The Reference Electrification Model: A Computer Model for Planning Rural Electricity Access

by

Douglas Ellman
A.B. Physics, Princeton University (2009)

Submitted to the Engineering Systems Division
in Partial Fulfillment of the Requirements for the Degree of

MASTER OF SCIENCE IN TECHNOLOGY AND POLICY
AT THE
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2015

Signature of Author: ____
Engineering Systems Division
May 14, 2015

Certified by: ____
Ignacio Pérez-Arriaga
Visiting Professor, Engineering Systems Division
Thesis Co-Supervisor

Certified by: ____
Claudio Vergara
Postdoctoral Associate, MIT Energy Initiative
Thesis Co-Supervisor

Accepted by: ____
Dava Newman
Professor of Aeronautics and Astronautics and Engineering Systems
Director, Technology and Policy Program

The Reference Electrification Model: A Computer Model for Rural Electrification Planning

by

Douglas Ellman

Submitted to the Engineering Systems Division
on May 14, 2015 in Partial Fulfillment of the
Requirements for the Degree of Master of Science in
Technology and Policy

ABSTRACT

Despite efforts from governments and other organizations, hundreds of millions of people—primarily in Africa and South Asia—still have no electricity service. Electrification efforts have historically been focused on extension of the main electric grid, but technology developments have made off-grid power systems, such as microgrids and home systems, viable alternatives for some areas. Especially since rural electrification typically depends on limited subsidies, if universal electrification is to be achieved in a timely manner, smart planning is essential to ensure that resources are directed towards cost-efficient technical solutions.

Since the areas requiring electrification are expansive, the technology choices are many, and experience with off-grid systems is limited, planners struggle to evaluate tradeoffs between technology choices and estimate project costs. This thesis demonstrates that computer models that can automatically produce cost-efficient designs to the individual customer level can provide significant value to the planning process. The development of such a model by the author and collaborators at MIT and Comillas University, called the Reference Electrification Model (REM), is described. REM uses a series of heuristics to process input data, identify areas better suited for on-grid or off-grid electrification, and produce technical designs for recommended grid-extension and off-grid projects. In addition to the current state of REM, the rationale for model design choices and recommendations for future developments are described. The process and results of a pilot application of REM to Vaishali District, in Bihar, India are also described.

REM will only be useful if it is actually incorporated into planning processes. In this spirit, concepts for how models like REM can benefit the regulation of rural electrification are presented, with a focus on India.

Thesis Co-Supervisor: Ignacio Pérez-Arriaga
Title: Visiting Professor, Engineering Systems Division

Thesis Co-Supervisor: Claudio Vergara
Title: Postdoctoral Associate, MIT Energy Initiative

Acknowledgements

Firstly, thank you to my advisor and thesis co-supervisor, Ignacio, who provided the opportunity for me to get involved with this exciting and important project. He provided great guidance and support, was a wealth of knowledge and experience, and challenged me to accomplish and learn more than I thought I would or could in my time at MIT. Along the way we also had a lot of fun.

Thank you to my other thesis co-supervisor Claudio, who worked most closely with me in implementation of REM, and made significant contributions in coding parts of the model. I depended on many conversations with Claudio to develop a better understanding of power systems and to generate ideas to help me solve big and small problems. Claudio's positive attitude helped to keep me and the team on track through some stressful times.

Thank you to Yael, who worked closely with me in all aspects of my research. We tallied an impressive record of world travels and presentation tag-teams over the last two years. Yael served as a consistent source of common sense whenever we risked getting lost in the weeds.

Thank you, also, to the rest of the Universal Access team, including Reja, Andrés, Lily, Vivian, and Patricia. We shared many great discussions about energy access, a lot of hard work, and many fun times.

Thank you to Rob for his exceptional guidance and support, especially considering his many other obligations. Thank you also to the rest of the Tata Center and the Tata Trusts for the financial, educational, and administrative support.

Thank you to the team at IIT Comillas, who provided feedback to my ideas and provided access to and support with using their Reference Network Models. Thank you especially to Carlos, who put many hours into helping me to understand and make effective use of the Reference Network Models.

Thank you to the Technology and Policy Program for providing me with the opportunity to learn a new set of skills to understand the world.

Thank you to all of the people in India who provided time and effort in support of my work. Exceptional support was provided by people at Tata Power DDL, the North Bihar Power Distribution Company, the Central Electricity Regulatory Commission, SELCO, and Tata Consultancy Services. Special thanks to the people of rural India who let us into their homes and communities as we sought to understand the electricity access situation.

Thank you, finally, to my family and friends, whose love and support have been crucial over the last two years. Special thanks to my parents and to my girlfriend, Elyse, who have been especially supportive. I am looking forward to seeing more of you all when I finish school!

CONTENTS

1 INTRODUCTION

1.1 MOTIVATION AND RESEARCH QUESTION

The scale of the global electricity access problem is enormous. As of 2010, 17% of the global population lacked access to electricity—that is 1.2 billion people (World Bank & International Energy Agency, 2014). Most of these people are in South Asia (418 million) and Sub-Saharan Africa (590 million). This issue is an important part of development generally and is related to other development goals—for example electricity access is correlated with the Human Development Index (Bhattacharyya & Palit, 2013).

There is a strong commitment by countries and the international community to dramatically improve the electricity access situation, including the United Nations' Sustainable Energy For All initiative, which set targets including universal electricity access by 2030 (World Bank & International Energy Agency, 2014).

