Technology, a tool was developed in MATLAB for finding a DC power flow solution for a given set of households, each capable of producing, consuming and eventually storing energy, connected by a set of edges [73]. The various elements of the network are modelled into this tool following electrical equations and simplifying their representation to a limited set of parameters. While the real topological layout of a micro-grid is more complicated then just a set of edges connecting punctual nodes, assumptions can be made in order to simplify the computation without sacrificing accuracy. Through topology analysis, it is possible to reduce to single punctual elements all the components internally present in a single household, transforming it into a single node that can be treat as a single entity with given parameters for computational purposes. By using this simplification, a micro-grid layout can be directly represented by a set of nodes and set of edges, which makes this tool perfect for processing the outputs of the optimisation performed in Chapter 5.

Line elements can be easily represented as resistors, hence needing only the value of cable resistance R in [Ω] as input for the tool. The resistance of each cable can be calculated with some basic information and assumptions on each single line using the formula:

$$R = \rho \cdot \frac{l}{A} \tag{6.1}$$

in which ρ is the resistivity of the cable measured in [Ω/m], l is the length of the cable in [m] and A is the cross-section area of the cable in [m^2].

Moving on to the nodes of the analysed networks, they can be modelled as source nodes, load nodes or hybrid nodes, depending on their capability of producing or consuming energy. Source nodes can be represented by characteristics of constant power, constant voltage or constant impedance, according to type of generator and the electrical configuration of the node. Into the scope of this project, photovoltaic production is the only generation option considered and it can be modelled as a constant power node [73]. To avoid confusion, note that constant power does not mean that the power output is constant in time, but that the power output is given as a fixed input and is not dependent on the configuration of the layout and other parameters. Load nodes can be modelled as constant power, constant current or constant impedance nodes. For the purpose of this thesis work, all loads in nodes were considered constant power loads and they were given as inputs through consumption profiles throughout typical daily periods. Hybrid nodes are nodes that can alternatively inject power into the grid or absorbing power from it, for example energy storage systems such as batteries, which can be modelled as droop nodes. Droop nodes can be defined by the following characteristic equation [74]:

$$V - V_{ref} = k_d \cdot I_s \tag{6.2}$$

in which V_{ref} is the reference voltage of the node, k_d is the droop coefficient of the node and I_s is the current flowing into the node. The droop behaviour can also be restructured using a Norton equivalent circuit to a combination of constant current and constant impedance, with a value for the current source given by the formula:

$$I_{ci} = \frac{V_{ref}}{Z} \tag{6.3}$$

in which Z is equal to the droop coefficient k_d.

For computational reasons it is then possible to characterise batteries using two simple parameters, the reference voltage and the droop coefficient [73].

Each node can also contain a combination of these three type of nodes, which is the case for an household equipped with a Solar Home System, which has a PV system working as a source node, electrical appliances working as a load node and a battery working as an hybrid droop node. In such a way, households can be considered as single node elements. The way to implement this efficiently in the power flow tool is going to be described in the section 6.3.

Once the set of lines and nodes has been defined and modelled, the other necessary input for the DC power flow calculation is a representation of the way those edges and nodes are connected. This is done through a so-called oriented incidence matrix, which is a matrix of dimension mXn, in which m is the number of edges, n is the number of nodes, and each row indicates the starting and ending point of each line with respectively -1 and 1 values in the matrix. A summary of all the elements of the networks and how they are reflected into the tool is shown in 6.1.

Table 6.1: Elements of a network in the DC power flow tool.

Real element	modelled element
PV system	Constant P node
Loads	Constant P node
Battery	Droop node (Reference V and Droop coefficient)
Edge	Constant R line
Layout structure	Incidence Matrix

Once the layout of the network and the parameters of each of the network elements have been defined, the tool proceeds to linearise the parameters of each node and solve the DC power flow at by using three sets of linear equations that relate voltages and currents with the known variables of the system, based on Kirchoff's and Ohm's laws [73]:

$$V_i - V_j - I_{ij} \cdot R_{ij} = 0 \tag{6.4}$$

in which V_i and V_j are the voltages in the starting and ending node of the line respectively, I_{ij} is the current flowing in the line and R_{ij} is the line resistance.

$$\sum_j I_j + I_{ci} = 0 \tag{6.5}$$

in which I_j is any current leaving or entering the node and I_{ci} is the node current for nodes modelled as constant current nodes.

$$\sum_j I_j - \frac{V_i}{Z_i} = 0 \tag{6.6}$$

In which Z_i is the node impedance for nodes modelled as constant impedance nodes.

Any node which presents a combination of different behaviours is reduced by the tool to these three basic linear equations to solve the power flow.

The tool is programmed so that, given a set of inputs, it calculates statically the DC power flow for each desired instant of time in a given period. The inputs are given at the beginning of the computation and the power flow of each time step is not dependent on the outputs of the previous time steps. The measured and display parameters for each time step are node voltages, line currents and line power losses, together with an analysis of the iterations needed to reach voltage convergence, which is not needed for the purpose of this application.

In the next section, some changes that were applied to the tool, to make it more adequate to the specific purpose of this thesis work, are going to be listed and explained.

6.2. Tool Modifications

Before performing the DC power flow analysis on the optimised layouts calculated in Chapter 5, some modifications had to be made to the code to include some aspects that were not considered or to retrieve some additional parameters from the calculations already made in the tool. Some of these changes were simply done by further processing of the outputs of the simulation, while other required a deeper and bigger effort in internally tweaking the tool and adding extra functions.

The outputs that were already defined in the original version of the tool developed by Chaifouroosh are: voltage levels per each node, current flowing in each line and power losses in each line. For the purpose of this thesis project, other interesting parameters that may be analysed are: node power input/output, efficiency of the network, battery throughput, state of charge of the battery and number of overvoltage and undervoltage occurrences.

Concerning the net power input/output of each node, all the needed technical parameters were already calculated inside the tool. It was only needed to add a piece of code per iteration, calculating the power injected (or absorbed) by each node i, using the following formula:

$$P_i = (\sum_j I_{ij}) \cdot V_i \tag{6.7}$$

which was already used internally in each time step to perform a power balance check, but without storing the values for the single time steps.

An additional output that was thought useful for this analysis and for other applications of the tool is the efficiency of the network, which indicates how much power is wasted in network losses over the power injected into the network and can be calculated using the formula:

$$\eta_{\mathrm{MG}} = 1 - \frac{\sum_i \sum_j P_{\mathrm{loss},ij}}{\sum_i P_i^+} \tag{6.8}$$

in which P_{loss} are the power losses along the cable lines, already calculated inside the tool using Equation 5.1 and P_i^+ is the power injected by each injecting node at a given time step. It is important to note that negative values (absorbing nodes) are not directly used in the calculation. It is also important to note that this efficiency only takes into account cable losses, without considering any other type of loss in the system or inside the single households. That is why it is more adequate to name η_{MG} *network layout efficiency* and its value will be higher than typical expected micro-grid efficiency.

Another fairly direct parameter to consider was the number of occurrences of overvoltages and undervoltages in the network. The sensitivity of islanded micro-grids to voltage fluctuations is a major issue, given the absence of a stabilising element such as the connection to the main grid and even more with the penetration of renewable energy systems such as the SHSs considered in this study. The number of occurrences was simply calculated counting all the values of node voltage V_i which were respectively above or below a tolerance threshold chosen to be 10% of the nominal voltage of the micro-grid.

apart from the just mentioned simple additional calculations, the major change applied to the original tool was including the presence of non-infinite energy storage systems to cope with supply and demand mismatches during the time span of the simulation. In the original tool, energy storage systems were already present and modelled as droop sources, in the same way in which they are modelled in this modified version, but no limits were given for the amount of energy which could be stored. This assumption is acceptable for the case studies and applications in Chaifouroosh's master thesis, while for the particular application to islanded micro-grid, where the buffer of the connection to the grid is missing, it is important to evaluate battery behaviour. In particular, "infinite batteries" would lead to a too high level of self consumption, reducing actual power flows in the lines and reducing the relevance of this DC power flow analysis in terms of evaluation of different layout topologies.

Unfortunately, the modelling of the battery as a droop source posed some limitations to its flexibility of operation and it was not possible, for example, to define it in such a way that the battery would have been available only to charge or discharge, so an alternative approach had to be found. First of all, a new input had to be defined in the set of parameters of each node, namely the value of the battery capacity, in [Wh]. This value is clearly not dependent on the time step, so it could be given as a single value per node. Secondly, an initial state of charge of the battery also needs to be defined, which is going to be used as input for the first time step of the computation and then automatically updated after each time step. Assumptions on the initial state of charge are going to be described in the next section. In order to update the state of charge of the battery at each time step, it is needed to first calculate the energy available in each battery at the beginning of the time step, using the formula:

$$E_{\mathrm{batt}_{i,t}} = SoC_{i,t} \cdot E_{\mathrm{max},i} \tag{6.9}$$

in which $SoC_{i,t}$ is the State of Charge of the battery of each node i (if present) at the time step t and $E_{\mathrm{max},i}$ is the battery capacity in [Wh].

