

A PRIZE COLLECTING STEINER TREE APPROACH TO LEAST COST EVALUATION OF GRID AND OFF-GRID ELECTRIFICATION SYSTEMS

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF ENGINEERING AND SCIENCE
OF BILKENT UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR
THE DEGREE OF
MASTER OF SCIENCE
IN
INDUSTRIAL ENGINEERING

By
Gizem Bölükbaşı
July 2017

A Prize Collecting Steiner Tree Approach to Least Cost Evaluation of
Grid and Off-grid Electrification Systems

By Gizem Bölükbaşı

July 2017

We certify that we have read this thesis and that in our opinion it is fully adequate,
in scope and in quality, as a thesis for the degree of Master of Science.

Ayşe Selin Kocaman(Advisor)

Özlem Karsu

Gültekin Kuyzu

Approved for the Graduate School of Engineering and Science:

Ezhan Karahan
Director of the Graduate School

ABSTRACT

A PRIZE COLLECTING STEINER TREE APPROACH TO LEAST COST EVALUATION OF GRID AND OFF-GRID ELECTRIFICATION SYSTEMS

Gizem Bölükbaşı

M.S. in Industrial Engineering

Advisor: Ayşe Selin Kocaman

July 2017

The lack of access to electricity in developing countries necessitates spatial electricity planning for guiding sustainable electrification projects that evaluate the costs of centralized systems vis-a-vis decentralized approaches. Heuristic approaches have been widely used in such electrification problems to find feasible, cost effective solutions; however, most of the time global optimality of these solutions is not guaranteed. Our thesis through its modeling approach provides a new methodology to find the least cost solution to this electrification problem. We model the spatial network planning problem as Prize Collecting Steiner Tree problem which would be base for a decision support tool for rural electrification. This new method is systematically assessed using both randomly generated data and real data from rural regions across Sub-Saharan Africa. Comparative results for the proposed approach and a widely used heuristic method are presented based on computational experiments. Additionally, a bi-objective approach that permits to take carbon emission level into the account is implemented and experimented with numerical data.

Keywords: Energy Infrastructure Planning, Grid, Isolated Systems, Rural Electrification, Minimum Spanning Tree, Spatial Electricity Planning, Prize Collecting Steiner Tree, Sub-Saharan Africa, Bi-objective Optimization.

ÖZET

İZOLE VE ŞEBEKEYE BAĞLI ELEKTRİK SİSTEMLERİNİN DEĞERLENDİRİLMESİNDE ÖDÜL TOPLAYAN STEINER AĞACI YAKLAŞIMI

Gizem Bölükbaşı

Endüstri Mühendisliği, Yüksek Lisans

Tez Danışmanı: Ayşe Selin Kocaman

Temmuz 2017

Gelişmekte olan ülkelerde elektrik erişiminin olmaması, sürdürülebilir elektrikleştirme projelerine rehberlik etmek amacıyla merkezi ve merkezi olmayan sistemlerin maliyetlerini değerlendiren elektrik planlaması yapılmasını gerektirmektedir. Sezgisel yaklaşımlar, uygulanabilir, uygun maliyetli çözümler bulmak için bu tür elektrikleştirme problemlerinde yaygın şekilde kullanılmaktadır. Bununla birlikte, çoğu kez bu çözümlerin küresel en iyiliği garanti edilmez. Bu çalışmada modelleme yaklaşımıyla, bu elektrikleştirme problemine en düşük maliyetli çözümü bulmak için yeni bir metodoloji sunmaktadır. Uzamsal ağ planlama problemi, kırsal alan elektrikleştirme projeleri için bir karar destek aracı olarak kullanılması hedeflenerek Ödül Toplayan Steiner Ağacı problemi olarak modellenmiştir. Bu yeni yöntem, rasgele oluşturulmuş veriler ve Sahra Altı Afrika boyunca kırsal bölgelerden gelen gerçek veriler kullanılarak sistematik olarak değerlendirilmiştir. Önerilen yaklaşım ve yaygın olarak kullanılan sezgisel bir yöntem için karşılaştırmalı sonuçlar, sayısal analizlere dayanarak sunulmuştur. Buna ek olarak, hesaba karbon emisyonu seviyesini de katmaya izin veren, iki amaçlı bir yaklaşım uygulanmış ve sayısal verilerle denenmiştir.

Anahtar sözcükler: Enerji altyapı planlaması, Şebeke, İzole enerji sistemleri, Kırsal elektrikleştirme, Ödül Toplayan Steiner Ağacı, İki amaçlı optimizasyon.

*To my Grandma,
and other women who were deprived of educational opportunities.*



Acknowledgement

It is a genuine pleasure to express my most sincere sense of gratitude to Asst. Prof. Ayşe Selin Kocaman, my advisor, for all her support, patience and kindness throughout my research journey. She has always been an understanding and encouraging advisor. I also owe a deep sense of gratitude to all professors of our department for their support and kindness during the master's courses.

I would like to thank Asst. Prof. Özlem Karsu and Asst. Prof. Gültekin Kuyzu for accepting to read and evaluate my thesis.

I also would like to extend my sincere thanks to my colleagues from EA-307 and my friends who have supported me in various ways. I am also grateful to my roommate, Elfe for reminding me to smile during hardest times of the master's journey. The graduate study would not be bearable without them.

I couldn't have achieved anything without my family. Finally, I am extremely thankful to my mother, my father and my sister for their support and love. I always feel their precious support in my life and words can't explain my love for them.

Contents

1	Introduction	1
2	Literature Review	6
2.1	Cost Minimizing Rural Electrification Planning Problems	6
2.2	Cost and CO ₂ Minimizing Rural Electrification Planning Problems .	12
3	Problem Definition	16
4	Solution Approach	21
4.1	An Existing Approach: Modified Kruskal's Algorithm	21
4.2	Proposed Approach: Prize Collecting Steiner Tree	23
4.3	The Bi-Objective Model	27
5	Numerical Analysis	31
5.1	Experiments with Synthetic Data	32

5.2	Experiments with Case Studies from Network Planner	34
5.3	Sensitivity Analysis on Distance Parameters	38
5.4	Computational Analysis for the Bi-Objective Problem	41
6	Conclusion	49
A	Results	59



List of Figures

3.1	A Representative Example of a Result of the Electrification Project (Source: Sustainable Engineering Laboratory, Columbia University) .	18
5.1	Final Grid obtained by PCST method of 102-node instance a)with MK, b)with PCST	37
5.2	Difference between the tails of two trees obtained by MK and PCST approached on 102-node example	38
5.3	a) Randomly generated 100 noded sample(Base Scenario), and its re- duced forms by b) 75%, c) 50% and d) 25%, respectively	39
5.4	Pareto Solution Set of 10-node Synthetic Data	42
5.5	Pareto Solution Set of 20-node Synthetic Data	42
5.6	Pareto Solution Set of 50-node Synthetic Data	43
5.7	Pareto Solution Set of 100-node Synthetic Data	43
5.8	Histogram of 141-node Real Demand Data	45

5.9	Pareto Solution Set of 141 noded Real Life Instance	46
5.10	Network Result Obtained with ϵ -Constraint Method One Iteration Before the Removal of the Extreme Node	48
5.11	Network Result Obtained with ϵ -Constraint Method Right After the Removal of the Extreme Node	48



List of Tables

5.1	Comparative Analysis of Two Methods Based on Randomly Generated Instances	33
5.2	Comparative Analysis of Two Methods Based on Real Life Instance Results	35
5.3	Comparative analysis of two methods with respect to distance variations	39
A.1	Carbon Emission and Total Cost(\$) Results Obtained by Epsilon-Constraint Method for the Synthetic Data n=10	59
A.2	Carbon Emission and Total Cost(\$) Results Obtained by Epsilon-Constraint Method for the Synthetic Data n=20	60
A.3	Carbon Emission and Total Cost(\$) Results Obtained by Epsilon-Constraint Method for the Synthetic Data n=50	61
A.4	Carbon Emission and Total Cost(\$) Results Obtained by Epsilon-Constraint Method for the Synthetic Data n=100	62
A.6	Carbon Emission and Total Cost(\$) Results Obtained by Epsilon-Constraint Method for the Real Data n=141	64

Chapter 1

Introduction

After the industrial revolution, continuous increase of greenhouse gas concentration in the atmosphere leads to significant changes in the Earth's climate. According to monthly analysis of global temperatures by scientists at NASA's Goddard Institute for Space Studies (GISS) in New York, May 2017 was the second-warmest May in 137 years of modern record-keeping [1]. This fact shows us how global warming is becoming more and more important. The concept called "The Greenhouse Effect" explains the current situation: the percentage of greenhouse gases in atmosphere increases which trap the heat that is coming from sunlight. First, the sunlight arrives the surface of the Earth and this is where it is absorbed and radiated back to the atmosphere. The greenhouse gases in the atmosphere do not let the heat spread through the space, in other words, the more greenhouse gases are in atmosphere, the more heat preserved in the Earth. Earth's climate is warming as a result of emissions of greenhouse gases, particularly carbon dioxide (CO_2) from fossil fuel usage. Emissions of non- CO_2 greenhouse gases, such as methane, nitrous oxide and ozone-depleting substances also contribute significantly to warming; however, most of them has shorter lifetimes than CO_2 [2].

Scientific studies show that human activity is the main cause of global warming

caused by carbon emission due to electricity generation, deforestation, agriculture and transport [3]. After the Industrial Revolution CO₂ levels increased from 280 parts per million (ppm) to 400 ppm because of human related activities, particularly fossil fuel combustion [4]. The current global energy system is dominated by fossil fuels [5]. To decrease the speed of global warming, carbon emission caused by human activity should be prevented by replacing old technologies with eco-friendly options. Renewable energy technologies offer a significant opportunity for mitigation of greenhouse gas emission and reducing global warming through replacing traditional energy sources [6].

Today's conventional energy sources are dominated by centralized electricity systems. In centralized systems, electricity is produced at large scale centralized facilities and distributed to the users through a transmission and distribution network. These systems require huge amount of infrastructural investments and generally work with fossil fuels [7]. Approximately 90% of human produced CO₂ emissions occur when burning fossil fuels [8]. Considering the environmental damage and human health problems due to carbon emission, there is a significant need for alternative energy sources. In addition to that, fossil fuels are limited. Shafiee and Topal estimates that coal will be the only fossil fuel after 2042 and in 2112 all the fossil fuels will run out [9]. These facts explain why alternative energy systems are becoming increasingly important for effective electrification.

Unlike centralized energy systems, decentralized energy systems are generally running with eco-friendly renewable energy sources. There are various renewable energy resources that can be utilized to decrease carbon emission level such as solar energy, biomass, wind power, hydro power, tidal power, wave power, geothermal power, biomass and biofuel. In the past 30 years, especially solar and wind power systems have become highly preferred as a result of their declining investment and electricity generation costs, and its improved performance characteristics [10]. Decentralized systems are especially appropriate for isolated rural communities since they are convenient for small demand scales. In general, they are located in areas where renewable

sources are eligible. They can be implemented both in presence of grid or as a stand-alone system isolated from the grid network.

Stand-alone systems are generally useful for the small sized local energy demand of remote locations. These kind of systems are not generally convenient for large scaled demands due to intermittency of natural sources such as wind or solar. One solution to this problem can be to support the stand-alone systems with energy storage systems. Moreover, to provide the base-load power, diversified systems using various renewable energy sources together can be considered. As an alternative, using natural gas and/or biomass-based power together with the renewable technologies can be implemented if there is a huge risk of shortage [10]. In contrast to the stand-alone systems, grid-connected forms of decentralized systems eliminate the risk of energy shortage since they use the grid connection as a support mechanism when there is power shortage that cannot be satisfied by the renewable sources. These kinds of systems are connected to the grid; therefore, they should be located close to the main network. To sum up, there are several decentralized technology alternatives and they should be reviewed in details while making the electricity planning. Motivated by that need, Kaundinya et al. made a detailed analysis of 102 different articles investigating both grid-connected and stand-alone decentralized systems analysing both economic, environmental and technological aspects [11].

Despite the fact that modern technology offers us cleaner and cheaper energy options, absence of electricity still remains as a problem for under-developed countries. International Energy Agency estimated that nearly 1.2 billion people do not have access to electricity in 2015 [12] and a majority of those are in Sub-Saharan Africa countries [13]. In this region, electricity coverage percentage is approximately 30% and in some countries such as Chad, Liberia, and Sudan it is lower than 10% [14]. However, developments in health, education, environmental sustainability, and agriculture depend on the electricity coverage [15]. Therefore, lack of access to electricity is a fundamental problem, which requires spatial electricity planning projects that evaluate the costs of centralized approaches vis-a-vis decentralized approaches

in these countries. Although cost is the major factor to be considered, in the light of global warming issues we can point out that ecological concerns should also be taken into account in the sustainable development plans.

One of the most common definitions of sustainable development is as follows: “development that meets the needs of the present without compromising the ability of future generations to meet their own needs” [16]. Sustainable development has various aspects but it is obvious that it cannot be considered without providing effective and efficient usage of sustainable energy resources. The fast-growing world economy and population are enough to show that energy needs would rise rapidly. This need should be satisfied by using maximum amount of clean energy supply since traditional sources are related “not only to global warming, but also to such environmental concerns as air pollution, acid precipitation, ozone depletion, forest destruction, and emission of radioactive substances” [17]. That’s why there is a strong connection between building a sustainable world and using clean renewable energy sources.

In order to implement cleaner power system technologies to the developed countries, we should first deal with changing existing conventional approaches. However, it is faster and easier to adapt new technologies to the under-developed countries since we can consider bypassing centralized approaches entirely for some regions and using directly renewable technologies. This concept is known as leapfrogging [18]. It should be underlined that this shift to the new technologies requires detailed spatial electricity network planning that can be realized through well-designed electrification projects.