It is also needed to calculate the power output/input of the battery for a given time step, which is not directly calculate by the tool but can be derived per each node using the following formula:

$$P_{\mathrm{batt}_{i,t}} = P_{\mathrm{net}_{i,t}} - P_{i,t} \tag{6.10}$$

in which $P_{i,t}$ is the power injected by the node into the grid, while $P_{\mathrm{net}_{i,t}}$ is the net power of the node without taking the battery into account. The amount of net power produced depends on the difference between power produced (if the node has a source) and power consumed (if the node has loads). With this value and knowing the duration of each time step (in this case one minute), the new value of energy level of the battery and the new state of charge can be calculated using the following formulas:

$$E_{\mathrm{batt}_{i,t+1}} = E_{\mathrm{batt}_{i,t}} + \frac{P_{\mathrm{batt}_{i,t}}}{\Delta t} \tag{6.11}$$

$$SoC_{i,t+1} = \frac{E_{\mathrm{batt}_{i,t+1}}}{E_{\mathrm{max},i}} \tag{6.12}$$

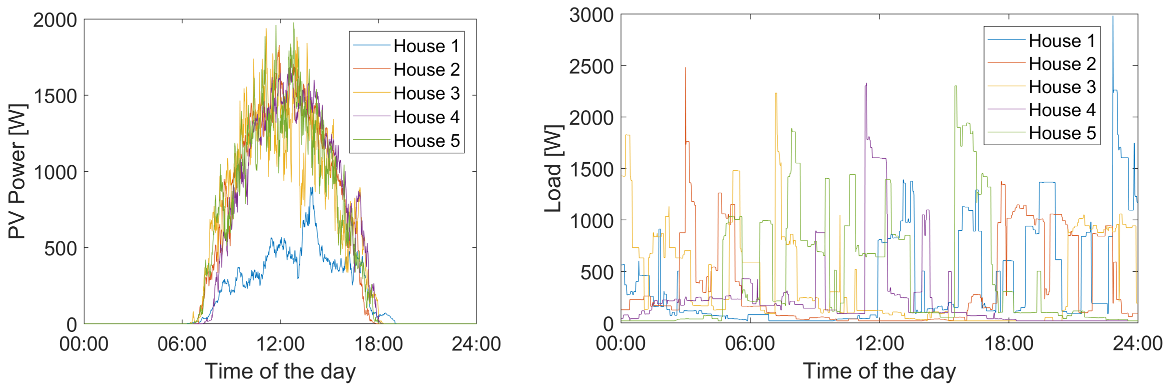
Contrarily to most of the other parameters involved in the tool, the State of Charge is directly modified by the previous time step, adding a dynamic time-dimension to this DC power flow computation.

Batteries are generally meant to be used not at their full energy capacity, but only inside a range that spans between 20 % and 80% of their total energy capacity. To comply with that constraint, the state of charge must be forced to stay within the values of 0.2 and 0.8 in every time step of the computation. Since with the current code structure it is not possible to predict *a priori* if the state of charge limits are going to be respected, the approach used was a feedback *a posteriori*. The time step is computed normally and at the end of each time step the check on the state of charge is performed. If the limits are not respected, the time step is re-initiated, disconnecting just for that single time step the battery which would have trespassed the limits, and the power flow is calculated again with the changed assumptions. In this way, all the batteries are kept within the limits, while still taking into account their state of charge and the maximum energy capacity available. It was also then possible to calculate the battery throughput of the battery, summing the inward and outward energy flows.

After having applied these modifications to the original tool, it was possible to proceed to the actual DC power flow computation for some selected optimised micro-grid layouts.

6.3. Case Study Assumptions

Before starting to describe the definition of the case studies, the technical parameters used and the assumptions made, it is important to clarify the purpose of this analysis. The main reason to perform a DC power flow computation on micro-grids layout in the context of this study is to identify how their technical operating parameters compare to each other in conditions of high energy exchange along the network. Some of the assumptions made may then be slightly exaggerated or not fully persistent to reality, with the purpose of



(a) **(b)**

Figure 6.1: Sample of (a) daily PV production profile and (b) daily load demand profile for a set of 5 households.

highlighting the difference between different layouts, while keeping all other parameters the same to have a fair comparison.

All the geometrical and geographical information were retrieved directly from the outputs of Chapter 4 and Chapter 5 in terms of coordinates, length of the edges and incidence matrix. The only bit of processing needed for the incidence matrix was transforming it from an undirected full matrix to a directed sparse matrix to be imported into the MATLAB tool. To derive the cable resistance per each edge, Equation 6.1 has been used, with typical values of 4.17 mm^2 of cross section, $0.0171\Omega \cdot mm/m^2$ and using the length of the cable stored as attribute in the ESRI shapefiles. In line with those assumptions, the maximum allowed current per cable is 10 A [75].

To model the households, three important components had to be defined in their characteristics: the PV system, the battery and the loads. It was chosen to perform the DC power flow calculation on a time period of one day (1440 minutes), in order to have a full simulation of a cycle of PV production and demand typical one day. The simulation could have been performed per longer periods, but it was chosen to keep it limited to just one day in order to have a better resolution and clearer results, avoiding the risk of dispersion of information along a period longer than needed.

To model the PV system production profile on a daily basis, data from the work done by another master thesis student of the DCE&S group, Ali Chamseddine, was used as starting point [76]. Estimated production profiles for a 20 W PV system in rural application in a location in India were scaled up by a factor which was suitable to comply with the requirements of the different Energy tiers shown in Figure 1.4, resulting in five different yearly profiles, each corresponding to one energy tier. In order to differentiate as much as possible between different households, always with the purpose of promoting energy exchange in the micro-grid, the daily profile of each household was selected randomly from one of the 365 days of the year. An example of the daily PV production profile used in the simulations is shown in Figure 6.1a.

Concerning the load profiles, inputs were taken from the work of Nishant Narayan, PhD candidate in the DCE&S group, who developed a stochastic methodology to derive typical load profiles of rural households using off-grid DC appliances [77]. This specific work was already meant to fit into the multi-tier framework proposed by ESMAP [14], so it was possible to use it directly without need of any modification, since it was already matching the categories of the 5 energy tiers. As per the production profiles, also for the demand a process of randomisation was used to diversify among different households. In addition, to further boost energy exchanges in the micro-grid, the load profiles were shifted, in order to have a better distribution during the daily period of analysis. This modification actually takes away some realism from the load profiles designed by Narayan, but it enhances the possibility of comparison between different layouts. An example of the daily load demand profile used in the simulations is shown in Figure 6.1b.

Since each node in the tool can only have a single power parameter input, the calculation of the net power produced/absorbed by each single household needs to be done before feeding the inputs to the code. This was done with a simple formula, per each time step and per each household:

$$P_{\mathrm{net}_{i,t}} = P_{\mathrm{PV}_{i,t}} - P_{\mathrm{L}_{i,t}} \tag{6.13}$$

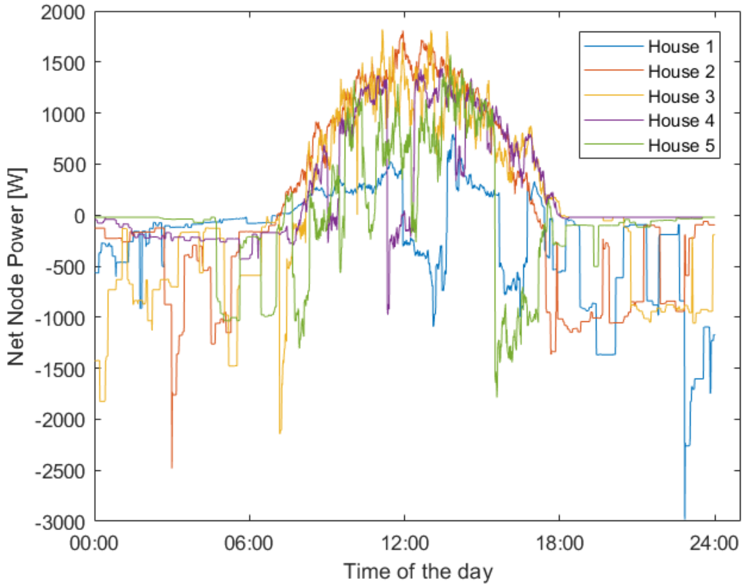


Figure 6.2: Sample of daily net power profile for a set of 5 households.

in which $P_{\mathrm{PV}_{i,t}}$ is the output of the PV system at time step t in node i and $P_{\mathrm{L}_{i,t}}$ is the load demand in the same node and time step. An example of the resulting node net power daily profile used in the simulations is shown in Figure 6.2.

Another pair of parameters that need to be set before running the code is the initial state of charge of each battery and its energy capacity. The set of State of Charge was chosen in such a way that there was a clear difference of initial SoC between different households, with all batteries falling within the limits allowed (0.2-0.8). Concerning the sizing of the batteries, namely their energy capacity, it was fine tuned so that the micro-grid would not suffer any loss of load nor would dump any energy, but so that at least some of the batteries would reach the limits of allowed state of charge during operation. The battery capacity was then set at 11200 Wh, homogeneously in all the SHSs. This assumption was made to boost once again energy flow into the grid, but it would require more accurate consideration. It is anyway outside of the scope of this thesis investigating the sizing of SHSs and their components and to analyse loss of load probability and dumped energy. These aspects are analysed in other scientific publications from Delft University of Technology [25][76][77]. To complete the full setup of the battery in terms of model for the tool, as explained in section 6.1, a reference voltage and a droop constant need to be set. A typical reference voltage of 380 V for DC micro-grids was chosen and, following the modelling guidelines of the DC power flow tool, a droop constant of 0.6 was chosen [73].

In order to have homogeneous assumptions for all the different simulations performed on different layouts, after the first randomisation of load profiles, production profiles and set up of the node parameters, the inputs were kept the same throughout the set of simulations, except for the incidence matrix and the edge vectors which are clearly typical and unique per each geometrical layout. For some layouts, namely the Spider diagrams, the Bus topology and some of the Radial networks, in case of non-household nodes, some modifications had to be made in the set up to model those nodes as passive nodes without any power input nor output.