The electrification projects aim to find the least cost alternative to electricity expansion which can be provided with either centralized or decentralized electrification systems. The main goal is to evaluate these options based on factors such as demographic conditions, investment costs, geographic features of the regions or the greenhouse gas emissions to find the least cost scenario. Considering all these parameters, using large data sets makes the optimization problems difficult to solve.

Moreover, it is very important to know the quality as well as the running time of the solution approaches to be able to use them as a decision support tool in large projects. Therefore, the optimization methods that the decision support mechanisms will be based on have significant impact on spatial electricity network planning. This thesis, through its modeling approach, provides a new methodology to find the least cost solution of electricity networks that are designed as a combination of centralized and decentralized options rapidly and optimally. Taking into account both the cost of internal and external costs of the centralized option and cost of the off-grid option, our method determines the grid-compatible nodes and the network connections between them. This method aims to be used in electrification projects of unelectrified rural regions as a decision support mechanism and offers to build a network design from scratch. For that purpose, the problem is implemented as a Prize Collecting Steiner Tree problem and it is tested by the help of computational experiments both with synthetic and real data. As a second step, a bi-objective version of the model is introduced so that the decision-maker will be able to take into account carbon emission in addition to cost of the project.

The rest of this thesis is organized as follows: In Chapter 2, the literature review of single and multi-objective electrification planning problems are presented. In Chapter 3, our problem definition is introduced. In Chapter 4, existing solution approach and our solution methodology along with its single-objective and bi-objective version are explained in details. In Chapter 5, a computational analysis of both methods performed with synthetic data and real life data is introduced. Finally, in Chapter 6, concluding remarks can be found.

Chapter 2

Literature Review

In this chapter, studies on electrification planning problem will be reviewed. In the first part studies that are focusing on single objective versions will be introduced. Secondly, multi-objective solution methodologies applied on the rural electrification planning will be explained in detail.

2.1 Cost Minimizing Rural Electrification Planning Problems

An extensive number of studies that focus on electrification planning incorporating both centralized and decentralized systems can be found in the literature. Most of these studies consider economic and spatial parameters using heuristic approaches. Since the energy systems planning is a widely studied issue in the literature, in this review we will address mainly the issue of network design in electrification planning.

Jebaraj and Iniyar studied on a review paper on energy models in general [19]. In

this paper, studies from the literature about various topics related to energy modeling problems were presented under the topics of “energy planning models, energy supply-demand models, forecasting models, renewable energy models, emission reduction models, optimization models”. This article can be considered as a brief summary of energy modeling studies done until 2006.

Abdul-Salam et al. developed a heuristic algorithm based on hierarchical lexicographic programming to examine the politico-economics of electrification planning of developing countries using a case study of unelectrified communities in Ghana [20]. This algorithm yields flexible solutions in terms of both cost efficiency and political economy. At first, it finds the grid compatible nodes and then it assigns simultaneous routes that provides connection between grid compatible nodes and existing network. In this research, they found out that different preferences cause major regional differences in grid access.

A study focusing on the regional version of the problem presented a method that combines modified Minimum Spanning Tree and simulated annealing algorithm [21]. In this article, the small-sized settlement’s electrification problem was considered as a two-level optimization problem: upper level and lower level design. The upper level design is focusing on the optimum placement of transformers that are connected to each other with MV(medium-voltage) wire. The lower level part aims to find the optimum grid network design taking into account the transformers distribution. The simulated annealing algorithm was used for both levels of the problem. Additionally, in order to slow down the algorithm, minimum spanning tree approach that provides a starting point was implemented. To do the verification of the methods, exhaustive enumeration algorithms for both upper and lower level are used.

Another research suggested a variation of Prim’s algorithm for the Minimum Spanning Tree Problem that guarantees highest prize in each step [22]. Their results highlight that decentralized electrification methods that use renewable energy sources will

be an essential part of expanding rural electricity access. They called the energy planners for action to detect the possible developments in expanding electricity access in Sub-Saharan Africa using especially the renewable energy sources.

In 2012, Weighted Composite Prim's algorithm was suggested to make the choice between centralized and decentralized energy systems [23]. In this research, this analysis was conducted for 150 countries taking into account parameters such as population distribution, electricity consumption, transmission cost, and the cost difference between decentralized and centralized electricity generation. The effects of electricity generation and transmission costs of both decentralized and centralized electrification systems while finding the cost-optimized networks with their algorithm were investigated in details. In addition to that, the results for Uganda, Botswana and Bangladesh were presented. As a result of the experiments made by 150 different countries, they indicated that centralized electricity is the least cost option for a significant majority of the world. However they underlined the fact that decentralized options can be cost-effective in some countries such as countries in Africa due to their population distribution.

Motivated by emissions impacts, increasing fossil fuel prices and importance of energy independence, Hoesen and Letendre [24] proposed a GIS based model for the rural electrification which offers an eco-friendlier alternative to the current United States central distribution network by supporting biomass, wind, and solar resources. In order to evaluate the regional potentials of the selected areas, GIS-based analyses were conducted which allow digitally visualised expressions of spatial specifications of the regions.

Szabó et al. introduced a spatial grid cost model that evaluates the diesel generators, photo-voltaic systems or grid extension [25]. In this study, the simple cash flow model which provides electricity cost information for each mini-grid option was used to guide policy-makers while making spatial electrification planning. In 2013, a new

research on spatial-economic analysis that determines the least cost rural electrification in Sub Saharan Africa was conducted [26]. In this study, the important “unique opportunity” of using the new innovative and eco-friendly tools without necessarily using the old centralized methods that have high infrastructural costs was underlined.

Underlining the importance of reducing the energy poverty in the world in a cost-efficient way, Urpelainen made a policy analysis of coordination for grid extension and off-grid electrification[27]. The cost of mini hydro, off-grid PV and diesel generators options were compared with cost of the electricity grid extension in order to find the least cost option for every location. As a result, they came up with the idea that these two methods complement each other.

Zeyringer et al. analyzed the cost-effective electrification solution for Kenya comparing grid extension with stand-alone PV systems [28]. In this research, firstly they estimated the electricity demand for non-electrified households using an exponential regression model. As a next step, they used the results obtained from the regression model as an input to their supply-sized optimization model. By the help of this model they determined the least cost option by evaluating the cost of extension of the electricity grid and the implementing stand-alone PV for every demand node. As a result, they concluded the fact that stand-alone option is cheaper for remote and low-demand areas. Therefore, this study underlines the importance of regional based electrification planning.

In [29], a hybrid system optimization done by HOMER (Hybrid Optimization Model for Electric Renewables) which is based on both conventional and renewable energy such as biomass gasifier, photo-voltaic(PV), diesel generator and grid is introduced. After determining a break-even distance limit, the comparison between grid extension and off-grid system was performed. As a result, the optimization results conclude that grid extension is preferable for the villages within the distance limit. Contrarily for the remote ones, it is possible to provide energy with hybrid energy systems. In addition to that, it was shown that biomass gasification option dominates

the PV system for all scenarios.

Ruijven et al. made a model based analysis of global rural electrification scenarios that makes the assessment of future trends in electrification and investment plans required [30]. They used a two-step modeling approach: the first step makes a econometric analysis to set the electrification rates and the second step calculates the cost of electrification. In that research, it is underlined that for Sub-Saharan Africa, the potential in mini-grid and off-grid technologies is expected to be high since the grid electrification will have high infrastructural costs due to low population densities in rural Africa.

There are some studies focusing on particularly decentralized energy systems, grid connected centralized systems or rural electrification infrastructural design for energy planning. Hiremath et al. published a review on decentralized energy systems planning including both modeling approaches and their applications [7]. Kaundinya et al. made a literature search on decentralized energy systems in which they underlined the difference between grid-connected and isolated decentralized systems [11]. In this paper, they made a review of 102 articles that study decentralized power options, modeling approaches and analysis of economic, environmental and technological aspects of both grid-connected and stand-alone systems. There are various studies focusing on grid-connected [31, 32, 33], whereas isolated decentralized systems are [34, 35, 36, 37].

While evaluating cost of grid connected and isolated systems, geographical conditions such as spatial distribution of demand nodes are very important. These conditions both affect the cost of network and therefore the infrastructural design of the electricity. Zvoleff et al. developed a methodology that calculate an estimation of the cost for local-level distribution systems and some other cost parameters that will be useful for decision-makers such as the “marginal cost of connecting additional households to a grid as a function of the penetration rate” [38]. This method takes into account geographical patterns of the location while calculating the least-cost

network. In [39] a greedy heuristic algorithm that can be used as a “pre-feasibility design tool” was introduced. This algorithm provides a two level (MV, LV) grid network for green-field settings. This algorithm does not require candidate locations of transformers since it assigns place to them simultaneously.

In this research, we focus on a similar problem studied by Parshall et al., in which a heuristic solution approach is presented for guiding the electrification projects of regions with low electricity coverage [40]. This approach is then updated to be used also for the regions which do not have any existing electricity coverage and used as the basis of a decision support mechanism to explore grid and off-grid electrification options in rural communities. This open-source mechanism is called “Network Planner” and can be reached through <http://networkplanner.modilabs.org>. This mechanism is capable of testing different scenarios, demonstrating performance comparisons and maps. Network Planner is used in national electrification studies of countries such as Senegal [41] and Ghana [42]. The underlying solution approach in Network Planner is an adaptive and fast heuristic approach which can provide solutions in a short amount of time; however, its optimality performance (i.e. the closeness of the results to the optimal solution) is unknown. Abdul-Salam et al. [43] presented a non-linear discrete modeling approach for solving this problem optimally and compared the results with [40] for a maximum of 40-node instance. In this study, we approach to this problem as a Prize Collecting Steiner Tree problem for the first time in the literature and propose a mathematical programming model of the problem. Our approach can provide the optimal solution for large size problems.

The outcome of this research is expected to be an important decision support tool for the electrification of underdeveloped and developing countries, having the potential of contributing to the socio-economic development of these countries. This study is expected to make a significant contribution to the literature as it includes both theoretical modeling that can solve the problem optimally for large instances and applications using real data for a very important social problem.

2.2 Cost and CO₂ Minimizing Rural Electrification Planning Problems

Multi-objective optimization methods in energy systems have gained importance especially in recent years due to the global ecological harm caused by carbon emission. Those techniques are gaining popularity since they let researchers coalesce conflicting objectives such as environmental aspects and economical conditions. Generally, the energy applications of multi-objective optimization methods are used in the design of hybrid systems for solar, wind, diesel with energy storage of various types and different combinations of them for a specific settlement unit. Although the number of studies focusing on multi-objective optimization applications in energy systems are relatively high, surprisingly there are only limited number of studies about applications on network design in the literature. In that section, some examples of multi-objective optimization methods in energy systems will be presented. Afterwards, we will introduce applications of multi-objective optimization methods of network design, which narrows the scope further.

Pohekar and Ramachandran presented a review paper on multi-criteria decision making for sustainable energy system's planning [44]. In this paper, a classification and a detailed explanation of 90 published papers was presented. The methods based on "weighted averages, priority setting, outranking, fuzzy principles and their combinations" are reviewed and discussed. In the paper, the reasons behind the increasing usage in multi-objective decision making tools in sustainable energy planning is also explained. In our introduction part, we have explained the ecological reasons behind the sustainable development need; however, usage of renewable energy can also be considered as a business need. In [44], it is explained that the oil shock of 1973 thought a lesson to the world that it is essential to find ways to store energy and substitute in case of shortage. After this date, motivated by the need in finding alternative sources that can replace conventional fuel usage, planners and policy-makers got attracted by the unused potential of renewable energy sources. Aiming to find the

conditions preventing the penetration of renewables and to develop plans for overcoming those obstacles, energy planners started to use decision sciences in energy field and multi-objective linear programming is of course, one of them.

In [45], a review about the use of multi-optimization algorithms for hybrid renewable energy system design was presented. The research materials reviewed in this article contain multi-optimization algorithms for renewable and sustainable energy aspects such as “placement, sizing, design, planning and controlling”. In this paper, it also underlined that despite the large number of optimization methods for sustainable energy development, heuristic algorithms developed for multi-objective stand alone systems are very limited.

In [46], Abbas et al. introduced a triple multi-objective optimization approach applied to a hybrid PV-wind-battery system for a residential house. This approach considers both the life cycle cost, embodied energy and loss of power supply probability. In order to find the best compromise, a controlled elitist genetic algorithm has been implemented.

A study was conducted by Ren et al. about the distributed energy systems using resources such as photo-voltaic (PV), fuel cell and gas engine using a multi-objective optimization model [47]. This model has two objectives: the minimization of energy cost and the minimization of environmental impact which is determined by CO₂ emissions’ level. This method was implemented to a real life case in Japan which is a eco-campus selected as an example for this study. By the help of this example, the trade-off between economic objective and the carbon emission objective is demonstrated.

The following research is focusing on the multi-objective optimization of a stand-alone PV, wind and diesel systems with batteries storage using The Strength Pareto Evolutionary Algorithm [48]. The method considers minimizing both levelized cost of energy and the equivalent carbon dioxide and life cycle emissions. As a result of this study, taking into account both environmental and economic aspects, it is shown that

the photo-voltaic (PV) panels play a very important role in electrification in Europe since in some of the best solutions from the Pareto front, PV solar panels can be used alone.

Aiming to minimize the total cost of the system by determining the size of hybrid renewable energy system, Sharafi and ELMekkawy proposed an optimization-simulation based approach [49]. The system is searching to find the solution that minimize the cost and CO₂ emission while maximizing reliability of the system. This has been achieved through a mixture of the ϵ -constraint approach with PSO method. As a result, a new method that allows the decision makers to balance between criteria such as satisfying demand, reducing air pollution caused by burning fossil fuels and minimizing the total cost of the system was proposed.