In the next section, the results of a particular set of simulation, namely on optimised layouts for village number 8, are going to be reported, described and discussed, as an example on how the last step of the proposed methodology can be applied to a specific case.

6.4. Results and Discussion

The DC power flow tool was used to calculate the power flow of different layout topologies obtained as a results of the optimisation described and discussed in Chapter 5. In order to give an example of the application of the DC power flow, developed by Chaifouroosh [73] and modified during this thesis project as described

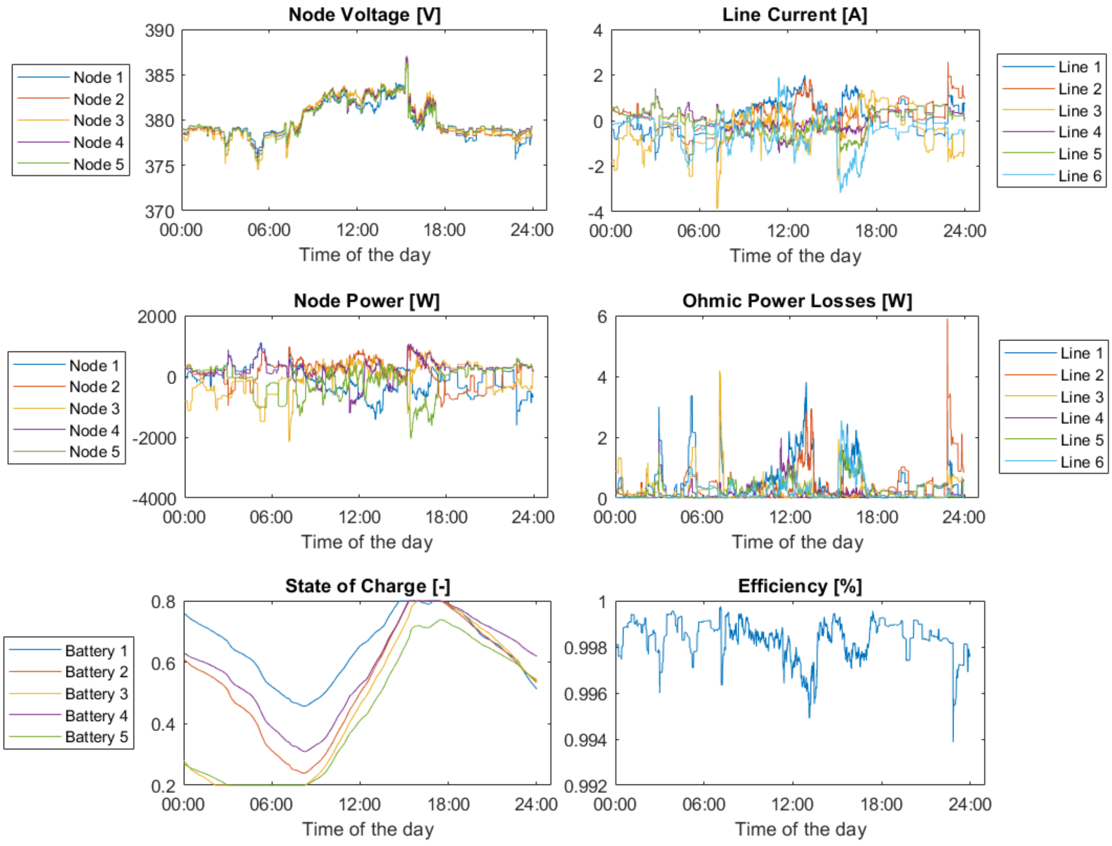


Figure 6.3: Example of graphical output results of the DC power flow simulation for village number 8, with MST as starting topology and wf = 0.2.

in section 6.2, the results for a set of simulations run on the specific optimised layouts for village 8 are going to be described here. The set of simulations were performed with consistent assumptions and parameters, as just described in section 6.3, on all the layouts obtained from the optimisation from four different starting topologies using a range of weight factor from 0.01 to 2. Such a wide range is kept and analysed in order to give a complete overview on how the operational parameter can change depending on the starting topology and on the chosen weight factor. It was chosen not to calculate separately also the result of the optimised network obtained starting from the Radial topology because of its high similarity with the Minimum Spanning Tree option. A summary of the optimisations performed to reach the layouts which were fed to the DC power flow tool can be seen in Table 5.2.

6.4.1. Results Visualisation

The DC power flow simulation performed on each of the proposed layouts gave a set of results, part of which are directly represented by the graphical output shown as an example in Figure 6.3. The simulations are run for a time period of a single day, divided into time steps with one minute of duration, resulting then in 1440 time steps. For each time step intermediate results and parameters are calculated and in these six final plots their fluctuation along the simulated period can be seen.

The three plots shown on the left side of the figures are relative to the single households, showing respectively the voltage fluctuations in each node, the power injected into or absorbed from the micro-grid in each node and the State of Charge of the batteries which are installed in each single SHSs connected to the micro-grid. It is important to mention that, while the number of batteries is always the same (5) in all the layouts, since the households (and the SHSs) are kept the same, the number of nodes of the network is higher for

Table 6.2: Summary of characteristics of proposed optimised layouts and operational parameters resulting from the DC power flow analysis for case study of village 8.

Starting topology	Weight Factor [-]	Redundant Edges [-]	Network Length [m]	Efficiency [%]	Total Losses [W]	Max Voltage [V]	Min Voltage [V]	Max Current [A]	Total Battery Throughput [Wh]	Single Battery Size [Wh]
MST	0.01	5	369.94	99.86	1823	386.86	374.95	2.94	55050	11200
MST	0.1-0.2	2	208.47	99.81	2495	387.07	374.54	3.88	55025	11200
MST	0.3-1.5	1	161.26	99.78	2788	387.26	374.34	3.72	55008	11200
MST	2	0	113.92	99.6	4853	393.37	367.32	6.16	55107	11200
Ring	0.01	5	369.94	99.86	1823	386.86	374.95	2.94	55050	11200
Ring	0.1-0.3	2	208.47	99.81	2495	387.07	374.54	3.88	55025	11200
Ring	0.5-2	1	164.07	99.76	3110	394.29	374.48	5.07	55022	11200
Bus	0.01	8	430.37	99.88	1648	387.6	374.28	2.31	54990	11200
Bus	0.1	5	279.55	99.83	2260	388.09	373.39	3.28	54995	11200
Bus	0.2	4	232.21	99.82	2439	388.17	373.38	3.3	55002	11200
Bus	0.3-0.5	2	174.27	99.73	3461	387.07	372.27	4.34	55024	11200
Bus	1	1	161.94	99.71	3686	387.07	372.07	5.74	55053	11200
Bus	1.5-2	0	118.26	99.55	5668	392.19	369.28	6.81	55072	11200
Spider	0.01	8	458.72	99.88	1613	387.41	374.61	2.01	54996	11200
Spider	0.1	4	258.5	99.82	2472	388.08	373.4	3.08	55005	11200
Spider	0.2-0.3	3	211.16	99.8	2640	388.1	373.39	3.04	55009	11200
Spider	0.5-2	2	167.48	99.74	3383	389.66	372.48	3.66	55014	11200

other kind of starting topologies (eg. 8 for the Bus starting topology and 6 for the Spider starting topology). The voltage fluctuations are going to be measured also for the additional passive nodes not representing real households (and SHSs), as well as the power in each node, which is supposed to be null for passive nodes.

The three plots on the right side of Figure 6.3 are related to line elements of the network, showing respectively the current flow in each separate line segment of the micro-grid, the losses at each time step and the overall efficiency of the network per time step, as calculated in Eq. 6.8. The number of lines in the network greatly varies from layout to layout, ranging from a minimum of 4 lines to a maximum of 16 lines. It is then not important to compare directly the ohmic power losses of each single line, while it is more relevant to compare the cumulative sum of the losses along the network at a given instant, which is represented by the instantaneous efficiency shown in the last plot. It is though of great importance to separately examine the current flow in each of the single line segment, regardless of its direction, to check if it stays within the technical limits for the chosen type of wire and to understand if there are specific lines which are more congested and stressed in a given layout. Examples of these considerations are going to be given later in this analysis.

6.4.2. Results Analysis

Before diving into a detailed analysis, an overall summary of some of the parameters and characteristics resulting from the DC power flow simulations of the different optimised layouts is given in Table 6.2. In the table it is possible to see, for each of the analysed layouts, which was the starting topology from which the optimised layout was found and using which weight factor(s). It is important to note that for some starting topologies identical results were reached using different weight factors, so they have been grouped together both in the simulation and in the results analysis. Moreover, the number of redundant edges are listed to give an idea of the level of redundancy and fault resistance of each proposed layout, as discussed in Section 5.1. In addition to that, even if it is not a result of the DC power flow analysis, it is important to take into consideration the actual network length of each of the proposed layouts, which is a direct representation of the costs. The other parameters listed in Table 6.2 are a direct output of the DC power flow simulations run with the described assumptions and modifications: network efficiency, calculated as defined in Eq. 6.8, the power losses throughout the simulated day time, the maximum peak of voltage experienced in the network, the minimum dip of voltage, the maximum current flow through any segment of the network and the cumulative energy throughput experienced by the batteries in the SHSs connected to the micro-grid. To avoid confusion, it must be said that applying a weight factor of 0.01, the layouts resulting from a starting MST and from a starting Ring topology end up being completely identical, that is why the parameters in the table perfectly match.