As it has been already explained at the beginning of this section, there are just a few number of studies dealing with the network design with multi-criteria optimization methods. In [50], an urban road network design problem with two objectives was introduced. The objectives of this novel problem are maximization of consumer surplus, and maximization of the demand share of the bus mode. In this paper, various methods are proposed to solve this problem: a hybrid of genetic algorithm and simulated annealing, a hybrid of particle swarm optimization and simulated annealing, and a hybrid of harmony search and simulated annealing. By the help of the numerical experiments, proposed methods were investigated.

A supply chain network design problem for forest biomass was investigated in [51]. After detecting the problem of energy wood accessibility caused by rapidly increasing demand and high supply costs, a multi-criteria optimization model including decisions such as chipping location, transport mode and volume and terminals used was proposed. This multi-criteria optimization problem was solved by the weighted sum scalarization approach. In this approach, in each step by alternating weights, the Pareto optimal solution is obtained. In this paper, they also made the analysis of how much profit maximization affected increase in CO₂ emission and transport

distance.

Wang et al. studied a multi-objective optimization problem for green supply chain network design that incorporates two conflicting objectives: total cost and environmental [52]. This method provides a portfolio of different alternatives with changing objectives that can guide supply chain planners in big firms. The strategic insights obtained from the Pareto optimal set ameliorated by a sensitivity analysis. This research reveals the fact that it is more effective to focus on lowering CO₂ emissions for larger supply chain networks with higher demand capacities.

Farmani et al. used three different multi-objective evolutionary algorithms for the water distribution systems network design [53]. The multi-objective optimization method provided in that research captures the trade-off between the cost and the maximum head deficiency for each network. This can be considered as a comparative study of different approaches since they were examined by the help of the two case studies and a comparison is presented by using two performance indicators.

To sum up; the network design problems that evaluate the centralized and decentralized options have never been solved by a method giving the optimal solution similar to ours. To our knowledge, this kind of problems are generally solved with heuristic methods. Therefore, this thesis is aiming to provide a novel solution method that can be used as a reference tool for existing methods. Secondly, for the first time in the literature, the problem will be approached with a bi-objective method, denoting the changes in the system design taking into account different criteria. In the next chapter, the electricity network design problem that we study will be introduced in details.

Chapter 3

Problem Definition

In this study, we work on a spatial network planning problem which would be a base for the decision support tools for rural electrification. In this problem, two options are considered for the rural communities which may be of varied sizes, such as a country, region, or city. The first option is to design an isolated decentralized system that generates electricity using clean renewable energy sources. The other option is to connect the community to the grid where the electricity is generated in a centralized way and transmitted using transmission lines. Depending on the total electricity demand of the community, estimating the cost of a centralized system is more difficult than that of a decentralized option because it requires considering the spatial distribution of the communities and the optimal placement of the network infrastructure between them.

Three main cost aspects are considered for the decision-making process of rural electrification. The first two of these costs are related to grid option and are referred to as grid internal and grid external costs for the communities. The external costs are medium or high voltage line costs that connect the communities on the network to each other and to a central source. The internal costs consist of the transformers that convert medium voltage to low voltage and the cost of the low voltage lines that

provide distribution of the electricity from the transformers to the demand points such as households, hospital, school etc. While grid-connected systems have both internal and external costs, a system that is not connected to the grid, such as an isolated solar or wind power system, will include costs related to electricity generation and storage. This decentralized system can be called an off-grid system. The problem is to provide each community with electricity access in a way that the total cost of the overall system is minimized. This will be achieved either by connecting communities to the grid or by suggesting an isolated decentralized option such as a solar energy system for them.

If the grid internal cost within a community is higher than the cost of an isolated system, then it would never be cost effective for this community to be electrified with the centralized grid option. If the cost of isolated system is higher for a community, the difference between the cost of the isolated system and the grid internal cost determines the amount that this community can afford to be connected to the grid via transmission lines. Our solution approach aims to determine the communities that will have off-grid option, and the network topology between the communities that will be electrified with the grid option minimizing the cost of the entire system.

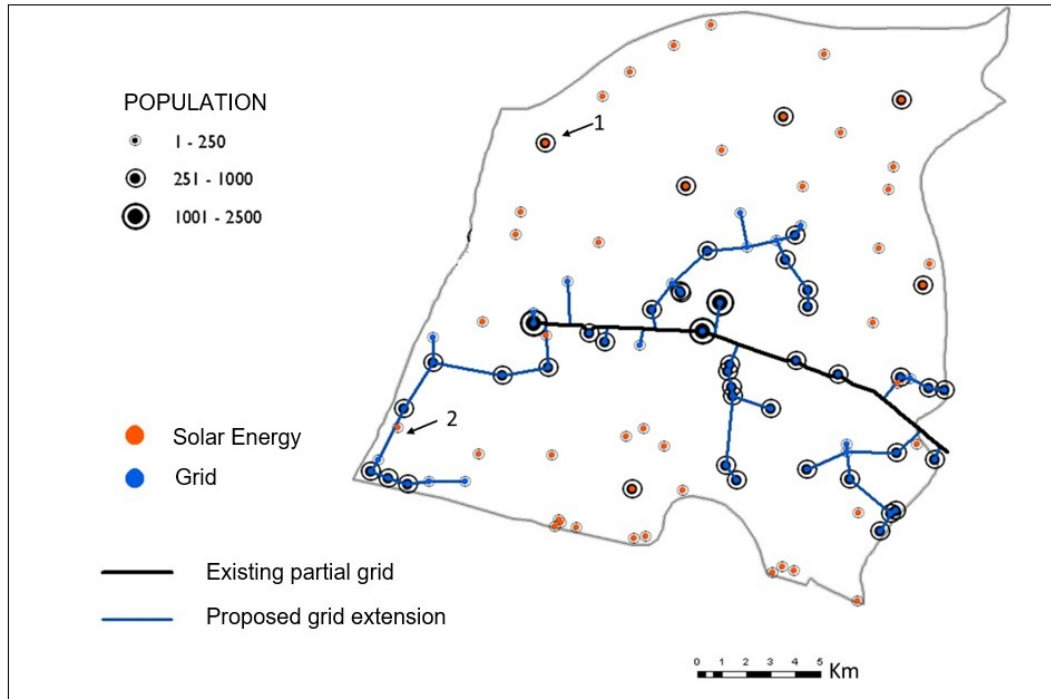


Figure 3.1: A Representative Example of a Result of the Electrification Project (Source: Sustainable Engineering Laboratory, Columbia University)

In Figure 3.1, the demand points with different populations of a sub-Saharan African settlement (Leona, Senegal) that has low electricity access are represented by different sized circles. There are two options for a demand point waiting for electricity: Either a partially existing network will be extended to cover this point, or an isolated energy system will be installed. Expansion of the network requires high infrastructure investments, especially for off-center settlements, due to mentioned internal and external costs. In addition to that, centrally produced electricity is cheaper due to the economies of scale. In solar or wind energy systems that are not connected to the grid, production costs are expected to be high in today's conditions and it can be very costly to provide electricity continuously in these systems especially for large scale demand points. For these reasons, it is more cost efficient to connect the large scale demand point to the network especially if there is an existing network nearby,

as seen in Figure 3.1, while the solar energy system can still be more cost effective for the demand points that are far away from the network. For the demand point marked with 1 in Figure 3.1, grid electrification is more expensive than installing solar energy system even though the demand is high. The demand point marked 2 is very close to the network but because the demand is very low, the solar energy system can be less costly than extending the network. In this sense, the problem aims to add the trade-off between the electricity generation and transmission costs to the account and aims to find the following elements:

- The demand points that will be connected to the central network(grid-compatible nodes), and the ones that will be isolated.
- The network topology of the network between the points to be connected to the network.

The example in the Figure 3.1 is an electrification example of a region with low-electricity coverage. This means that there is already an existing grid and the aim is to increase the level of access to electricity. Since the Network Planner can handle the existing grid network, it can be used for both regions without electricity and regions with low electricity coverage. However, in this thesis we focus on the problem with no existing grid and we design a system from scratch.

In the project, it is clear that the decisions made by system designers can change when considering the cost as well as the amount of emissions to be released from the system. With today's technology, isolated power systems operating with renewable resources cost more to install and although they produce clean energy, the production cost is higher. Electricity from the grid is usually produced at low cost with economies of scale from central and high capacity thermal power plants. However, it is produced using fossil fuels, it causes harmful gases such as CO_2 to be released. Thus, for rural electrification problems, more than one solution to system planners instead of a single solution is needed, which will allow the decisions to be made cost or just CO_2

emissions. In these cases, it is important to have the “most preferred” solution rather than the optimal solution [54]. In multi-objective decision making problems, the optimality concept is substituted with the concept of Pareto optimality or efficiency. Therefore, in the bi-objective problem, the CO₂ emission and total cost criterion will be the two objectives and a Pareto solution set for this bi-objective model will be found. In this thesis, we are seeking to obtain the Pareto solution set that can be considered as a catalog for the decision makers that guides them to determine the pay-off between conflicting objectives.



Chapter 4

Solution Approach

The heuristic approach used in Network Planner to solve the problem defined in the previous section is known as the modified Kruskal's algorithm. In this section, we firstly review the modified Kruskal's algorithm and present our solution approach to the problem that can provide optimal results in a reasonable time.

4.1 An Existing Approach: Modified Kruskal's Algorithm

As an input, Network Planner takes the data that are specific to the communities such as the cost, population, finance, average distance between houses etc. As a result, it calculates the total cost of the decentralized systems (off-grid) and the cost of the network for the centralized system (grid). Since the cost of the centralized option includes two parts (internal grid cost and external grid cost), selecting the communities which will be connected to the grid is not an easy task. If the cost of the decentralized option is lower than the internal grid cost (cost of low voltage

lines) for a community, then it would never be cost-effective for this community to be connected to the grid as the connection to the other grid-connected communities also requires external grid cost. Therefore, this community can be directly assigned to the off-grid option. However, if the internal grid cost is lower than the cost of decentralized option for a community, some external grid cost may be compensated with the difference. If the difference between the cost of the decentralized option and grid internal cost is sufficient for a community to be connected to other communities on the network via medium-voltage lines, then this community can be considered as a grid-compatible node, i.e. this node can be connected to the network. To define the maximum length of medium voltage line that can be afforded by each community, a metric called $MVmax$ is introduced in [40] as follows:

$$MVmax = (Cost\ of\ decentralized\ system - Internal\ grid\ cost) \\ / Unit\ cost\ of\ medium\ voltage\ (MV)\ line$$

The $MVmax$ values calculated for each community are given as inputs to the heuristic approach. The approach used in the Network Planner is modified to include the $MVmax$ into the Kruskal's algorithm [55], which optimally solves the Minimum Spanning Tree (MST) problem. In Minimum Spanning Tree problem, there is a given connected graph with positive edge weights and the aim is to find a minimum weight set of edges that connects all of the nodes. Kruskal's algorithm is a famous method that provides the optimal solution for MST problem, in which as a first step, distances between all points are sorted in an ascending order. The algorithm starts with an empty tree and repeats itself by adding the shortest available edge to the tree as long as it does not create a cycle until all of the points are spanned. In the modified version of Kruskal's algorithm, $MVmax$ criteria is added as an additional constraint. At each iteration, the $MVmax$ of the two points considered to be added to the tree is compared with the distance between the two points, and if the $MVmax$ of both points is sufficient to meet the distance between the points, the edge between these points is added to the network. Whenever a new connection is added, the $MVmax$ of the new set containing the connected i and j points is updated as follows:

$$MVmax_k = MVmax_i + MVmax_j - distance_{ij}$$

Modified Kruskal's (MK) algorithm explained above is a very fast solution method that gives practical solutions. However, it is a heuristic method without a proven optimality bound and therefore its reliability is questionable. Our new solution approach, which is explained in the next section, assures the optimal solution and therefore permits to evaluate the mathematical reliability of MK method.

4.2 Proposed Approach: Prize Collecting Steiner Tree

In previous sections, the rural electrification problem that aims to minimize the cost of the entire system by combining centralized and decentralized isolated systems is presented and the widely used modified Kruskal's Algorithm is explained in details. One of the most important contributions of our study to the literature is to define the mentioned rural electrification problem as a Prize Collecting Steiner Tree problem for the first time and provide solutions with known optimality performance.

In the general version of the Prize Collecting Steiner Tree problem, there is a prize to be earned specific to each node if the node is included in the tree. The goal in the problem is to collect the highest prize while reducing the cost of the tree that connects the nodes. In this sense, there is a trade-off in the problem. In the Geomans and Williamson (GW) version of the problem, a penalty is defined instead of a prize for each node [56]. If this point is not included in the tree, the penalty determined for the node should be added to the system cost. Therefore, the problem is to minimize the cost of the tree and the penalty of the ones that are not included in the tree in the GW version. Note that a Steiner tree may contain non-terminal nodes which do not have associated penalty/prize and these are referred to as Steiner points. In our problem, all the of nodes have non-negative penalties as there is the

cost of decentralized systems which will be paid unless the nodes are included in the tree.

In the general formulation of Prize Collecting Steiner Tree Problem, one is given a non-directional graph $G = (V, E)$, a non-negative edge cost $c(e)$ for each edge $e \in E$, a non-negative node prize/penalty $p(v)$ for each node $v \in V$. The Goemans-Williamson version of the PCST problem aims to find a subgraph $T' = (V', E')$ of G , $V' \subseteq V$, $E' \subseteq E$ that minimizes the objective function $c(T')$ which includes cost of the edges in the tree and the penalty of the nodes not in the tree as follows:

$$c(T') = \sum_{e \in E'} c(e) + \sum_{v \notin T'} p(v) \quad (4.1)$$

The mathematical model of Goemans and Williamson's Prize Collecting Steiner Tree problem is as follows [57]:

$$\min \sum_{e \in E} c(e)x_e + \sum_{v \in V} p(v)s_v \quad (4.2)$$

subject to

$$\sum_{e \in \delta(S)} x_e + \sum_{v: v \supseteq S} s_v \geq 1 \quad \forall S \in V; \quad (4.3)$$

$$x_e \in \{0, 1\} \quad \forall e \in E; \quad (4.4)$$

$$s_v \in \{0, 1\} \quad \forall v \in V; \quad (4.5)$$

In this formulation, x_e is a binary variable which indicates whether the connection between two node points is included in the tree. If the edge is included it takes the value 1, otherwise it takes 0. The s_v variable specifies whether any node subset is included in the tree. The variable takes the value 1 if the subset is decided not to be included in the tree and 0 if it is included. In the model, the set denoted by $\delta(S)$ is the a set of edges that have only one endpoint in S .

The above mathematical model is not a compact formulation as it is defined on subsets. Therefore, a flow-based mathematical model which is an alternative to the most general formulation above was implemented in our study. We model the PCST problem using single commodity flow formulation [58]. Following the same logic, non-negative edge cost corresponds to the cost of external grid system and the penalty corresponds to the difference between the decentralized system cost and internal grid cost of the communities. The parameters and decision variables used in the model are as follows:

Parameters :

- $CoffGrid_i$: The cost of the decentralized system to be installed at the point i
- $CgridInternal_i$: Cost of distribution network within point i (internal grid cost)
- $CgridExternal_{ij}$: Medium or high voltage line cost between points i and j (external grid cost)
- N : Number of demand points

Variables :

- x_i : 1 if point i is connected to the network, 0 otherwise
- s_i : 1 if point i is decentralized, 0 otherwise
- f_{ij} : flow amount flowing from point i to point j
- u_{ij} : 1 if point i is connected to the point j , 0 otherwise
- v_i : 1 if point i is selected as the source node, 0 otherwise
- ϕ_i : flow amount coming to the point i

$$\min \sum_{i=1}^N s_i CoffGrid_i + \sum_{i=1}^N x_i CgridInternal_i + \sum_{i=1}^N \sum_{j=1}^N u_{ij} CgridExternal_{ij} \quad (4.6)$$

subject to

$$x_i + s_i = 1 \quad \forall i \quad (4.7)$$

$$\sum_{i=1}^N v_i \leq 1 \quad (4.8)$$

$$Nv_i \geq \phi_i \quad \forall i \quad (4.9)$$

$$\sum_{j=1}^N f_{ij} + x_i = \sum_{l=1}^N f_{li} + \phi_i \quad \forall i \quad (4.10)$$

$$Nu_{li} \geq f_{lj} \quad \forall i, l \quad (4.11)$$

$$\sum_{l=1}^N u_{li} + v_i = x_i \quad \forall i \quad (4.12)$$

$$u_{ij} \leq x_i \quad \forall i, j \quad (4.13)$$

$$x_i \in \{0, 1\} \quad \forall i \quad (4.14)$$

$$s_i \in \{0, 1\} \quad \forall i \quad (4.15)$$

$$v_i \in \{0, 1\} \quad \forall i \quad (4.16)$$

$$u_{ij} \in \{0, 1\} \quad \forall i, j \quad (4.17)$$

$$f_{ij} \geq 0 \quad \forall i, j \quad (4.18)$$

$$\phi_i \geq 0 \quad \forall i \quad (4.19)$$

In the objective function of the above model, the total cost of the system is minimized. Constraint (4.7) guarantees for each community either the establishment of a decentralized system or the connection to the grid. Constraints (4.8) - (4.13) model the tree structure between network-connected units. In (4.8) it is assured that there will be only one source node. Constraint (4.9) indicates that the additive flow amount passing from any node cannot be greater than total number of nodes. The constraint (4.10) provides the flow between two consecutive nodes that will be included in the tree. It assures that if the node is included in the tree, the amount of the flow coming from its predecessor will increase by one. The flow passing from an edge cannot be greater than the total number of the nodes. This condition is satisfied by the constraint (4.11). The constraint (4.12) assures that if a node is selected as a source node, it cannot be linked to the tree with a predecessor node that is included in the tree. The constraint (4.13) guarantees that if a node is not included in the tree, its outgoing edges cannot also be included in the tree. Constraints (4.14) - (4.19) are constraints that define the decision variables and their bounds.

4.3 The Bi-Objective Model

The proposed solution for the bi-objective model is the ϵ -constraint method which is one of the most commonly used methods for bi-objective models [59]. This method solves a multi-objective model using single-objective subproblems. In each subproblem, one of the objectives is optimized while the other objective is controlled by a constraint. The general form of subproblems that is being involved in ϵ -constraint

method is as follows. In this model, x denotes the decision variable vector. S is the feasible set of solutions and C_k is the multi-dimensional objective function for some fixed k and varying ϵ_i values [60].

$$\begin{aligned} & \min C_k(x) \\ & \text{subject to} \\ & x \in S, \\ & C_i(k) \leq \epsilon_i, \forall i \neq k \end{aligned}$$

We define the first objective ($OF1$) as total cost and the second one ($OF2$) as carbon emission which can be found below. The second objective will be controlled by ϵ value that will change in each iteration. Suppose that ($OF2^*$) is the optimal solution for the second objective obtained from the last iteration. The ϵ value is defined as $OF2^* - \beta$. We solve the same model iteratively, each time tightening the constraint by the amount of predetermined step size β . The number of Pareto solutions found can be controlled by changing the step size. It is obvious that the number of steps will increase the size of the obtained Pareto solution set, but at the same time it will increase the solution time.

$$OF1 : \sum_{i=1}^N s_i C_{offGrid_i} + \sum_{i=1}^N x_i C_{gridInternal_i} + \sum_{i=1}^N \sum_{j=1}^N u_{ij} C_{gridExternal_{ij}}$$

$$OF2 : \sum_{i=1}^N x_i d_i$$

($OF2$) represents the amount of emissions released to the atmosphere under the assumption that the amount of emissions is directly proportional to the demand of the grid-connected settlements using fossil fuels. The new parameter d_i is the demand value of the nodes. The objective function 2 ($OF2$) minimizes the total demand that is met by fossil fuels and therefore the total amount of emission gases released in the system. The algorithm works as follows:

Step (1) Solve:

$$\begin{aligned} &\min OF1 \\ &\text{subject to} \\ &x \in S \end{aligned}$$

Suppose that the solution of the single objective minimization model of $(OF1)$ is found as $(OF1^*, OF2^*)$ in **Step (1)**.

Step (2) Solve:

$$\begin{aligned} &\min OF1 \\ &\text{subject to} \\ &x \in S \\ &OF2 \leq \epsilon \end{aligned}$$

If *infeasible*, the algorithm stops. Otherwise, supposing that the solution of the model are $(OF1^*, OF2^*)$, the algorithm sets $\epsilon = OF2^* - \beta$ and go back to **Step (1)**.

As we move from a more grid-connected network design we expect to have higher carbon emission values. In other words, this yields “a set of optimal trade-offs” in which one objective cannot be improved without deteriorating other [61]. Visualizing the trade-offs between different solutions, decision makers can easily see the impact of changes of the project budget in terms of carbon emission. This set of Pareto solutions can help the decision makers see how much extra money they should spend for a more eco-friendly solution. In rural electrification projects, cost is always an issue to be considered. Large amount of extra investment costs for environmental friendly energy generation may discourage policy makers. Therefore, the main contribution of this solution method will be detecting solutions that require only small increase in budget but yield an effective reduction in carbon emission.

Finally, it should be underlined that the solution to the problem would be a weakly Pareto efficient solution. This means that if there are more than one solution that is optimal with the same $(OF1)$ value, the one with the smallest $(OF2)$ value may

not be obtained from these solutions. In that case a better solution option could be missed. To prevent this, in each iteration, lexicographic optimization can be used. However, this doubles the number of problems to be solved in each step and increases the solution time. For our problem, we do not need to use that approach because we believe that it is a very low possibility to find different combinations of trees that has the same total cost but different emission values since each node has its own unique internal, external and off-grid cost data. Therefore we only use the classical ϵ -constraint method for our bi-objective model.



Chapter 5

Numerical Analysis

Modified Kruskal's algorithm has been widely used in national and local electrification projects for developing and underdeveloped countries to provide feasible results that offer the cost-effective balance between decentralized and centralized systems. In this heuristic approach global optimality is not guaranteed. Our new approach, on the other hand, can provide the optimal solution to the same problem. We use these optimal solutions of the PCST formulation as a benchmark to evaluate the performance of Modified Kruskal's algorithm, which we refer as MK in the rest of the thesis. In this section the computational experiments required for the comparative analysis of those two approaches are presented.

Firstly, we tested these two solution approaches by using synthetic data and secondly, we made further analysis by the help of real life instances obtained from Network Planner. We followed the same approach also for the bi-objective version, too. Experiments are performed on a dual core computer with Intel(R) Core(TM) i7-4510U CPU @ 2.00GHz, 2001 Mhz. Modified Kruskal's algorithm is implemented in MATLAB R2016a and PCST model is solved by CPLEX Studio IDE 12.6.

5.1 Experiments with Synthetic Data

We created test instances at different sizes to test the PCST approach and compare it with the MK approach. We generated our own sample data rather than using the ones from literature since our problem includes only terminal nodes and the existing data sets that are frequently used in the literature for PCST problems contain both non-terminal and terminal nodes. Both the distance and off-grid costs are generated randomly using uniform distribution. This analysis provides information about mathematical performance of MK. Moreover, experiments demonstrate the computational time required for the PCST approach. In Table 1, the size of the instances used in the computational experiments can be found in the first column. In the second and third columns, percentages of grid-compatible nodes obtained by MK and PCST methods are presented, respectively. A node is called grid-compatible if it is selected to be included in the tree for the least cost scenario. In the following six columns, decentralized system cost, centralized system cost and total cost of both methods are reported. In the following columns, the CPU time requirement of the both approaches and cost difference percentages are provided. Note that some of the instances did not give optimal results in our 3 hours time limit when we use PCST method. For these results, the optimality gap reported by CPLEX 12.6 can be found next to the result in the Total Cost column.

Table 5.1: Comparative Analysis of Two Methods Based on Randomly Generated Instances

Number of nodes	Grid percentage(%) ^a		Decentralized System Cost		Centralized System Cost		Total Cost		CPU Times (seconds)		Cost Diff.(%) ^b
	MK	PCST	MK	PCST	MK	PCST	MK	PCST	MK	PCST	
5	40	40	45	45	19.2	19.2	64.2	64.2	4.5	0.1	0
10	50	50	50	50	62.7	62.7	112.7	112.7	4	0.2	0
20	55	60	97	86	112.4	119.5	209.4	205.5	6.8	1.4	1.90
50	46	56	209.5	182.8	216.5	233.8	426	416.6	5.4	7.4	2.26
100	40	58	326	206	275.1	374	601.1	580	8.4	76.3	3.64
200 ^c	43	62	2836	1648	2949.7	3886	5785.7	5534 (0.95%)	13.6	10800	4.55
300	49	62.3	2538	1728	3576.7	4152.7	6114.7	5880.7	22.4	5135.2	3.98
500 ^c	44	60.6	3052	1972	3406.8	4233.1	6458.8	6205.1 (5.6%)	81.5	10800	4.09

^aGrid percentage = Number of grid-compatible nodes/Total Number of Nodes

^bCost Difference = (Total Cost of MK- Total Cost of PCST) * 100/ (Total Cost of PCST)

^cThe result is not optimal, percentages within the parenthesis show optimality gap reported by CPLEX. Time limit is 3 hours.

Given that PCST approach provides the optimal solution to the problem, results in Table 5.1 show that MK approach is practical and reliable for electrification projects since results are close to the optimal solution and CPU times are very small even for the large instances comparing to the PCST approach. It can also be concluded that PCST formulation can be preferred especially for small sized instances since there is no significant difference in terms of the solution time. It should also be underlined that, in this study we directly solve the model via CPLEX. However, alternative methods that solve PCST problem optimally and in a faster way can be found in the literature. Implementations of branch-and-cut algorithm [62, 63], using reduction techniques and valid inequalities [64] and relax-and-cut algorithm [65] are some of them.

5.2 Experiments with Case Studies from Network Planner

In this section, we test the performances of the two approaches using the real data obtained from the Network Planner tool. Note that Network Planner is an open source platform in which all the data required such as internal costs of centralized systems, cost of unit MV-line, off-grid, and mini-grid costs, coordinates of each node can be found. We have chosen six instances to make a comparative analysis. Diversity in terms of sizes and geographic features are our main criteria while choosing our samples. The results of the two methods can be found in Table 5.2. Similar to the Table 5.1, the size of the instances used in the computational experiments can be found in the first column. In the second and third columns, percentages of grid-compatible nodes obtained by MK and PCST methods are presented, respectively. In the following six columns, decentralized system cost, centralized system cost and total cost of both methods are reported. In the following columns, the CPU time requirement of the both approaches and cost difference percentages are provided. Note that for some instances, optimal results could not be obtained within our 3 hr solution time limit in our computational setting. For these results, the optimality gap reported by CPLEX is indicated next to the result in the Total Cost column.