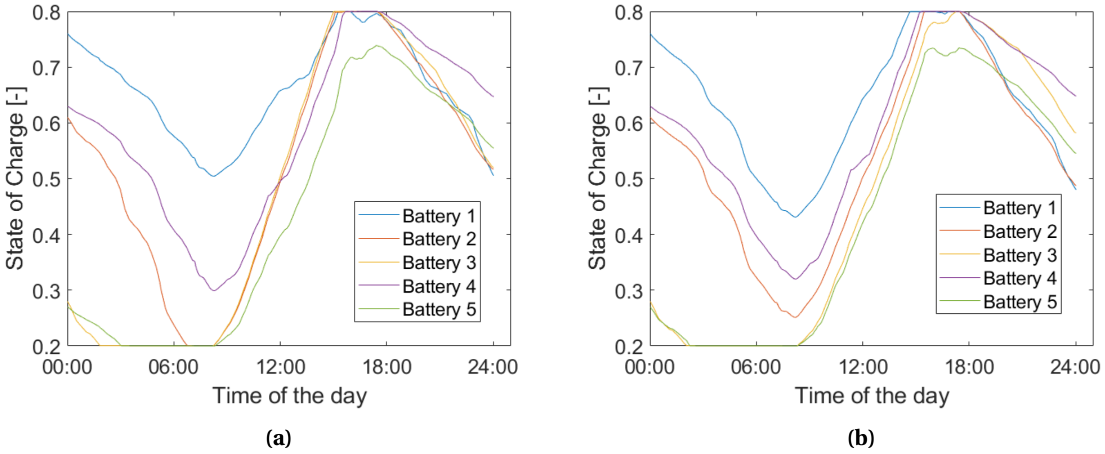


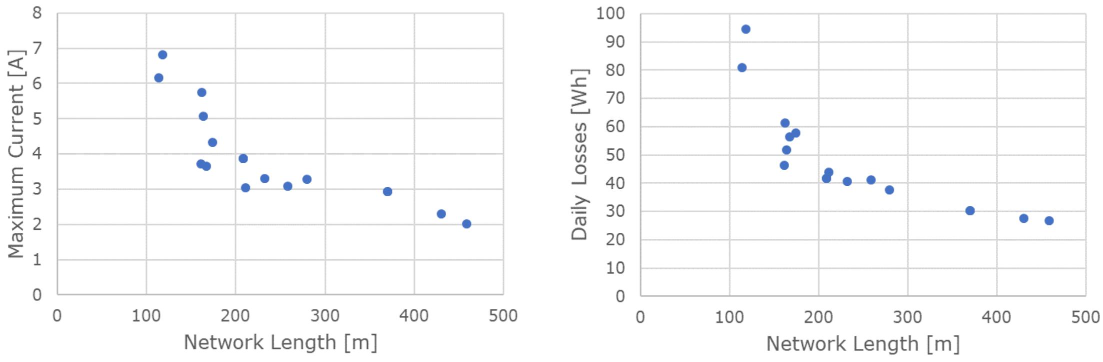
Figure 6.4: Comparison of evolution in time of the State of Charge of SHS batteries in two different layouts for village 8: (a) MST-based with wf = 2 and (b) Bus-based with wf = 2.

Battery throughput

A first consideration which can be done looking at these aggregated figures is that the layout of the micro-grid is practically not affecting at all the total amount of energy which is being absorbed and injected into the grid by the batteries installed into the single SHSs, which is represented by the battery throughput, showing very little variation in the sample of layouts analysed and listed in the table. This means, trivially, that the micro-grid layout is almost neutral to the exchange of energy happening inside the micro-grid, but it can be stated that it affects the way these energy exchange happens, to and from which batteries is the energy flowing and the paths along the network that it will take. This can be seen for example comparing the evolution in time of the State of Charge of the batteries in two different layouts, namely an MST-based and a Bus-based optimisation with the same weight factor used (2), shown in Figure 6.4. It is possible to see how the profiles are different and how the choice of which batteries are filled/discharged first depends on the layout itself, namely on how the batteries are connected to each other. It can also be noted how, even if the initial SoC configuration is identical, the final SoC of the batteries after one day of simulation show some differences between the two layouts. The exchange of energy itself will also be based on other parameters, which in this case are kept constant along the simulations, such as the load and production profiles, the tier of each households, or that are external to technical considerations, such as the pricing system for energy trade internally to micro-grid, which is the focus of the work of Guarino [25].

Losses

Focusing now on considerations on the energy losses and on the efficiency, it can be seen how all the layouts present quite high values of efficiency. It has to pointed out that this is due to the fact that only losses physically happening in the cable are taken into account. Hence, all the losses happening inside the single SHSs and in the interaction between PV systems, batteries and loads are not taken into account. That is the reason why, rather than comparing directly the efficiency, it is more relevant to compare the power losses, already listed in Table 6.2. As expected, the losses are higher in networks with fewer redundant edges (and consequently lower network length), with the best performance being achieved by an over-redundant (wf = 0.01) spider-based optimised layout and the worst one by the Bus topology itself without added edges (wf = 2). It can be noted that the spider-based best performing topology is also the one with highest overall network length among all the considered alternatives. The increased losses in less redundant topologies is due to the presence of higher current flows into the lines (keeping in mind that cable losses are proportional to the the square of the current, as shown in Eq. 5.1). This is consistently proven by the correlation which can be identified between losses and network length and maximum current measured and network length, as shown in Figure 6.5. Apart from an increase in losses and a decrease in efficiency, all the considered layouts were still falling inside the allowed current ranges for the selected cables, which was up to a maximum of $10A$. It must be noticed, anyway, that, comparing different starting topologies among each other, the Bus topology presents the highest values of current flow in its segments, especially if considering the segments of network which make up the main bus line. This is a clear sign that the risk of congestion in such a topology



(a) (b)

Figure 6.5: Correlation between network length and (a) maximum current measured in the network and (b) daily losses along the network.

is higher than in other considered topologies, since a great amount of power is forced to flow into the main bus to move from one household to another, causing also a dis-homogeneous distribution of the current flow within the network. This can be seen in the comparison between current flows and line losses of Bus topology with another layout with same weighting factor, namely a Ring layout, as shown in Figure 6.6.

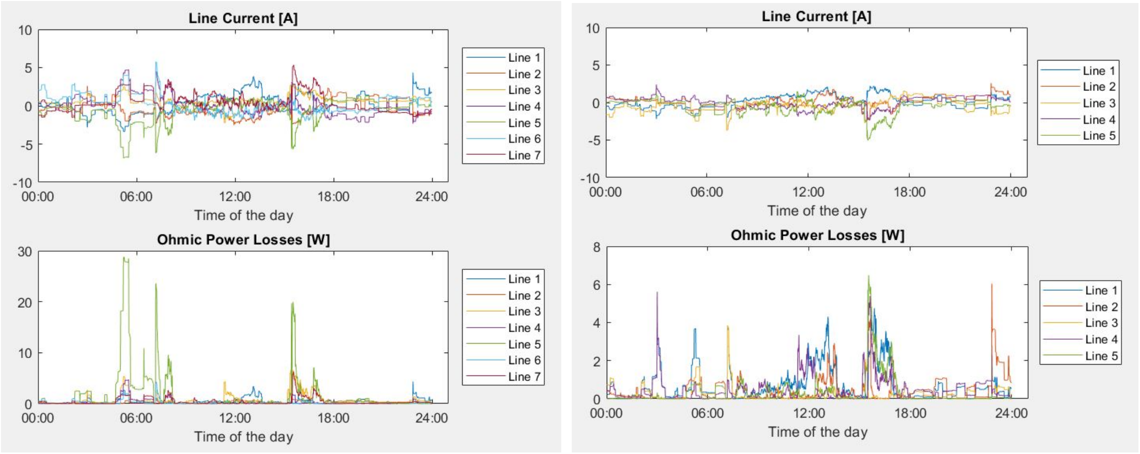
Voltage Fluctuations

Shifting the focus to voltage fluctuations, even though none of the layouts experiences any overvoltage or undervoltage outside of an allowed range of 10% variation around the nominal voltage of 380V, it is possible to notice how more meshed micro-grids (lower weight-factor and more redundancy) present smaller voltage fluctuations than shorter, less redundant network, as it had already been mentioned in literature [46] and [48].

6.4.3. Conclusions

After this first example of application of the modified DC power flow tool, it is possible to draw some preliminary conclusions, even though they would need to be confirmed or confuted by future work and future applications of this same methodology:

- As expected, meshed micro-grid, while being more costly in terms of cable length, grant better stability in terms of voltage fluctuations



(a) (b)

Figure 6.6: Comparison of current flow and line losses of two different layouts for village 8: (a) Bus-based with wf = 2 and (b) Ring-based with wf = 2.

- Bus topology is more keen to congestion along the main bus line than other adopted starting topologies, because of an uneven power flow distribution along the network
- Total amount of battery energy throughput in the network is not strictly depending on the kind of layout adopted, while the path and layout of single household-to-household exchange seems to vary
- Current flow in single lines and cable losses are higher for less redundant, shorter network layouts
- Considering at the same time operational and geometrical parameters, it can be stated that similar technical performances can be reached starting from geometrically optimal topologies (MST) with lower total network length, hence with lower costs

7

Conclusions and Recommendations

In this conclusions and recommendations chapter, the work carried out during this thesis project and the results obtained and described in this report are going to be briefly summarised and discussed. The initial research goals that had been set are going to be repeated, to check to what extent the expectations for this work have been fulfilled. The results of each step of the designed methodology are going to be summarised and interpreted. Finally, some recommendations are going to be suggested for future improvements starting from this piece of work, in order to further expand and perfect the contribution that this methodology can give to scientific research and project realisation in the field of rural electrification.