Table 5.2: Comparative Analysis of Two Methods Based on Real Life Instance Results

Number of nodes	Grid percentage(%) ^a		Decentralized System Cost		Centralized System Cost		Total Cost		CPU Times (seconds)		Cost Diff.(%) ^b
	MK	PCST	MK	PCST	MK	PCST	MK	PCST	MK	PCST	
21	90.5	90.5	677,478.4	677,478.4	4,205,351.7	4,205,351.7	4,882,830.1	4,882,830.1	6.8	6.1	0
94 ^c	23.4	70.2	2,247,648.9	760,542.6	2,118,952.7	3,473,335.8	4,366,601.6	4,233,878.4 (4.39%)	9.3	10800	3.14
141	7.1	14.2	7,004,088.9	5,824,077.2	32,405,490.2	33,502,097.7	39,409,579.1	39,326,169.6	10.5	2578.8	0.21
102	43	58	3,549,189	2,322,924	3,339,708	4,104,273	6,888,897	6,427,197	17.0	91.83	7.18
230 ^c	91.3	100	2,418,714.9	0	62,920,938	65,316,843.7	65,339,652.9	65,316,843.7 (30.81%)	21.2	10800	0.04
274 ^c	84.7	89.1	1,398,680.7	1,087,633	68,124,740.9	68,463,021.1	69,523,421.6	69,550,654.1 (4.96%)	19.6	10800	-0.04

^aGrid percentage = Number of grid-compatible nodes/Total Number of Nodes

^bCost Difference = (Total Cost of MK- Total Cost of PCST) * 100/ (Total Cost of PCST)

^cThe result of PCST is not optimal, percentages under the results show optimality gap reported by CPLEX. Time limit is 3 hours.

Table 5.2 shows that the optimal solution of the PCST problem for some instances could not be obtained within our 3 hr time limit given our computational environment. However, we observe that the results with positive optimality gaps are still better than the result of MK approach except for one instance. In that instance, the MK approach provides 0.04% better result than PCST result with 4.96% optimality gap. The optimal results of the instances resulted with an optimality gap in our time limit and computational environment can be easily obtained in better computational settings or without a time limit. We observe that MK approach provides solution for all instances in units of seconds. However, the solution quality of these results are unknown unless we compare these results with the results of PCST approach. On the other hand, PCST approach provides results in units of hours and the solution time highly differs based on grid percentage in the final solution. Since the tree topology between the grid-compatible nodes makes the problem very complicated, the solution time of the model is also affected by the grid-compatible nodes. For example, the optimal result for 94-node instance is not obtained within 3 hr time limit, whereas the optimal result for 141-node example is obtained. As the number of grid-connected

nodes is higher in the 94-node example, the solution process includes more decisions about the connections made between these nodes.

Another interesting results that Tables 5.1 and 5.2 show is that MK approach is more conservative in terms of the number of grid connections, i.e. grid percentage of the MK approach is always less than or equal to the PCST approach. To examine the reason behind the potential differences between the two methods in detail, we have chosen an instance and investigated the steps of MK approach iteratively to detect, where in contrast to the PCST solution, the MK chooses the decentralized option. In other words, this analysis aims to explain the behavioral tendencies of MK approach that makes it result with different solutions than PCST. We worked on 102-node example to investigate the reasons behind differences in results. The final trees obtained by both methods are different as presented in Figure 5.1. In Figure 5.2, we zoom into the right-bottom corner of the 102-node example. Remember that MK approach, first sorts the edges of the given graph in an ascending order of their lengths. Then, each edge is considered to be added to the tree by checking if the new edge creates a cycle and if the $MVmax$ of the nodes at the end points of the edge exceeds the length of the edge. In this example, the distance between node 1-2 is smaller than 1-4 with values 1398.8 and 2178.8 meters, respectively. However, in Figure 5.2 it can be observed that MK approach chooses to connect 1-4 instead of 1-2 because when edge 1-2 was considered, the length of the edge 1-2 was greater than $MVmax$ of node 2. Recalling the fact that $MVmax$ is a metric that is getting larger proportionally to the number of nodes included in the tree, $MVmax$ of node 2 could have been larger if it had made another connection before. Likewise, initially the $MVmax$ of node 4 was 1670, however it reached the value of 3926 by getting connected with nodes 5 and 3 before the algorithm checked its connection with node 1. The distance between node 1-4 is 2179.8 meters and it was impossible for node 4 to be connected to 1 if they got checked before node 4 made other connections. This example clearly shows us how MK approach shows favor to the locations that has close neighbors in terms of building connections.

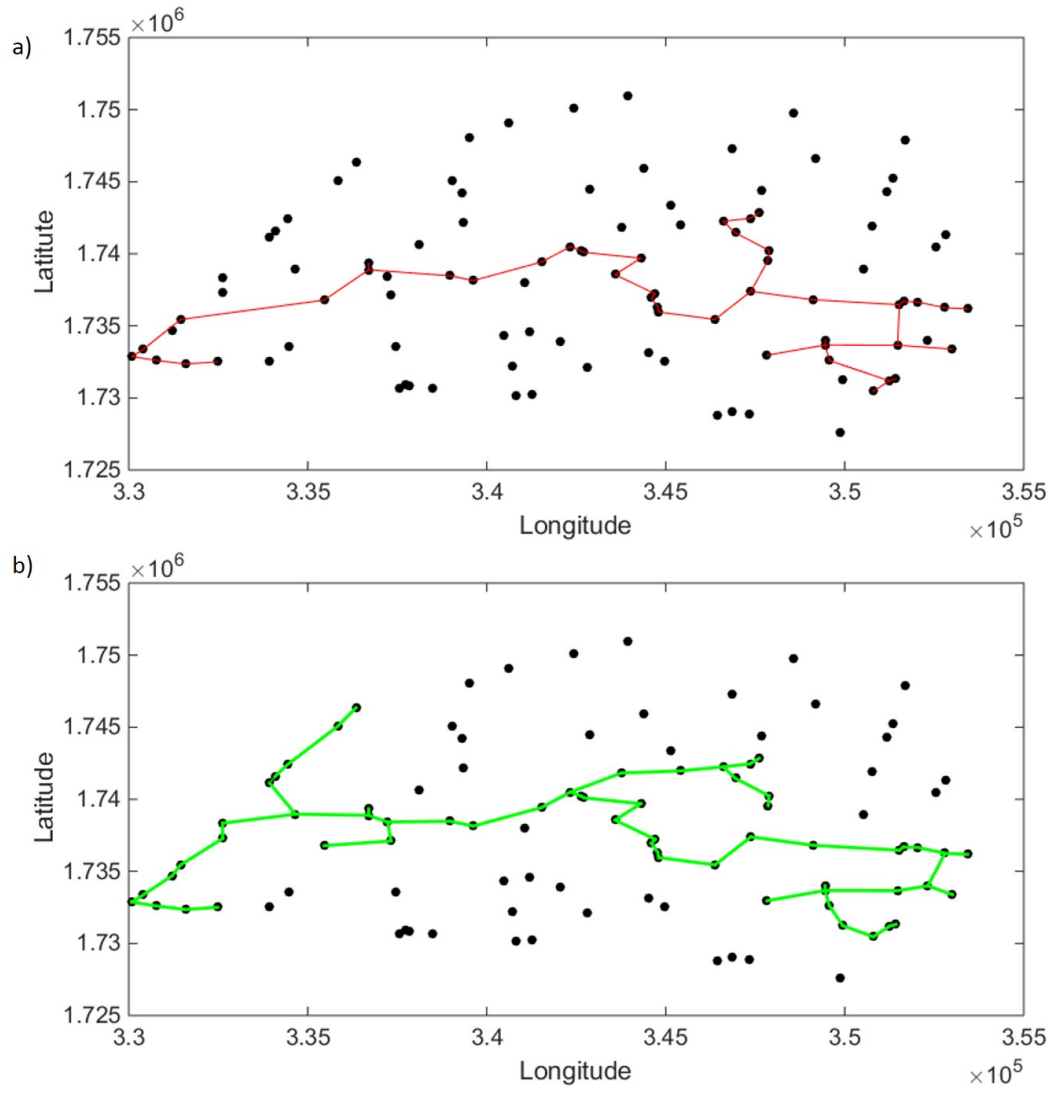


Figure 5.1: Final Grid obtained by PCST method of 102-node instance a)with MK, b)with PCST

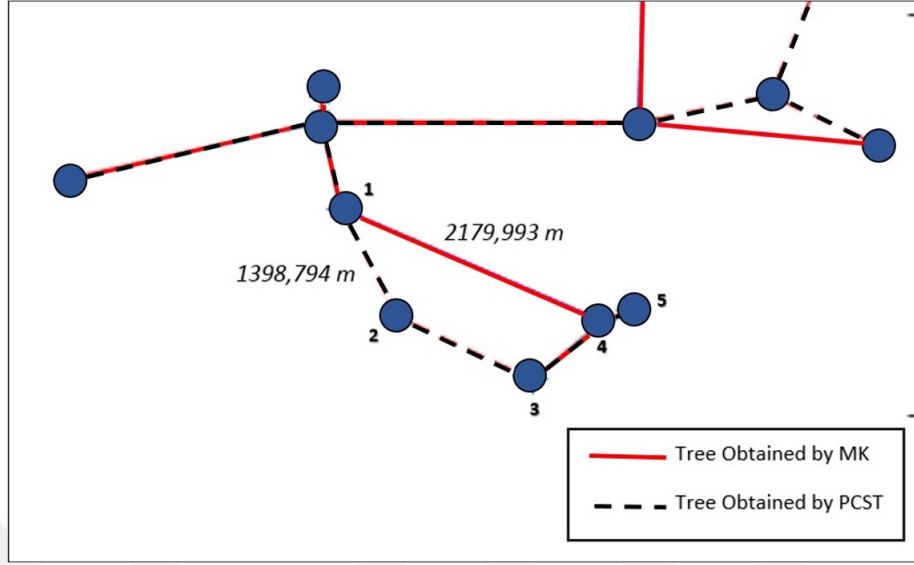


Figure 5.2: Difference between the tails of two trees obtained by MK and PCST approached on 102-node example

5.3 Sensitivity Analysis on Distance Parameters

We conducted a sensitivity analysis to observe how results alter with variations of distances (or similarly prices on external network costs). To assess the potential impact of variations on total cost and grid percentage, we created a 100-node test instance and its reduced forms as shown in Figure 5.3. The results of both MK and PCST methods are summarized in Table 5.3.

Table 5.3: Comparative analysis of two methods with respect to distance variations

	MK:Cost	PCST: Cost	MK : Decrease in cost (%)	PCST: Decrease in cost (%)	Cost Diff.(%)	MK : Grid percentage	PCST : Grid percentage
Base Scenario	601.103	580			3.64	40 %	58 %
Reduce distances by 25%	472.825	465.36	21.340	19.766	1.60	74%	80 %
Reduce distances by 50%	328.138	320.90	30.601	31.042	2.26	83%	92 %
Reduce distances by 75%	164.578	164.11	49.845	48.859	0.29	95%	96 %

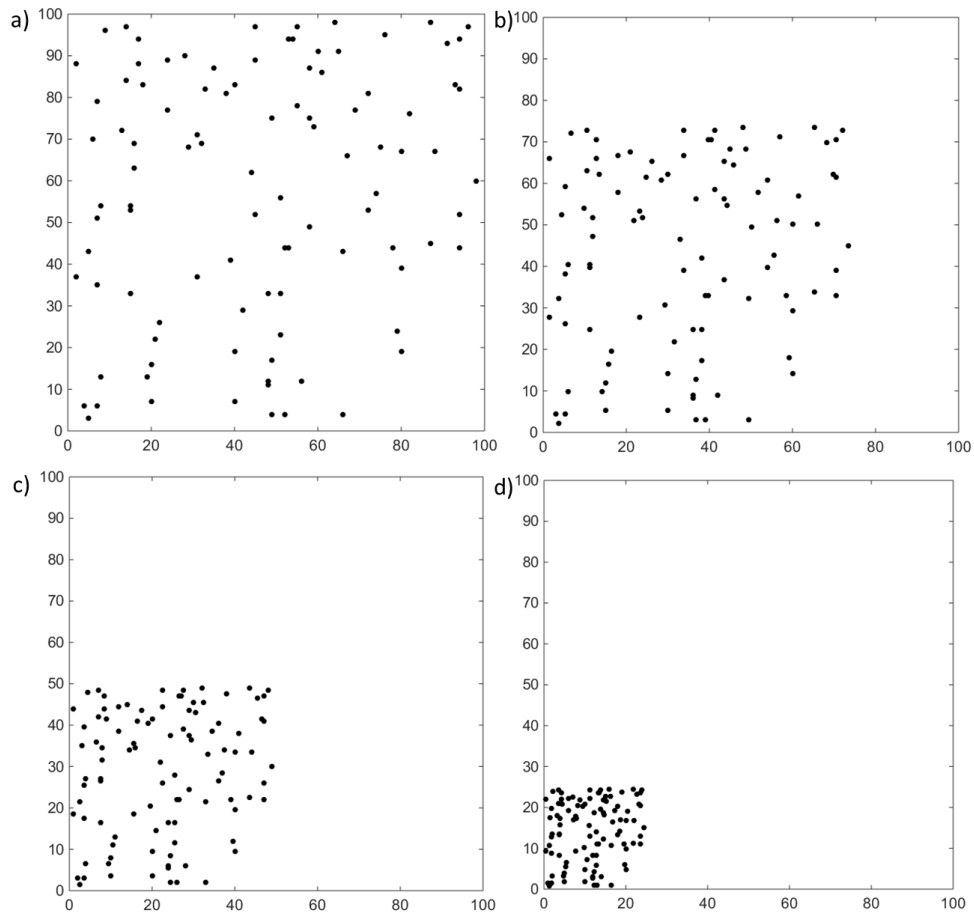


Figure 5.3: a) Randomly generated 100 noded sample(Base Scenario), and its reduced forms by b) 75%, c) 50% and d) 25%, respectively

In Table 5.3, in the first two columns, total costs obtained by MK and PCST approaches are presented. In the following two columns, how much the total costs decrease with respect to distance reduction is presented. The percentage based cost comparison of MK and PCST approaches can be found under the cost difference column. Finally, the grid percentages of the solution trees of both methods are presented in the last two columns. As it can be expected, when the distances between the communities are reduced (i.e. communities are clustered), the total costs of the final systems decrease and the number of nodes connected to grid increases for both methods as the $MVmax$ values of the communities will be more than the lengths of the higher number of edges. Similar to previous results, MK approach tends to connect less communities to the grid than PCST approach. However, as the clustering of the communities increase, the difference between the solutions of the both approaches reduce. In the 100% grid connection percentage case, both methods are supposed to find a minimum spanning tree between the communities and this tree will be the optimal tree which can span all points. Therefore, it is expected that MK approach will provide solutions which are close to the optimal result for the cases when grid connection percentages are high.