7.1. Achieved goals

This whole thesis work was started with the purpose of contributing to a series of research efforts, done in the DCE&S group of Delft University of Technology, in the field of rural electrification. Energy access issue is one of the most urgent problem that the modern world is facing, together with a strong need of finding sustainable and environmentally-friendly solutions to boost development and improve living conditions in developing countries, especially in remote off-grid areas.

In this context, after a preliminary literature review, the main scope of the project was then identified as:

> *"To develop a well-structured methodology, based on the use of Geographic Information Systems to gather, process and analyse geographical data, for planning and evaluation of remote DC micro-grid networks to boost rural electrification."*

In order to reach this scope, the following research questions were formulated and subsequently addressed throughout the thesis work:

1. Which has been the approach used so far in addressing rural electrification projects and which are the micro-grid layouts suitable for Solar Home Systems interconnection in rural communities?
2. How can Geographic Information Systems be used to gather, analyse and process ground-level geographical data with the purpose of facilitating micro-grid planning?
3. How can graph theory and network analysis concepts be used to find optimised layouts for DC solar micro-grids?
4. What are the DC power flow operational parameters of different optimised micro-grid layouts and how do they compare to each other?

While answering the first research question, various gaps to fill and additional aspects to be considered were found, becoming then the main focus of this thesis project. Even though a consistent amount of research work was already available in the scientific community, the following points required additional attention:

- Limited implementation of a bottom-up approach rather than a top-down imposed electrification process, not taking into proper consideration the real needs and the dimension of the single villages and communities.

- Lack of access to and use of real ground-level geo-spatial data.

- Underestimation of the usefulness of a thorough evaluation and comparison at internal micro-grid level, based on the characteristics of the single case study, on a geometrical, geographical, economical and technical point of view.

- Lack of a generalised methodology, making use of modern tools and useful concepts, to be applicable to different case studies in different areas of the world.

In order to tackle these issues, it was chosen to focus on developing a step-by-step methodology for micro-grid layout planning, making joint use of Geographic Information Systems and Graph Theory. GIS tools are meant to fill the gap in ground-level data gathering and processing, while graph theory concepts, implemented also directly inside the GIS environment, are used to design in a smarter and optimised way the geo-spatial layout of micro-grids. The resulting methodology is composed by three separate but consequent steps, which are singularly described, along with their intermediate results, each of them directly referring to one of the remaining research questions 2-3-4.

The first step of the methodology which was performed, described in Chapter 4 and referring to research question 2, consists on the use of Geographic Information Systems, namely QGIS, to gather and process data on a sample set of villages scattered in different areas of the world, designing and comparing on a purely geometrical basis different potential micro-grid layouts. The topologies chosen to be compared, on the basis of previous examples found in literature and using additional concepts adopted from graph theories, were the following: Spider topology, Radial topology, Bus topology, Ring topology and Minimum Spanning Tree topology, with the additional use of Steiner Tree topology as a universal benchmark.

The second step of the methodology, described in Chapter 5 and referring more heavily to research question 3, consists in the use of graph theory concepts and algorithms to develop a code, written in Python, to design, using as starting points the topologies just mentioned, optimised micro-grids layouts for the selected villages. The code is based on a dual-objective optimisation, trying to minimise a customisable target function which considers at the same time cost parameters, represented by the average cable length, and operational parameters, represented by the average shortest path length of the network.

The third and last step of the methodology, described in Chapter 6 and referring to research question 4, consists of a DC power flow analysis performed on a set of optimised micro-grid configurations resulting from the previous step. The DC power flow simulation is used to compare the operational parameter of different optimised solutions, using a MATLAB tool developed inside TU Delft, with the implementation of some modifications and tweaking to make it more adequate to the considered case study.

Overall, the designed methodology is presented as an integrated solution, harnessing and combining concepts of different, but related, fields of science. The three steps of the methodology are self-standing and can be used singularly for specific tasks and purposes along the process of rural electrification solutions evaluation, but they can as well be easily integrated and performed subsequently, as done in this thesis work. It leads, starting from simple ground level data on village layout, to a full power flow simulation on optimal suggested layouts, analysing along the process technical, geometrical and economical aspects. This methodology was designed to be as flexible and customisable as possible, in order to easily adapt to single case studies in different locations, environment and with different conditions.

Even though the main purpose of this thesis work was not to already obtain results applying the proposed methodology, it is anyway important and interesting to describe and discuss the outcome reached with this preliminary application to a sample of 42 villages. In the next sections, these intermediate and final results are going to be summarised and discussed as an example of how this methodology can be used in rural electrification projects for specific locations.

7.2. Discussion of results and conclusions

Each single step of the methodology, applied to the identified sample of 42 villages, lead to some intermediate conclusions, which are worth mentioning singularly, before wrapping up the project as a whole and draw some general conclusions based on the overall results obtained.

Starting from the GIS analysis performed in Chapter 4, in which the aim was to compare different suggested topologies on a geometrical length point of view (which is an indicator of network costs), the main takeaways from the results can be summarised as follows:

- **Denser villages mean lower costs.** The internal distribution of the households in a village strongly affects the average network length needed to build an interconnected system. Denser settlements will have to face lower investments per household.

- **Superimposing topology structure leads to higher length.** The comparison between different suggested topologies showed how most of the superimposed layouts which were typically applied in rural electrification projects in the past, such as Spider diagram, Ring topology and Bus topology, are generally more expensive in terms of cable costs. Layouts proposed using application of Graph theory concepts, without structural constraints, especially the Minimum Spanning Tree, lead to consistent savings in the majority of cases, with the Radial topology being the second best alternative case.

- **Each village is different and requires case-specific considerations.**While general trends can be identified in terms of comparison between different topologies, it is of foremost importance to make a non-biased comparison, considering all the options available, because factors such as geometrical layout, shape, density and geographical features can strongly affect the convenience of one topology over the other. Hence the optimal topology in terms of length is extremely case-specific.

From the results of the application of the proposed optimisation algorithm to some case studies, performed and described in Chapter 5, some preliminary considerations can be done, both on the functioning of the algorithm itself and on the optimised layouts:

- **Shorter types of topologies are also better starting topologies for the algorithm.** Those topologies, like Minimum Spanning Tree and Radial topology, which resulted to be optimal in pure terms of network lengths, are also performing better as starting topologies for the dual-objective optimisation algorithms. Meshed topologies formed by edge addition to these two kind of topologies resulted to be the optimal solution in almost all the considered cases, independently from the weight factor used.

- **The choice of weight factor is crucial for the final results and it is case-specific.** Sensitivity studies changing the weight factor showed how significant is the change in optimised layout depending on the choice for this parameter. As expected, lower weight factors, which prioritise redundancy, lead to more edges added to the initial layout, and viceversa. Further tuning and modifications to the target function of the algorithm may be needed to faithfully reflect the priorities and conditions of each single case.

- **Optimal edge addition follows specific criteria.** The edges that are more likely to be added to the initial topology to form meshed micro-grids are shorter edges or edges that tend to close loops inside the network, or cut through already existing large loops. These pieces of information can be used to further modify the edge addition criteria used in the algorithm to perform an even smarter optimisation.

Finally introducing more technical and operational aspects in the analysis of the proposed micro-grid layouts, in Chapter 6, the results of the DC power flow analysis lead to the following considerations to be drawn:

- **Meshed micro-grids grant better network stability.** Layouts with higher redundancy, more cable lines and more meshed structures, even if being more expensive in terms of cable lengths, show better stability in terms of voltage fluctuations in the network nodes. It must be said, though, that none of the optimised network layouts experienced severe overvoltages or undervoltages.

- **Single line congestion strongly depends on the layout.** Analysing single lines performances and current distribution along the network, it can be noticed how the kind of topologies used strongly affects the current patterns. As an example, bus topology leads to severe line congestion in the main bus line, while more meshed topologies grant a more even current distribution.

- **Total energy exchange is barely affected by network layout, but exchange patterns are.** While the total cumulative battery throughput of the households in the village does not change depending on the installed micro-grid layout, the profile of the State of Charge in the simulations show how the single batteries behaviour is affected by the installed layout. It can be stated, then, that the micro-grid layout affects the way of trading energy inside a village or community.

After this step-by-step analysis of intermediate results, more conclusive considerations can be done on the application of this methodology as a whole and on how planning and evaluation of rural electrification process can be made better or more efficient with it. The most important take-away conclusions of this study are summarised here:

- **Combined use of GIS and graph theory is a powerful tool in rural electrification projects.** In this study an integrated use of Geographic Information Systems and Graph Theory concepts has been proposed as main source of novelty. While GIS grant consistence to reality and potential access to huge amount of easy-to-process data, Graph Theory is an instrument to smartly and optimally design micro-grid layouts, as well as giving extra tools for their analysis. The use of these two science fields, too often underestimated in electrical engineering, can be even further exploited in the process of solving the energy access issue.

- **New topology layouts need to be considered in micro-grid planning.** Results from each step of this methodology show how newly proposed network layouts generally outperform, both in terms of technical operating parameters and in cable length savings, classical superimposed structure seen in literature. This does not mean that the micro-grids layouts applied so far need to be fully discarded, but a less uni-directional analysis and comparison need to be done.

- **Case-specific conditions play a fundamental role in the application of this methodology.** While this methodology is set to be universally applied to potentially any case study, each of its steps can be tweaked and customised in order to fully adapt to the single case-specific conditions. It was noticed, throughout each step of this work, that, even if general trends can be identified, exceptions and particular considerations are fundamental to reach a fully optimal solution for each specific case. This means that other external conditions and parameters, that were mentioned but not directly implemented in this specific application, will need to be taken into account for future, deeper applications of this methodology.

7.3. Recommendations for future improvements and applications

This work is part of a wider effort in the direction of addressing rural electrification issues in developing countries, focusing specifically on one of the three possible solutions identified in the literature review for solving the energy access problem: stand-alone micro-grids.

It is then important, as follow-up of this work, to integrate this part of analysis with a more complete and wide analysis of all the possible solutions, possibly taking inspiration from some of the research papers identified in the literature review, which already assess in a detailed and comprehensive way the options of grid extension and single stand-alone household systems. The integration of this work could then lead to a completely accurate comparison of all the available options, increasing the level of detail and accuracy of the cost considerations on the micro-grid level which was found to be missing in literature examples. Additional forms of integration of the work would be to include considerations on energy sharing and pricing mechanisms, as studied in [25], or considerations on optimal sizing of single components of SHSs, as studied in [76].

Apart from integrating this single piece of work into the bigger framework of rural electrification, also some improvements can be implemented in each of the single steps of the proposed methodology to make it more efficient, more accurate or more flexible.

First of all, all the geo-spatial data used throughout the implementation of the first step of the methodology were gathered and processed in a semi-automated way, using partly automated operations and algorithms and partially manual manipulation and processing. The level of automation of some processes, for example recognition of households from satellite imagery and generation of line shapefiles representing micro-grid options, can be improved to make the full operation quicker and more efficient, giving the opportunity to process a higher amount of data repeating automated operations in a subsequential way.

Moreover, while in the first step of the methodology presented in this report cable length of different options is compared directly without being translated to actual costs, the following step for a consistent comparison would be to implement assumptions and conditions typical of each case study to compare real costs of the network, in terms of LCOE or of cost per households. This would allow an easier comparison of the mciro-grid option with the other possible alternatives, namely grid extension and single SHSs.

Sticking to the use of GIS, only a small fraction of the huge potential of application of GIS tools to energy planning was used in this work. Different complementary tasks can be performed, as shown in the example given in Section 7.4 for the case of calculation of distance from the existing transmission infrastructure of the considered villages. Another potential use of GIS tools is going to be mentioned and briefly described in the next and last section of this report, while other alternative uses have been already mentioned in Section 2.4.

In the analysis carried out for this thesis project, a sample of 42 villages in different locations was considered and analysed in a comparative way, already showing some trends, just described in the previous section. The application of the same methodology to larger samples or to other location can give further insights on the usefulness and adequateness of this methodology and could confirm, or maybe disprove, the findings of this work concerning topology comparison.

Moving on to the second step of the methodology, the code is still perfectible, in particular on the definition of the target function to be minimised and concerning the potential addition of more parameters to optimise in order to better quantify the influence of redundancy, losses reduction and congestion avoidance in terms of economic gain (or savings). Moreover, even though some assumptions and smart edge-addition criteria are implemented to make the optimisation faster, the time required to reach fully optimal solutions when the number of households involved increases is still prohibitive. Hence, some modifications to the algorithm can be thought of to solve the problem.

Finally, about the third step of the methodology, which is the application of the DC power flow tool, some modifications were already implemented with respect to the original version developed in [73]. Further improvements can still be done, by including considerations on loss of load probability and energy dump depending on the sizing of each single SHSs. This sizing choice can also be based on the target electrification tier which is chosen for each single project in order to run more demand-specific simulations. The simulations during this thesis work were done with homogeneous electrification tiers along the different households, while in order to assess the operation of different proposed micro-grids layout it could also be interesting to run simulations with dis-homogeneous profiles, or with some households without storage options, in order to evaluate how the micro-grid layout reacts to higher energy exchange and higher stress in terms of power flow along its lines.

As can be seen by this quite extensive list, given the novelty of the methodology proposed, the amount of improvements and modifications which can be applied is relatively high, but the described approach can be considered an optimal starting point a step forward towards a more well-thought decision making process in rural electrification projects.

7.4. An additional example of GIS use: cost comparison with grid extension alternative

Apart from the methodology to compare different micro-grid layouts described in Chapter 4, Geographic Information Systems can be also used for other purposes inside the scope of this thesis work. One example is to compare micro-grid solution as an alternative to main grid extension. Basic steps of the methodology to compare these alternatives using QGIS are going to be described in this section, as an example of other potential uses of GIS outside than the main purpose of the thesis.

The first step of this task would the acquisition of data on the layout of the electricity transmission and distribution grid. An example of this step was performed using Africa as location. The data, in form of a polyline ESRI shapefile (.shp), was retrieved by an open-source database of energy-related data of the World Bank Group [78]. It was then sufficient to import the shapefile into the current QGIS project to have it visualised on the map (see Figure 7.1). It was then possible, using geo-spatial correlation tools inside QGIS, to measure the distance between each village and the nearest point on the electricity grid, and add this as an extra attribute field for the Villages shapefile.

In order to make a comparison between the two alternatives which are being considered (grid extension or stand-alone micro-grid), three cost components have to be quantified: cost of internal grid layout, cost of grid extension to reach the village, cost of solar home systems per household. Both alternatives have to cover the internal grid costs, so the comparison between the two alternatives depends on the other components. The cost of solar home systems depend on the energy tier which needs to be achieved, while the grid extension costs can be assumed to be dependent solely on the distance from the grid, just calculated in QGIS.

The cost of grid extension alternative could then be approximately calculated using the following formula:

$$C_{\text{gridExt}} = d \cdot c_{\text{grid}} + l \cdot c_{\text{mg}} \tag{7.1}$$

in which d is the distance from the grid, l is the length of the micro-grid, c_{grid} is the cost per unit length of grid extension and c_{mg} is the cost per unit length of internal micro-grid, assuming an homogeneous network with equal size of cable. The cost for the stand-alone micro-grid option can be calculated using the following formula:

$$C_{\mathrm{MG}} = hh \cdot c_{\mathrm{SHS}} + l \cdot c_{\mathrm{mg}} \tag{7.2}$$

in which hh is the number of households in the village and c_{SHS} is the cost per single Solar Home System.

The parameters needed for this cost comparison vary greatly depending on the location, the availability of resources and components, the national and local policies, the expected supply and demand of each household and many other factors. It is out of the scope of this thesis to investigate deeper this comparison for single case studies, but it is still relevant to set down a basic proposed methodology to show other potential application of GIS in rural electrification processes.

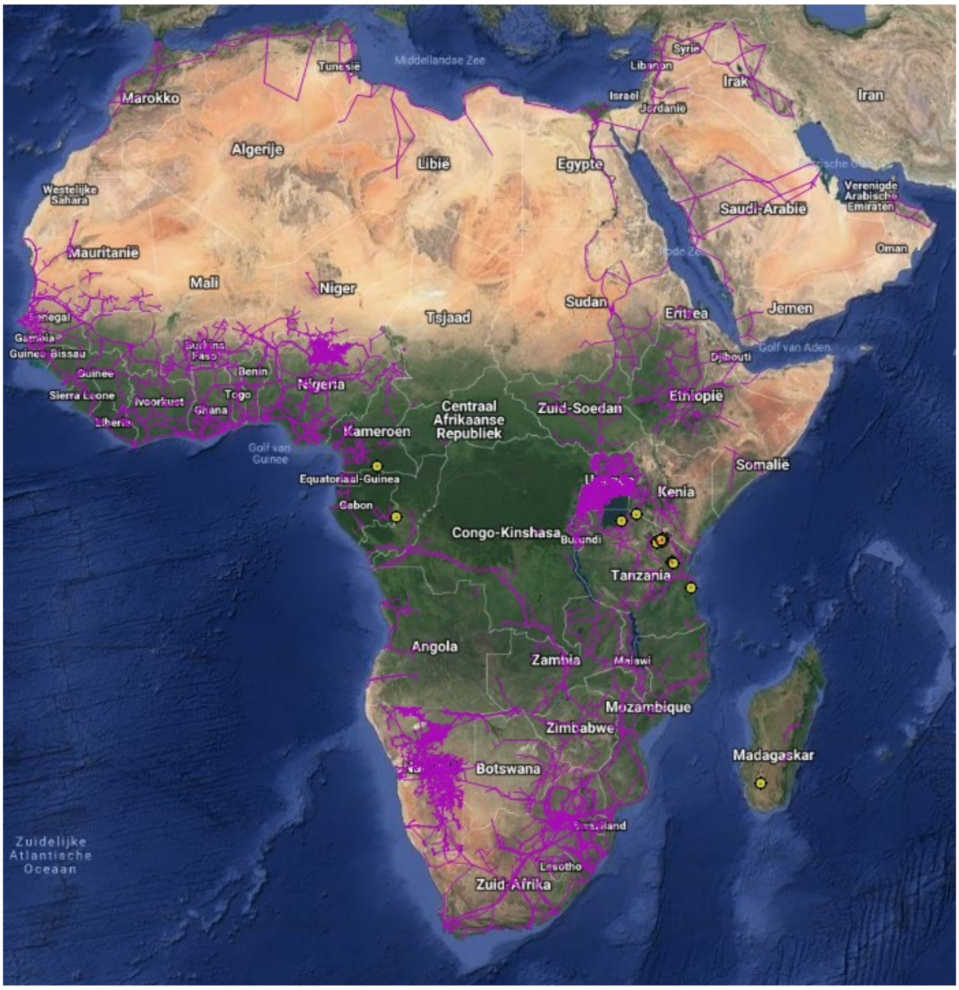


Figure 7.1: Visualisation on QGIS of the African transmission and distribution grid

Bibliography

[1] Hannah Daly and Molly A Walton. *Energy Access Outlook 2017: From Poverty to Prosperity*. IEA, Oct 2017.