Our results show that PCST model may be used for electricity planning problem for large scaled unelectrified regions including up to 500 given our computational setting. However, the number of nodes is not the only factor that determines the level of computational complexity and CPU time. Costs and the level of dispersion of the communities are some other factors contributing to them. We previously showed in Table 5.1 that PCST method solved the instance with 300-node instance optimally but resulted with an optimality gap for 200-node instance within 3 hr time limit. Similarly in Table 5.2, it is shown that for the PCST model is solved optimally for 141-node instance but resulted with an optimality gap for other 94-node instance within 3 hrs. Analyzing the spatial distribution of the communities can give us an idea about how the level of clustering effects the computational complexity.

5.4 Computational Analysis for the Bi-Objective Problem

In the computational study, we used several instances from the ones created as test instances for the PCST approach, explained in the previous section. As we have already mentioned, both the line and off-grid costs are generated randomly using uniform distribution. For the bi-objective problem, we picked the ones with size of 10, 20, 50 and 100 nodes from the synthetic data and as a real life case, the data with 141 nodes is chosen.

In Figures 5.4, 5.5, 5.6 and 5.7 you can find the Pareto solution set obtained by using synthetic data with node sizes equal to 10, 20, 50 and 100, respectively. Note that, in all instances, demand sizes are taken as 1. In other words, the demand sizes of all of the nodes are considered as identical. Therefore for these four examples, we can only see the results of ϵ -constraint method for regions that has identical demands. However, sometimes, for example in our real life example from Sub-Saharan Africa, this is not the case. You can find the Pareto solution set of real life example in Figure 5.9. In this example, the demand of the nodes are not identical. Therefore, we have the opportunity to see how the varied demand sizes of the communities affect the Pareto solutions that we obtain.

In ϵ -constraint method, predetermined step size defines the computational length of algorithm. If this value is high, there will be more iterations and therefore it requires longer time. However, that way, the risk of missing a Pareto optimal solution will be reduced. For the synthetic data, we took 1 as the step size. Therefore, all Pareto optimal solutions are found for the examples of synthetic data. For the real life data, we set the step size to the average value of demands. This is equal to 130546.70.

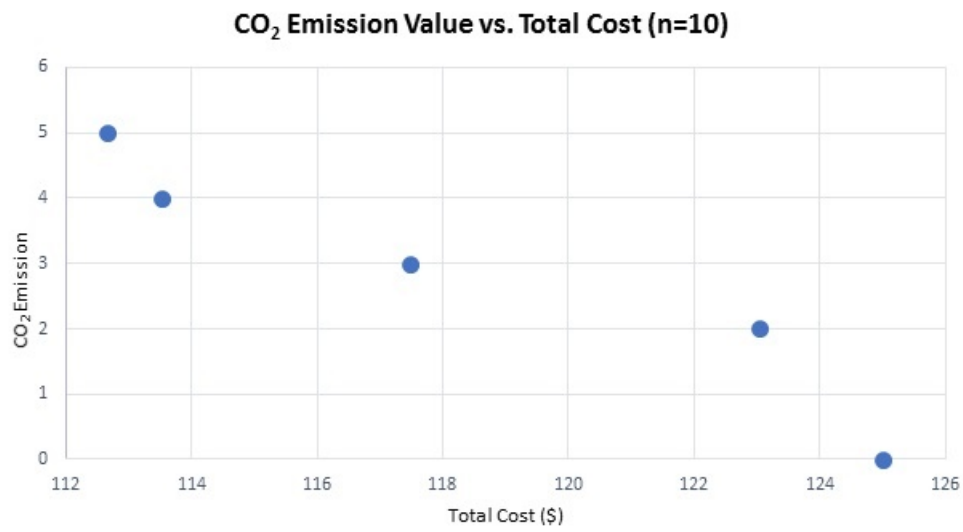


Figure 5.4: Pareto Solution Set of 10-node Synthetic Data

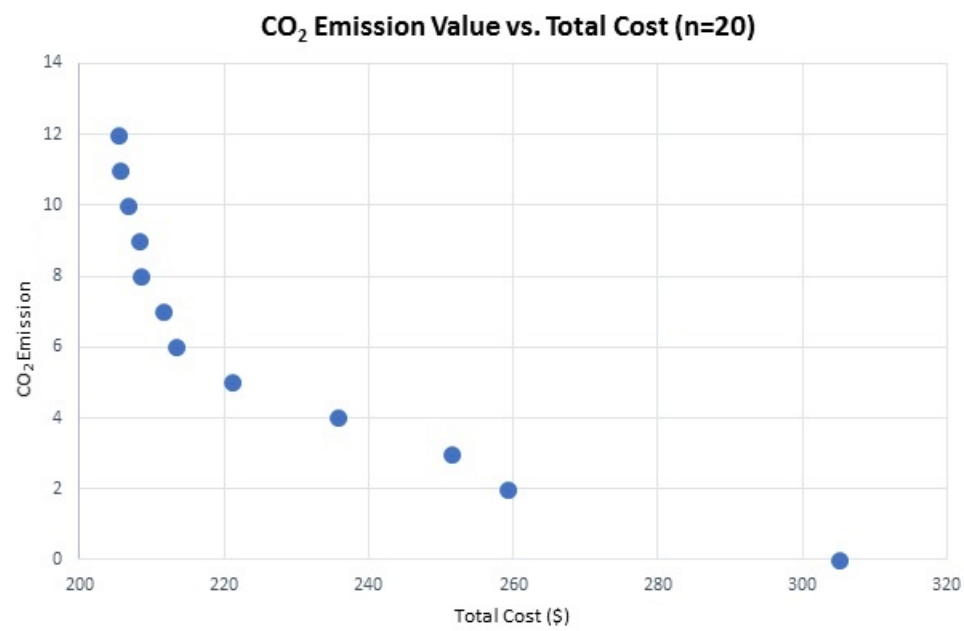


Figure 5.5: Pareto Solution Set of 20-node Synthetic Data

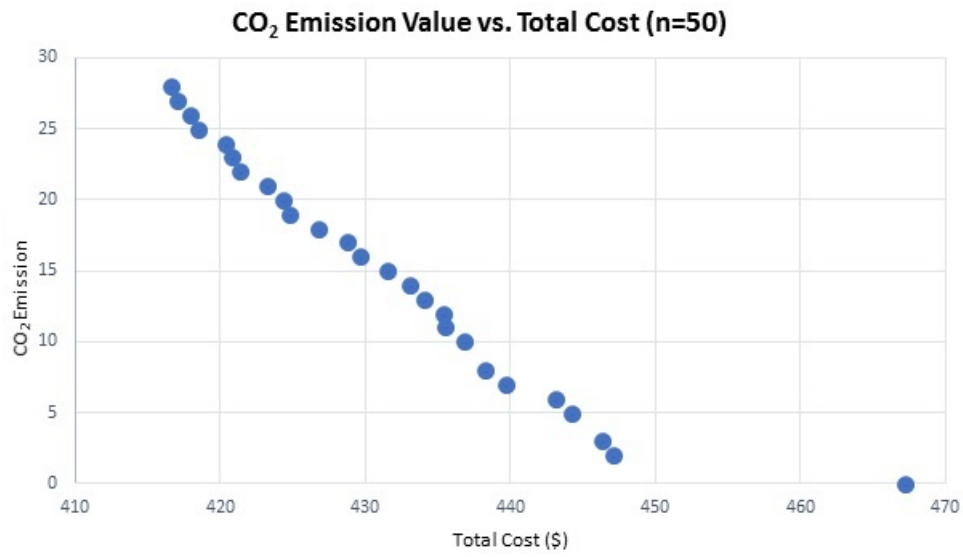


Figure 5.6: Pareto Solution Set of 50-node Synthetic Data

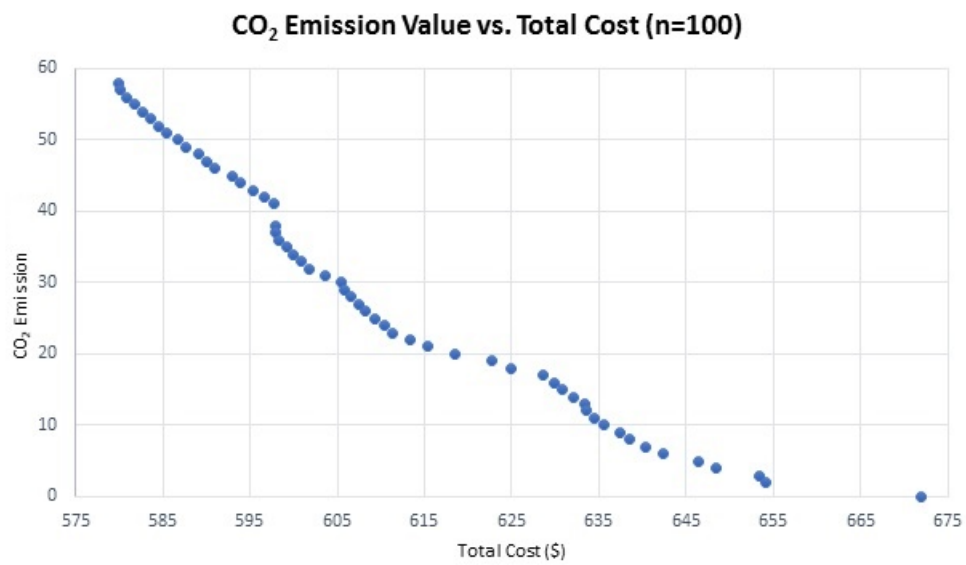


Figure 5.7: Pareto Solution Set of 100-node Synthetic Data

Figures above show us that even if the decision maker is not willing to increase the budget of the project too much for the environmental concerns, options with lower carbon emissions are also available in expense of small increases in the cost. When we look at Figure 5.4, we observe that if the decision maker agrees to pay only 1.8 % more, the carbon emission value can be decreased by 20%. If the decision maker wants a more balanced result in terms of both objective functions, only the cost will increase by 5%, to decrease the carbon emission by 25%. In the Figure 5.5, we can find a very good example of how a very small amount of change in cost can provide a huge difference in carbon emission. If the decision maker accepts to give only 7.7 % percent more money than he will pay for cost minimized version, he carbon emission can increase by 58.3%. In Figures 5.6 and 5.7, we observe a more linear Pareto front comparing to Figure 5.5. However, we can still say that, without giving too much sacrifices from the budget, reduction in carbon emission is possible.

Policy makers, governments and organizations should take into account the environmental issues in their decision making process and our experiments show that it is possible to provide it with small cost sacrifices, at least for the identical demand distributions. We should underline that demand parameter plays an important role in determination of the Pareto solution set. In some cases, the pay-off can be larger in a way that the decision maker is not willing to pay. However, our results obtained with synthetic data show us that it is worth to evaluate the bi-objective version of the problem since it is possible to reduce the environmental side effects of electrification with relatively small budgets.

We can now move to the real case data in which we can clearly observe the impacts of demand difference. In Figure 5.9, the Pareto solution set obtained by using ϵ -constraint method can be found. In that example the real demand data is used. In the data, there is one node that has a very high demand value compared to the other nodes with the value of 16,149,914.12. This node will be referred as extreme node for the rest of the paper. The others are changing between 0 and 208,872.25. The histogram of the demand data of the 141-node example can be found in Figure 5.8.

Note that in we did not put the extreme node in the histogram to be able to narrow down the horizontal interval. Otherwise, it is not possible to see the distribution of the demand sizes of the nodes clearly.