[2] World Bank Group. Access to electricity, rural (% of rural population), n.d.. Retrieved from: `https://data.worldbank.org/indicator/EG.ELC.ACCS.RU.ZS?end=2016&start=1990&view=chart&year=2014`.

[3] United Nations. Sustainable development goals, 2015.

[4] Bas J. Van Ruijven, Jules Schers, and Detlef P. Van Vuuren. Model-based scenarios for rural electrification in developing countries. *Energy*, **38**(1):386–397, 2012. doi: 10.1016/j.energy.2011.11.037.

[5] Michael H. Smith, Karlson Hargroves, and Cheryl Desha. *Cents and sustainability: securing our common future by decoupling economic growth from environmental pressures*. Earthscan, 2010.

[6] eia. Global access to electricity has increased over the past two decades, Jun 2017. Retrieved from: `https://www.eia.gov/todayinenergy/detail.php?id=31552#tab2`.

[7] Hisham Zerriffi. *Rural electrification strategies for distributed generation*. Springer, 2011.

[8] James Hazelton, Anna Bruce, and Iain Macgill. A review of the potential benefits and risks of photovoltaic hybrid mini-grid systems. *Renewable Energy*, **67**:222–229, 2014. doi: 10.1016/j.renene.2013.11.026.

[9] Doudou N. Luta and Atanda K. Raji. Decision-making between a grid extension and a rural renewable off-grid system with hydrogen generation. *International Journal of Hydrogen Energy*, **43**(20):9535–9548, 2018. doi: 10.1016/j.ijhydene.2018.04.032.

[10] Debajit Palit and Kaushik R. Bandyopadhyay. Rural electricity access in south asia: Is grid extension the remedy? a critical review. *Renewable and Sustainable Energy Reviews*, **60**:1505–1515, 2016. doi: 10.1016/j.rser.2016.03.034.

[11] Henry Louie. *Off-Grid Electrical Systems in Developing Countries*. Springer International Publishing, Jan 2018. ISBN 978-3-319-91889-1. doi: 10.1007/978-3-319-91890-7.

[12] Mathias Koepke and Sebastian Groh. Against the odds: The potential of swarm electrification for small island development states. *Energy Procedia*, **103**:363–368, 2016. doi: 10.1016/j.egypro.2016.11.300.

[13] Sebastian Groh and Mathias Koepke. A system complexity approach to swarm electrification. *International Symposium for Next Generation Infrastructure Conference Proceedings*, , Dec 2014.

[14] Mikul Bhatia and Niki Angelou. *Beyond Connections: energy access redefined*. ESMAP, World Bank, 2015. doi: 10.1596/24368.

[15] MWH Global. *Sustainable Energy Handbook: 5.1 Rural electrification*, chapter 5, 1–6. MWH, Feb 2016.

[16] Sudeshna G. Banerjee, Mikul Bhatia, Elisa Portale, Ruchi Soni, and Nicolina Angelou. The world bank's perspective on energy access. *Energy poverty: Global challenges and local solutions*, :77–105, Nov 2014.

[17] Sebastian Groh, Shonali Pachauri, and Narasimha D. Rao. What are we measuring? an empirical analysis of household electricity access metrics in rural bangladesh. *Energy for Sustainable Development*, **30**:21 – 31, 2016. ISSN 0973-0826. doi: https://doi.org/10.1016/j.esd.2015.10.007.

[18] Annemarije van Dijk. *The power to produce : the role of energy in poverty reduction through small scale enterprises in the Indian Himalayas*. PhD thesis, University of Twente, Netherlands, Dec 2008.

[19] Mathias Gustavsson and Anders Ellegård. The impact of solar home systems on rural livelihoods. experiences from the nyimba energy service company in zambia. *Renewable Energy*, **29**(7):1059 – 1072, 2004. ISSN 0960-1481. doi: https://doi.org/10.1016/j.renene.2003.11.011.

[20] Daniel Burmester, Ramesh Rayudu, Winston Seah, and Daniel Akinyele. A review of nanogrid topologies and technologies. *Renewable and Sustainable Energy Reviews*, **67**:760 – 775, 2017. ISSN 1364-0321. doi: https://doi.org/10.1016/j.rser.2016.09.073.

[21] Sebastian Groh, Daniel Philipp, Brian E. Lasch, and Hannes Kirchhoff. Swarm electrification: Investigating a paradigm shift through the building of microgrids bottom-up. In *Decentralized Solutions for Developing Economies*, 3–22, Cham, 2015. Springer International Publishing. ISBN 978-3-319-15964-5.

[22] International Energy Agency. *World Energy Outlook 2011*. OECD publishing, 2011. doi: https://doi.org/https://doi.org/10.1787/weo-2011-en.

[23] Mashood Nasir, Hassan A. Khan, Arif Hussain, Laeeq Mateen, and Nauman A. Zaffar. Solar pv-based scalable dc microgrid for rural electrification in developing regions. *IEEE Transactions on Sustainable Energy*, **9**(1):390–399, 2018. doi: 10.1109/tste.2017.2736160.

[24] Tania Urmee, David Harries, and Hans-Gerhard Holtorf. *Photovoltaics for rural electrification in developing countries: a road map*. Springer Verlag, 2016.

[25] Luca Guarino. *Decentralized market platforms for energy sharing in remote islanded DC micro-grids*. Master's thesis, TU Delft, 2019.

[26] Anurag Chauhan and R.P. Saini. A review on integrated renewable energy system based power generation for stand-alone applications: Configurations, storage options, sizing methodologies and control. *Renewable and Sustainable Energy Reviews*, **38**:99 – 120, 2014. ISSN 1364-0321. doi: https://doi.org/10.1016/j.rser.2014.05.079.

[27] Nisha Thirumurthy, Laura Harrington, Daniel Martin, Leena Thomas, Jigmet Takpa, and Reuben Gergan. Opportunities and challenges for solar minigrid development in rural india. *University of North Texas Libraries*, , Jan 2012. doi: 10.2172/1052904.

[28] Dimitrios Mentis, Manuel Welsch, Francesco Fuso Nerini, Oliver Broad, Mark Howells, Morgan Bazilian, and Holger Rogner. A gis-based approach for electrification planning—a case study on nigeria. *Energy for Sustainable Development*, **29**:142 – 150, 2015. ISSN 0973-0826. doi: https://doi.org/10.1016/j.esd.2015.09.007.

[29] Chigueru Tiba, Ana L.B. Candeias, Naum Fraidenraich, Elielza M. de S. Barbosa, Pedro B. de Carvalho Neto, and José B. de Melo Filho. A gis-based decision support tool for renewable energy management and planning in semi-arid rural environments of northeast of brazil. *Renewable Energy*, **35** (12):2921 – 2932, 2010. ISSN 0960-1481. doi: https://doi.org/10.1016/j.renene.2010.05.009.

[30] Francis Kemausuor, Edwin Adkins, Isaac Adu-Poku, Abeeku Brew-Hammond, and Vijay Modi. Electrification planning using network planner tool: The case of ghana. *Energy for Sustainable Development*, **19**:92 – 101, 2014. ISSN 0973-0826. doi: https://doi.org/10.1016/j.esd.2013.12.009.

[31] Francesco Fuso Nerini, Mark Howells, Morgan Bazilian, and Maria F. Gomez. Rural electrification options in the brazilian amazon: A multi-criteria analysis. *Energy for Sustainable Development*, **20**:36 – 48, 2014. ISSN 0973-0826. doi: https://doi.org/10.1016/j.esd.2014.02.005.

[32] Francesco Fuso Nerini, Roger Dargaville, Mark Howells, and Morgan Bazilian. Estimating the cost of energy access: The case of the village of suro craic in timor leste. *Energy*, **79**:385 – 397, 2015. ISSN 0360-5442. doi: https://doi.org/10.1016/j.energy.2014.11.025.

[33] Francesco Fuso Nerini, Oliver Broad, Dimitris Mentis, Manuel Welsch, Morgan Bazilian, and Mark Howells. A cost comparison of technology approaches for improving access to electricity services. *Energy*, **95**:255 – 265, 2016. ISSN 0360-5442. doi: https://doi.org/10.1016/j.energy.2015.11.068.

[34] Marianne Zeyringer, Shonali Pachauri, Erwin Schmid, Johannes Schmidt, Ernst Worrell, and Ulrich B. Morawetz. Analyzing grid extension and stand-alone photovoltaic systems for the cost-effective electrification of kenya. *Energy for Sustainable Development*, **25**:75 – 86, 2015. ISSN 0973-0826. doi: https://doi.org/10.1016/j.esd.2015.01.003.

[35] Dimitrios Mentis, Magnus Andersson, Mark Howells, Holger Rogner, Shahid H. Siyal, Oliver Broad, Alexandros Korkovelos, and Morgan Bazilian. The benefits of geospatial planning in energy access: A case study on ethiopia. *Applied Geography*, **72**:1 – 13, May 2016. doi: 10.1016/j.apgeog.2016.04.009.

[36] Gizem Bolukbasi and Ayse Selin Kocaman. A prize collecting steiner tree approach to least cost evaluation of grid and off-grid electrification systems. *Energy*, **160**:536 – 543, 2018. ISSN 0360-5442. doi: https://doi.org/10.1016/j.energy.2018.07.029.

[37] Francesco Dalla Longa, Teun Strikkers, Tom Kober, and Bob van der Zwaan. Advancing energy access modelling with geographic information system data. *Environmental Modeling & Assessment*, **23**(6):627–637, Dec 2018. ISSN 1573-2967. doi: 10.1007/s10666-018-9627-1.

[38] Sándor Szabó, Katalin Bódis, Thomas Huld, and Magda Moner-Girona. Energy solutions in rural africa: mapping electrification costs of distributed solar and diesel generation versus grid extension. *Environmental Research Letters*, **6**(3):034002, 2011. Retrieved from: `http://stacks.iop.org/1748-9326/6/i=3/a=034002`.

[39] Hannes Kirchhoff. Identifying hidden resources in solar home systems as the basis for bottom-up grids. In *Decentralized Solutions for Developing Economies*, 23–32, Cham, 2015. Springer International Publishing. ISBN 978-3-319-15964-5.

[40] Annette Werth, Mario Tokoro, and Kenji Tanaka. Bottom-up and recursive interconnection for multilayer dc microgrids. *2016 IEEE 16th International Conference on Environment and Electrical Engineering (EEEIC)*, , 2016. doi: 10.1109/eeeic.2016.7555492.

[41] Allen R. Inversin. *Mini-grid Design Manual*, volume 1 of *ESMAP technical paper*. ESMAP, World Bank, 2000.

[42] Mashood Nasir, Hassan A. Khan, Nauman A. Zaffar, Juan C. Vasquez, and Josep M. Guerrero. Scalable solar dc micrigrids: On the path to revolutionizing the electrification architecture of developing communities. *IEEE Electrification Magazine*, **6**(4):63–72, Dec 2018. ISSN 2325-5897. doi: 10.1109/MELE.2018.2871297.

[43] Björn Heinbokel, Hannes Kirchhoff, Tomislav Dragicevic, and Josep Guerrero. Zonal protection of dc swarm microgrids using a novel multi-terminal grid interface with decentralized control. *2015 50th International Universities Power Engineering Conference (UPEC)*, :1–6, Sep 2015. doi: 10.1109/UPEC.2015.7339889.

[44] Jackson John Justo, Francis Mwasilu, Ju Lee, and Jin-Woo Jung. Ac-microgrids versus dc-microgrids with distributed energy resources: A review. *Renewable and Sustainable Energy Reviews*, **24**:387 – 405, 2013. ISSN 1364-0321. doi: https://doi.org/10.1016/j.rser.2013.03.067.

[45] Annette Werth, Nobuyuki Kitamura, Ippei Matsumoto, and Kenji Tanaka. Evaluation of centralized and distributed microgrid topologies and comparison to open energy systems (oes). In *2015 IEEE 15th International Conference on Environment and Electrical Engineering (EEEIC)*, 492–497, Jun 2015. doi: 10.1109/EEEIC.2015.7165211.

[46] Rabiul F. Islam, Krishneel Prakash, Kabir A. Mamun, Avneel Lallu, and Hemanshu Pota. Aromatic network: A novel structure for power distribution system. *IEEE Access*, **5**:25236–25257, Oct 2017. ISSN 2169-3536. doi: 10.1109/ACCESS.2017.2767037.

[47] Tomislav Dragicevic, Xiaonan Lu, Juan C. Vasquez, and Josep Guerrero. Dc microgrids–part ii: A review of power architectures, applications and standardization issues. *IEEE Transactions on Power Electronics*, **31**:1–1, Jan 2015. doi: 10.1109/TPEL.2015.2464277.

[48] Krishneel Prakash, Avneel Lallu, Rabiul F. Islam, and Kabir A. Mamun. Review of power system distribution network architecture. In *2016 3rd Asia-Pacific World Congress on Computer Science and Engineering (APWC on CSE)*, 124–130, Dec 2016. doi: 10.1109/APWC-on-CSE.2016.030.

[49] 1000 gis applications & uses - how gis is changing the world, Dec 2018. Retrieved from: `https://gisgeography.com/gis-applications-uses/`.

[50] Dimitrios Mentis, Sebastian Hermann, Mark Howells, Manuel Welsch, and Shahid Hussain Siyal. Assessing the technical wind energy potential in africa a gis-based approach. *Renewable Energy*, **83**:110 – 125, 2015. ISSN 0960-1481. doi: https://doi.org/10.1016/j.renene.2015.03.072.

[51] Sanusi Ohiare. Expanding electricity access to all in nigeria: a spatial planning and cost analysis. *Energy, Sustainability and Society*, **5**(1):8, Mar 2015. ISSN 2192-0567. doi: 10.1186/s13705-015-0037-9.

[52] Petra Heijnen. Sen1241 optimal trees in network design, 2018.

[53] Marcus Brazil and Martin Zachariasen. *Optimal Interconnection Trees in the Plane*, volume 29. Springer, 05 2015. doi: 10.1007/978-3-319-13915-9.

[54] Joseph B. Kruskal. On the shortest spanning subtree of a graph and the traveling salesman problem. *Proceedings of the American Mathematical Society*, **7**(1):48–48, Jan 1956. doi: 10.1090/s0002-9939-1956-0078686-7.

[55] Robert C. Prim. Shortest connection networks and some generalizations. *Bell System Technical Journal*, **36**(6):1389–1401, 1957. doi: 10.1002/j.1538-7305.1957.tb01515.x.

[56] Ronald L. Graham and Pavol Hell. On the history of the minimum spanning tree problem. *IEEE Annals of the History of Computing*, **7**(1):43–57, 1985. doi: 10.1109/mahc.1985.10011.

[57] Pawel Winter. Steiner problem in networks: A survey. *Networks*, **17**(2):129–167, 1987. doi: 10.1002/net.3230170203.

[58] Florian Dörfler, John W. Simpson-Porco, and Francesco Bullo. Electrical networks and algebraic graph theory: Models, properties, and applications. *Proceedings of the IEEE*, **106**(5):977–1005, May 2018. ISSN 0018-9219. doi: 10.1109/JPROC.2018.2821924.

[59] Edsger W. Dijkstra. A note on two problems in connexion with graphs. *Numerische Mathematik*, **1**(1):269–271, 1959. doi: 10.1007/bf01386390.

[60] Robert W. Floyd. Algorithm 97: Shortest path. *Communications of the ACM*, **5**(6):345, Jan 1962. doi: 10.1145/367766.368168.

[61] Anita Amberg, Wolfgang Domschke, and Stefan Voß. Capacitated minimum spanning trees: algorithms using intelligent search. *Combinatorial Optimization: Theory and Practice*, **1**:9–40, 1996.

[62] Steven Nolan, Scott Strachan, Puran Rakhra, and Damien Frame. Optimized network planning of minigrids for the rural electrification of developing countries. *2017 IEEE PES PowerAfrica*, :489–494, Jun 2017. doi: 10.1109/PowerAfrica.2017.7991274.

[63] About arcgis, n.d. Retrieved from: https://www.esri.com/en-us/arcgis/about-arcgis/overview.

[64] Qgis, n.d. Retrieved from: https://qgis.org/en/site/.

[65] Rural electrification densification programme - all villages, n.d. Retrieved from: http://tudelft.maps.arcgis.com/home/webmap/viewer.html?url=http://services5.arcgis.com/HfoIeLKiUrwVPv5a/ArcGIS/rest/services/REDP_all_villages_25_04_2016/FeatureServer&source=sd.

[66] Hard to reach unions in bangladesh, n.d. Retrieved from: http://www.arcgis.com/home/webmap/viewer.html?url=http://services5.arcgis.com/76j9X5O0jqKLGy6f/ArcGIS/rest/services/HardtoReachUnionsofBangladesh/FeatureServer&source=sd.

[67] Google hybrid map, n.d. Retrieved from: https://mt1.google.com/vt/lyrs=y&x={x}&y={y}&z={z}.

[68] Online crs registry, n.d. Retrieved from: http://www.epsg.org/.

[69] Projected coordinate systems, map projections, parameters, and linear units of measure, n.d. Retrieved from: http://desktop.arcgis.com/en/arcmap/10.3/guide-books/map-projections/pdf/projected_coordinate_systems.pdf.

[70] Petra W. Heijnen, Emile J. L. Chappin, and Pauline M. Herder. A method for designing minimum-cost multi-source multi-sink network layouts. (under review), 2018.

[71] Zdzislaw A. Melzak. On the problem of steiner. *Canadian Mathematical Bulletin Bulletin canadien de mathématiques*, **4**:143–148, Jan 1961. doi: 10.4153/cmb-1961-016-2.

[72] Petra W. Heijnen, Emile J.L. Chappin, Gijsbert Korevaar, and Sabine Pelka. Design of integrated energy systems - sen1531 reader, 2018.

[73] Dario Chaifouroosh. *Direct Current Power Flow: Computational methods and low voltage applications*. Master's thesis, TU Delft, 2018.

[74] Tomislav Dragicevic, Xiaonan Lu, Juan Vasquez, and Josep Guerrero. Dc microgrids–part i: A review of control strategies and stabilization techniques. *IEEE Transactions on Power Electronics*, **31**(7): 4876–4891, 2015. doi: 10.1109/tpel.2015.2478859.

[75] ASTM international. Astm standard b 258 - 02: Standard specification for standard nominal diameters and cross-sectional areas of awg sizes of solid round wires used as electrical conductors, 2002.

[76] Ali Chamseddine. *Off-grid PV systems for rural electrification: Optimizing the sizing methodology for modular off-grid PV systems*. Master's thesis, TU Delft, 2018.

[77] Nishant Narayan, Zian Qin, Jelena Popovic-Gerber, Jan-Carel Diehl, Pavol Bauer, and Miroslav Zeman. Stochastic load profile construction for the multi-tier framework for household electricity access using off-grid dc appliances. *Energy Efficiency*, , Nov 2018. ISSN 1570-6478. doi: 10.1007/s12053-018-9725-6.

[78] World Bank Group. Africa: Electricity transmission and distribution grid map, n.d..