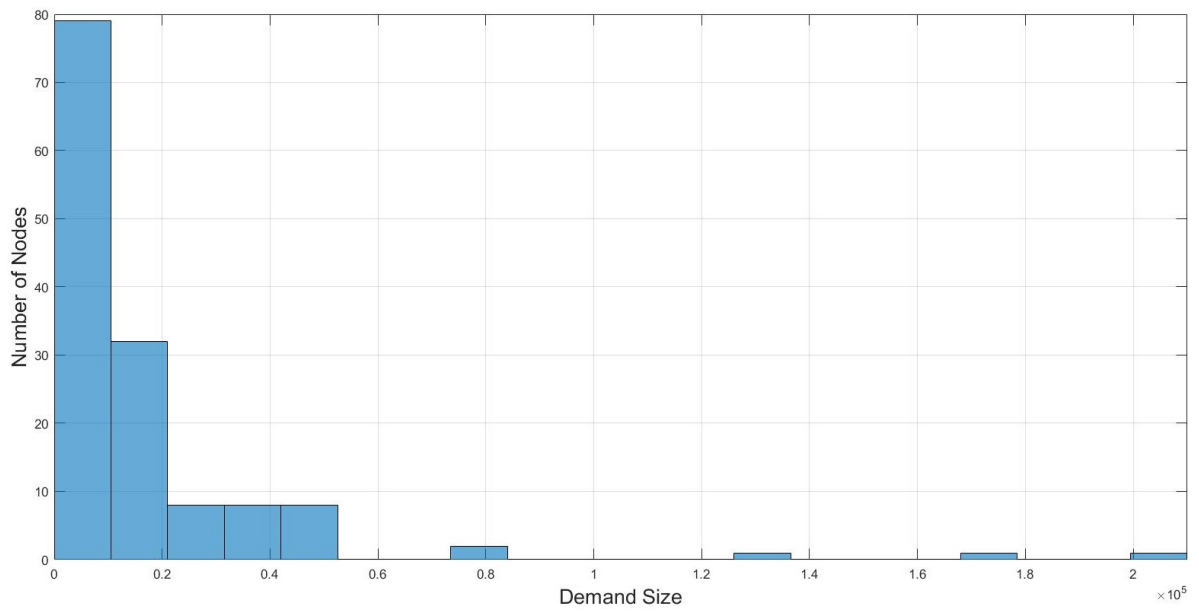


Figure 5.8: Histogram of 141-node Real Demand Data

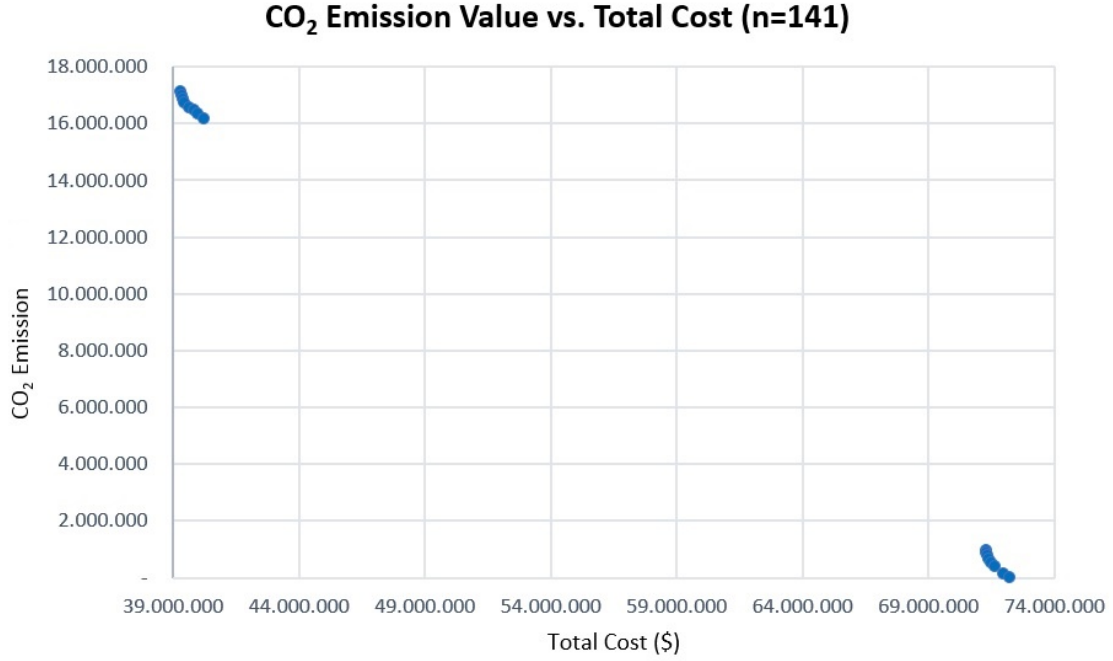


Figure 5.9: Pareto Solution Set of 141 noded Real Life Instance

In the Pareto solution set presented in Figure 5.9, we find two main groups that has close objective values within itself, but far from the other group. In this example, the decision maker should first decide on which objective is more important for the project and than he should think about how much pay-off he can give between different options. In other words, if the decision maker decides to satisfy the majority of the demand with grid electrification, then he should focus solutions on the top left corner. However it should also be underlined that, just like in the synthetic data examples, it is possible to reduce the carbon emission with little shifts in the cost. In this case, by only paying 2.3% more, a carbon emission reduction by 5.5% can be achieved.

When we analyzed results obtained in each iteration, we found out that the large gap between these two groups is caused by node that has the extremely large demand value of 16,149,914.12. In Table A.6 which can be found in the Appendix, the results

obtained in each iteration of the ϵ -constraint method for real example can be found. It can be observed that after hitting 16,185,117.15, there is a very significant decrease in carbon emission. This is because, in that step, for the first time ϵ reaches a value smaller than 16,149,914.12. The left hand-side of the ϵ -constraint is 16,080,170.41 and therefore it does not let the extreme node be involved in the tree. It is the game-changing step for the whole network design. After this connection is removed, a new network with 17 new grid connections has been built. In Figures 5.10 and 5.11, the proposed networks obtained before and after this iteration can be found. It is not surprising to see that, the node with the highest demand value has also the highest $MVmax$ value. This explains the reason why the algorithm pushes its limits for including this node in the grid network: Having the largest $MVmax$ value points out that the difference between off-grid cost and internal cost of this extreme node is larger than any node. Therefore, instead of paying the off-grid cost of this node, it is not hard to find a cheaper option for the model, in which a large amount of grid extension can be afforded. The amount is so large that the solution of the model included the extreme node in the tree until the ϵ -constraint forced to omit it. This extreme case shows us clearly how much differences between demand sizes can affect the Pareto solution set.

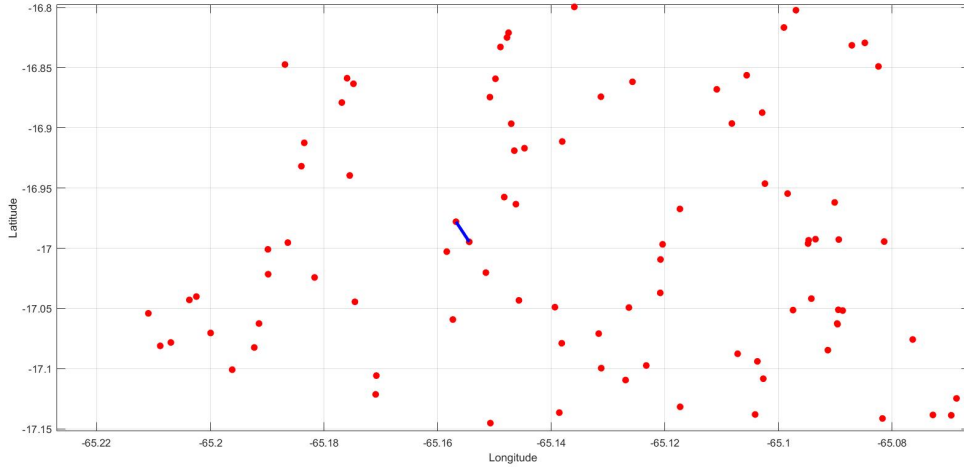


Figure 5.10: Network Result Obtained with ϵ -Constraint Method One Iteration Before the Removal of the Extreme Node

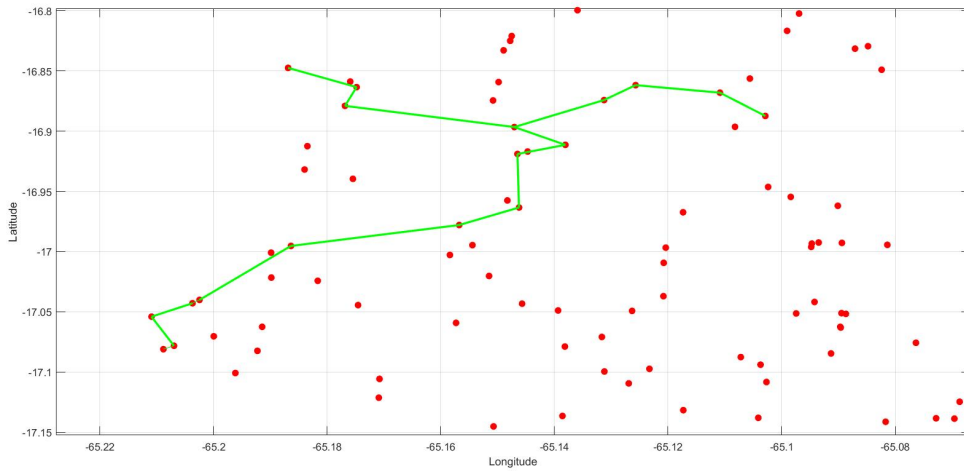


Figure 5.11: Network Result Obtained with ϵ -Constraint Method Right After the Removal of the Extreme Node

The numerical results of all of the experiments can be found in Appendix.

Chapter 6

Conclusion

Incorporating decentralized and centralized electrification systems for the rural electrification projects of developing countries helps to find better solutions in terms of both cost effectiveness and environmental responsibilities. However, due to computational complexity of the problem, especially for large scaled instances, solving the problem optimally requires a lot of time. Therefore, heuristic approaches are widely used since they can provide practical solutions in a reasonable amount of time. Yet, most of the time, they do not guarantee global optimality of the solution. In this paper, we focused on the problem introduced by Parshall et al. and the heuristic solution approach called Modified Kruskal's Algorithm [40]. Alternatively, we presented the Prize Collecting Steiner Tree model of the underlying problem, which yields the optimal solutions to the problem. Despite the optimality guarantee of the proposed approach, the long CPU times can make it difficult to use for larger instances. Therefore, faster methods solving PCST optimally that are based on more complicated algorithms can be implemented as a further research.

Our computational experiments were based on both randomly synthetic and real life instances. In this paper, our aim is to share the patterns and implications observed through our limited number of experiments. The empirical analysis showed that it is

possible to find the optimal solution of the least cost planning problem of centralized and decentralized electrification systems within a reasonable time using the PCST approach. Given that infrastructure planning problems are long term and capital intensive problems, most of the time, it is worth the wait to obtain the best solution. On the other hand, we observed that the existing MK approach can quickly provide reasonable solutions, which are close to the optimal solution.

Finally, recalling the ecological harm caused by traditional electrification methods, cost should not be the only objective to be considered. The new electrification planning policies for developing countries should therefore contribute both to sustainable development and to climate change mitigation. Motivated by that fact, a bi-objective optimization method for this problem was implemented. This permits taking into account not only the cost parameter but also the carbon emission. As a result of our computational experiments, we found out that it may be possible to have very large reductions in carbon emission with relatively small budget increase for some of the cases. Our results show that it is worth to make this kind of analysis since it is possible to spot very cost effective improvements. Best example of it is the 20-node example in which as a result of 7.7% increase in cost, the carbon emission can decrease by 58.3%.

Recalling that for the bi-objective model, the number of settlements in the problem affects the size of the models, and therefore the resolution times. If the problem can not be solved at reasonable times which can be the case for ϵ -constraint method with small predetermined step size, heuristic methods can be implemented for bi-objective optimization.

Bibliography

- [1] NASA's Goddard Institute for Space Studies, "May 2017 was second-warmest may on record." <https://climate.nasa.gov/news/2600/may-2017-was-second-warmest-may-on-record/>. Accessed June, 2017.
- [2] S. A. Montzka, E. J. Dlugokencky, and J. H. Butler, "Non-co₂ greenhouse gases and climate change," *Nature*, vol. 476, no. 7358, pp. 43–50, 2011.
- [3] N. Stern *et al.*, "What is the economics of climate change?," *World economics : Henley on Thames*, vol. 7, no. 2, p. 1, 2006.
- [4] S. Solomon, *Climate change 2007-the physical science basis: Working group I contribution to the fourth assessment report of the IPCC*, vol. 4. Cambridge University Press, 2007.
- [5] O. Edenhofer, R. Pichs-Madruga, Y. Sokona, K. Seyboth, P. Matschoss, S. Kadner, T. Zwickel, P. Eickemeier, G. Hansen, S. Schlömer, *et al.*, "Ipcc special report on renewable energy sources and climate change mitigation," *Prepared By Working Group III of the Intergovernmental Panel on Climate Change*, 2011.
- [6] N. Panwar, S. Kaushik, and S. Kothari, "Role of renewable energy sources in environmental protection: a review," *Renewable and Sustainable Energy Reviews*, vol. 15, no. 3, pp. 1513–1524, 2011.

- [7] R. Hiremath, S. Shikha, and N. Ravindranath, “Decentralized energy planning; modeling and application -a review,” *Renewable and Sustainable Energy Reviews*, vol. 11, no. 5, pp. 729 – 752, 2007.
- [8] C. Le Quéré, R. J. Andres, T. Boden, T. Conway, R. A. Houghton, J. I. House, G. Marland, G. P. Peters, G. Van der Werf, A. Ahlström, *et al.*, “The global carbon budget 1959–2011,” *Earth System Science Data Discussions*, vol. 5, no. 2, pp. 1107–1157, 2012.
- [9] S. Shafiee and E. Topal, “When will fossil fuel reserves be diminished?,” *Energy policy*, vol. 37, no. 1, pp. 181–189, 2009.
- [10] A. V. Herzog, T. E. Lipman, and D. M. Kammen, “Renewable energy sources,” *Encyclopedia of Life Support Systems (EOLSS). Forerunner Volume-Perspectives and Overview of Life Support Systems and Sustainable Development*, 2001.
- [11] D. P. Kaundinya, P. Balachandra, and N. Ravindranath, “Grid-connected versus stand-alone energy systems for decentralized power: a review of literature,” *Renewable and Sustainable Energy Reviews*, vol. 13, no. 8, pp. 2041–2050, 2009.
- [12] International Energy Agency, “World energy outlook,” 2015.
- [13] T. Winther, *The impact of electricity: Development, desires and dilemmas*. Berghahn Books, 2013.
- [14] World Bank, “Access to electricity (% of population).” <http://data.worldbank.org/indicator/EG.ELC.ACCS.ZS/countries>. Accessed June, 2017.
- [15] O. R. Davidson and Y. Sokona, *A new sustainable energy path for African development: Think bigger act faster*. Energy and Development Research Centre, University of Cape Town Cape Town, 2002.
- [16] J. Butlin, “Our common future. by world commission on environment and development. (london, oxford university press, 1987, pp.383 5.95.),” *Journal of International Development*, vol. 1, no. 2, pp. 284–287, 1989.

- [17] I. Dincer, “Renewable energy and sustainable development: a crucial review,” *Renewable and Sustainable Energy Reviews*, vol. 4, no. 2, pp. 157–175, 2000.
- [18] J. Goldemberg *et al.*, “Leapfrog energy technologies,” *Energy Policy*, vol. 26, no. 10, pp. 729–741, 1998.
- [19] S. Jebaraj and S. Iniyan, “A review of energy models,” *Renewable and Sustainable Energy Reviews*, vol. 10, no. 4, pp. 281–311, 2006.
- [20] Y. Abdul-Salam and E. Phimister, “The politico-economics of electricity planning in developing countries: A case study of ghana,” *Energy Policy*, vol. 88, pp. 299 – 309, 2016.
- [21] T. W. Lambert and D. Hittle, “Optimization of autonomous village electrification systems by simulated annealing,” *Solar Energy*, vol. 68, no. 1, pp. 121–132, 2000.
- [22] U. Deichmann, C. Meisner, S. Murray, and D. Wheeler, “The economics of renewable energy expansion in rural sub-saharan africa,” *Energy Policy*, vol. 39, no. 1, pp. 215–227, 2011.
- [23] T. Levin and V. M. Thomas, “Least-cost network evaluation of centralized and decentralized contributions to global electrification,” *Energy Policy*, vol. 41, pp. 286–302, 2012.
- [24] J. Van Hoesen and S. Letendre, “Evaluating potential renewable energy resources in poultney, vermont: A gis-based approach to supporting rural community energy planning,” *Renewable Energy*, vol. 35, no. 9, pp. 2114–2122, 2010.
- [25] S. Szabó, K. Bódis, T. Huld, and M. Moner-Girona, “Energy solutions in rural africa: mapping electrification costs of distributed solar and diesel generation versus grid extension,” *Environmental Research Letters*, vol. 6, no. 3, p. 034002, 2011.
- [26] S. Szabó, K. Bódis, T. Huld, and M. Moner-Girona, “Sustainable energy planning: Leapfrogging the energy poverty gap in africa,” *Renewable and Sustainable Energy Reviews*, vol. 28, pp. 500–509, 2013.

- [27] J. Urpelainen, “Grid and off-grid electrification: An integrated model with applications to india,” *Energy for Sustainable Development*, vol. 19, pp. 66 – 71, 2014.
- [28] M. Zeyringer, S. Pachauri, E. Schmid, J. Schmidt, E. Worrell, and U. B. Morawetz, “Analyzing grid extension and stand-alone photovoltaic systems for the cost-effective electrification of kenya,” *Energy for Sustainable Development*, vol. 25, pp. 75 – 86, 2015.
- [29] R. Rajbongshi, D. Borgohain, and S. Mahapatra, “Optimization of pv-biomass-diesel and grid base hybrid energy systems for rural electrification by using homer,” *Energy*, vol. 126, pp. 461–474, 2017.
- [30] B. J. Van Ruijven, J. Schers, and D. P. van Vuuren, “Model-based scenarios for rural electrification in developing countries,” *Energy*, vol. 38, no. 1, pp. 386–397, 2012.
- [31] J. Kaldellis, “Feasibility evaluation of greek state 1990–2001 wind energy program,” *Energy*, vol. 28, no. 14, pp. 1375–1394, 2003.
- [32] G. Bakos and N. Tsagas, “Technoeconomic assessment of a hybrid solar/wind installation for electrical energy saving,” *Energy and Buildings*, vol. 35, no. 2, pp. 139–145, 2003.
- [33] MNRE-Ministry of New and Renewable Energy India, “Baseline in grid connected renewable energy projects.,” 2004.
- [34] M. Ranaboldo, A. García-Villoria, L. Ferrer-Martí, and R. P. Moreno, “A heuristic method to design autonomous village electrification projects with renewable energies,” *Energy*, vol. 73, pp. 96–109, 2014.
- [35] M. Ranaboldo, A. García-Villoria, L. Ferrer-Martí, and R. P. Moreno, “A meta-heuristic method to design off-grid community electrification projects with renewable energies,” *Energy*, vol. 93, pp. 2467–2482, 2015.

- [36] M. Kolhe, S. Kolhe, and J. Joshi, “Economic viability of stand-alone solar photovoltaic system in comparison with diesel-powered system for india,” *Energy Economics*, vol. 24, no. 2, pp. 155–165, 2002.
- [37] S. Chakrabarti and S. Chakrabarti, “Rural electrification programme with solar energy in remote region—a case study in an island,” *Energy Policy*, vol. 30, no. 1, pp. 33–42, 2002.
- [38] A. Zvoleff, A. S. Kocaman, W. T. Huh, and V. Modi, “The impact of geography on energy infrastructure costs,” *Energy Policy*, vol. 37, no. 10, pp. 4066–4078, 2009.
- [39] A. S. Kocaman, W. T. Huh, and V. Modi, “Initial layout of power distribution systems for rural electrification: A heuristic algorithm for multilevel network design,” *Applied Energy*, vol. 96, pp. 302–315, 2012.
- [40] L. Parshall, D. Pillai, S. Mohan, A. Sanoh, and V. Modi, “National electricity planning in settings with low pre-existing grid coverage: development of a spatial model and case study of kenya,” *Energy Policy*, vol. 37, no. 6, pp. 2395–2410, 2009.
- [41] A. Sanoh, L. Parshall, O. F. Sarr, S. Kum, and V. Modi, “Local and national electricity planning in senegal: Scenarios and policies,” *Energy for Sustainable Development*, vol. 16, no. 1, pp. 13–25, 2012.
- [42] F. Kemausuor, E. Adkins, I. Adu-Poku, A. Brew-Hammond, and V. Modi, “Electrification planning using network planner tool: The case of ghana,” *Energy for Sustainable Development*, vol. 19, pp. 92–101, 2014.
- [43] Y. Abdul-Salam and E. Phimister, “How effective are heuristic solutions for electricity planning in developing countries,” *Socio-Economic Planning Sciences*, vol. 55, pp. 14–24, 2016.

- [44] S. Pohekar and M. Ramachandran, “Application of multi-criteria decision making to sustainable energy planning: a review,” *Renewable and Sustainable Energy Reviews*, vol. 8, no. 4, pp. 365–381, 2004.
- [45] M. Fadaee and M. Radzi, “Multi-objective optimization of a stand-alone hybrid renewable energy system by using evolutionary algorithms: a review,” *Renewable and Sustainable Energy Reviews*, vol. 16, no. 5, pp. 3364–3369, 2012.
- [46] D. Abbes, A. Martinez, and G. Champenois, “Life cycle cost, embodied energy and loss of power supply probability for the optimal design of hybrid power systems,” *Mathematics and Computers in Simulation*, vol. 98, pp. 46–62, 2014.
- [47] H. Ren, W. Zhou, K. Nakagami, W. Gao, and Q. Wu, “Multi-objective optimization for the operation of distributed energy systems considering economic and environmental aspects,” *Applied Energy*, vol. 87, no. 12, pp. 3642–3651, 2010.
- [48] R. Dufo-López, J. L. Bernal-Agustín, J. M. Yusta-Loyo, J. A. Domínguez-Navarro, I. J. Ramírez-Rosado, J. Lujano, and I. Aso, “Multi-objective optimization minimizing cost and life cycle emissions of stand-alone pv–wind–diesel systems with batteries storage,” *Applied Energy*, vol. 88, no. 11, pp. 4033–4041, 2011.
- [49] M. Sharafi and T. Y. ElMekkawy, “Multi-objective optimal design of hybrid renewable energy systems using pso-simulation based approach,” *Renewable Energy*, vol. 68, pp. 67–79, 2014.
- [50] E. Miandoabchi, R. Z. Farahani, and W. Y. Szeto, “Bi-objective bimodal urban road network design using hybrid metaheuristics,” *Central European Journal of Operations Research*, vol. 20, no. 4, pp. 583–621, 2012.
- [51] C. Kanzian, M. Kühmaier, J. Zazgornik, and K. Stampfer, “Design of forest energy supply networks using multi-objective optimization,” *Biomass and Bioenergy*, vol. 58, pp. 294–302, 2013.

- [52] F. Wang, X. Lai, and N. Shi, “A multi-objective optimization for green supply chain network design,” *Decision Support Systems*, vol. 51, no. 2, pp. 262–269, 2011.
- [53] R. Farmani, D. Savic, and G. Walters, “Evolutionary multi-objective optimization in water distribution network design,” *Engineering Optimization*, vol. 37, no. 2, pp. 167–183, 2005.
- [54] G. Mavrotas, “Effective implementation of the ε -constraint method in multi-objective mathematical programming problems,” *Applied mathematics and computation*, vol. 213, no. 2, pp. 455–465, 2009.
- [55] J. B. Kruskal, “On the shortest spanning subtree of a graph and the traveling salesman problem,” *Proceedings of the American Mathematical society*, vol. 7, no. 1, pp. 48–50, 1956.
- [56] D. S. Johnson, M. Minkoff, and S. Phillips, “The prize collecting steiner tree problem: theory and practice,” in *SODA*, vol. 1, p. 4, 2000.
- [57] M. X. Goemans and D. P. Williamson, “A general approximation technique for constrained forest problems,” *SIAM Journal on Computing*, vol. 24, no. 2, pp. 296–317, 1995.
- [58] M. Haouari, S. B. Layeb, and H. D. Sherali, “Tight compact models and comparative analysis for the prize collecting steiner tree problem,” *Discrete Applied Mathematics*, vol. 161, no. 4, pp. 618–632, 2013.
- [59] K. Deb, *Multi-objective optimization using evolutionary algorithms*, vol. 16. John Wiley & Sons, 2001.
- [60] Y. Y. Haimes, “On a bicriterion formulation of the problems of integrated system identification and system optimization,” *IEEE transactions on systems, man, and cybernetics*, vol. 1, no. 3, pp. 296–297, 1971.
- [61] D. Trystram, *Multi-Criteria Optimization and its Application to Multi-Processor Embedded Systems*. PhD thesis, Université de Provence, 2011.

- [62] T. Koch and A. Martin, “Solving steiner tree problems in graphs to optimality,” *Networks*, vol. 32, no. 3, pp. 207–232, 1998.
- [63] I. Ljubic, R. Weiskircher, U. Pferschy, G. W. Klau, P. Mutzel, and M. Fischetti, “Solving the prize-collecting steiner tree problem to optimality,” in *ALENEX/ANALCO*, pp. 68–76, 2005.
- [64] M. Haouari, S. B. Layeb, and H. D. Sherali, “Algorithmic expedients for the prize collecting steiner tree problem,” *Discrete Optimization*, vol. 7, no. 1, pp. 32–47, 2010.
- [65] A. S. da Cunha, A. Lucena, N. Maculan, and M. G. Resende, “A relax-and-cut algorithm for the prize-collecting steiner problem in graphs,” *Discrete Applied Mathematics*, vol. 157, no. 6, pp. 1198–1217, 2009.

Appendix A

Results

Table A.1: Carbon Emission and Total Cost(\$) Results Obtained by Epsilon-Constraint Method for the Synthetic Data n=10

Total Cost	Carbon Emission Value
112.648	5
113.523	4
117.481	3
123.042	2
125	0

Table A.2: Carbon Emission and Total Cost(\$) Results Obtained by Epsilon-Constraint Method for the Synthetic Data n=20

Total Cost	Carbon Emission Value
205.45	12
205.62	11
206.69	10
208.22	9
208.39	8
211.58	7
213.26	6
221.14	5
235.70	4
251.40	3
259.28	2
305	0

Table A.3: Carbon Emission and Total Cost(\$) Results Obtained by Epsilon-Constraint Method for the Synthetic Data n=50

Total Cost	Carbon Emission Value
416.55	28
417.04	27
417.95	26
418.45	25
420.37	24
420.82	23
421.32	22
423.25	21
424.33	20
424.83	19
426.75	18
428.71	17
429.63	16
431.55	15
433.11	14
434.03	13
435.39	12
435.44	11
436.79	10
438.31	8
439.67	7
443.17	6
444.21	5
446.31	3
447.14	2
467.23	0

Table A.4: Carbon Emission and Total Cost(\$) Results Obtained by Epsilon-Constraint Method for the Synthetic Data n=100

Total Cost	Carbon Emission Value
580	58
580.06	57
580.89	56
581.76	55
582.63	54
583.56	53
584.55	52
585.49	51
586.68	50
587.61	49
589.09	48
590.08	47
591.02	46
592.92	45
593.95	44
595.45	43
596.56	42
597.68	41
597.86	38
597.92	37
598.29	36
599.13	35
600	34
600.86	33
601.80	32

603.70	31
605.46	30
605.74	29
606.57	28
607.44	27
608.30	26
609.24	25
610.43	24
611.37	23
613.37	22
615.36	21
618.47	20
622.76	19
624.84	18
628.54	17
629.93	16
630.86	15
632.14	14
633.41	13
633.60	12
634.53	11
635.50	10
637.50	9
638.42	8
640.42	7
642.41	6
646.42	5
648.41	4
653.33	3

654.16	2
672	0

Table A.6: Carbon Emission and Total Cost(\$) Results Obtained by Epsilon-Constraint Method for the Real Data n=141

Total Cost	Carbon Emission Value
39326169.55	17124544.02
39364338.98	16993412.09
39408667.14	16860878.16
39468004.73	16727409.97
39634120.38	16575825.44
39839095.83	16454352.10
39992668.28	16325735.28
4022929275	16185117.15
71276086.97	968615.03
71296830.08	851924.76
71343272.13	801880.54
71390186.59	669889.68
71476250.53	536277.94
71642367.12	384693.41
71983615.82	151584.53
72234581.00	0