Planning for the required investments to achieve universal electricity access represents a huge challenge. There are major uncertainties about the geographic distribution and nature of electricity demand in currently unelectrified areas. With the emergence of off-grid technologies, such as microgrids and home systems,[1] planners now have to consider more technical options than ever before. In fact, the IEA estimates 70% of rural areas in the developing world may be served by off-grid systems (International Energy Agency, 2013), which would require a totally different paradigm for rural electrification planning compared to traditional grid extension. And the sheer numbers of customers to be connected may render some traditional planning methods to be inadequate.

For all of these reasons, we now require new planning tools that are commensurate with the magnitude of the electricity access problem. Fortunately, new capabilities exist today which were unavailable to planners in the past. High-quality satellite imagery is now easily accessible, and image processing techniques make it possible to automatically locate important features—such as the locations of buildings—over large areas. Also, advancements in computing power have made it feasible to automatically produce power system designs for hundreds of thousands of customers in reasonable time.

In this thesis, I describe my attempt to take advantage of these capabilities by developing a new computer planning tool specifically catered to the context of planning for universal electricity access where extension of the grid and off-grid systems are being considered. In developing the model—called the Reference Electrification Model, or REM—and applying it to a district in India, I attempt to answer the following research question:

"Can computer models effectively support the joint planning of grid extension and off-grid systems for universal electricity access in a large region?"

[1] In this document, "microgrids" refer to power systems consisting of local generation sources and an electricity distribution network, which are not connected to the main electric grid. Terms such as microgrid and mini-grid have different meanings in different communities, sometimes relating to the size of the system and whether it is grid-connected. "Home systems" or "isolated systems" refer to power systems with local generation sources that serve only a single customer, and thus do not require an electricity distribution network (except any internal wiring).

1.2 MIT's Universal Energy Access Research Group

This thesis describes the products of my individual research, but my work was done as part of the Universal Energy Access Research Group at MIT. I worked collaboratively with other team members throughout my research, especially in the areas of gathering data and coding parts of the model. In this thesis I have emphasized the parts of the work in which I was the lead contributor, but also discuss the products of work in which I played a supporting role. Even where my teammates led in implementation in some parts of this work, I was fully intellectually engaged and a part of the decision making process about what to implement. I must acknowledge and thank my teammates who worked with me on this project, because it allowed me to tackle a much more ambitious project than I otherwise would have be able to.

The Universal Energy Access Research Group is supervised by Professor Ignacio Pérez-Arriaga, and includes students and researchers at MIT, as well as students and researchers at the Instituto de Investigación Tecnológica (IIT) at Comillas University in Madrid, Spain.

This thesis, focused primarily on the consideration of the technoeconomic factors in rural electrification planning through computer models, is complemented by the thesis of Yael Borofsky, another member of the Universal Energy Access Research Group who focused on non-technical factors.

1.3 Previous work in using computer models for rural electrification planning

1.3.1 Literature

There is a body of literature pertaining to planning of energy access in the developing world. Much of this literature considers practical lessons drawn from case-studies or personal experiences. This emphasis on practical lessons is useful because successful implementation of off-grid electricity systems critically depends on negotiating a wide range of non-technical challenges, such as community dynamics, payment collection, finance, and government policies. Context relevant to the design of rural electrification systems is provided in (Chakrabarti & Chakrabarti, 2002; Chaurey & Kandpal, 2010; Millinger, Marlind, & Ahlgren, 2012; Oda & Tsujita, 2011) but these papers do not emphasize the specification of algorithms for the facilitation of system design. The most important consideration from these papers is the electricity demand scenario, which is very different from the usual electricity demand considered in developed world power system design. Important demand considerations which are incorporated into the REM include allowing that not all demand is served (since this may be too expensive) and providing multiple levels of value for electricity service (for example representing the very large value of basic services like lighting).

One computer-based approach to the problem of determining least-cost electrification modes for rural electrification (comparing between grid-extension, microgrids, and home systems) has been addressed in (Kemausuor, Adkins, Adu-Poku, Brew-Hammond, & Modi, 2014). This approach considers generally the same type as information as REM, except that it makes decisions at a "community" level, and does not consider the location of individual loads. A comparison of cost of microgrids versus home systems for rural electrification was performed in (Chaurey & Kandpal, 2010). This assessment was based on general descriptions of a potential microgrid site including parameters such as dispersion, rather than considering specific household locations.

REM includes advances over these types of approaches in terms of the range of technical and design options considered, in the technical rigor of evaluating each option, and in the ability to produce actual design recommendations.

1.3.2 Existing solutions

The most comparable existing tool to REM that I have come across is called Network Planner. Network Planner is an online tool for the selection of electrification mode for rural communities. It was developed by the Modi Research Group and Columbia University and is available at http://networkplanner.modilabs.org/. An application of Network Planner is described in (Kemausuor et al., 2014).

Network planner has some advantages over the current implementation of REM in terms of usability. It already has a functioning web-interface for the public to access the model. By considering aggregated communities and having lower-fidelity technical representations, the data requirements and barriers to understanding the model's operation are lower.

However, REM has advantages in terms of being able to consider a wider range of design and technology choices at a higher level of technical fidelity. Thus, while the learning curve and data requirements may be steeper for REM, the quality and detail of results and derived insights should be better.

1.4 Overview of the Reference Electrification Model

The Reference Electrification Model, or REM, is a computer model which can assist in planning or studying electricity access in large rural areas. Given information about a region, the model identifies appropriate areas for extension of the main electric grid and development of off-grid systems, based on cost-minimization. The costs considered include financial costs and cost to the customer if electricity demand is not fully met. REM also produces preliminary technical designs for the grid extension and off-grid projects. These designs are used to estimate project costs and some operating characteristics of the systems.

REM is a static model, in that it considers a single future year, and produces system designs and cost estimates based on serving the electricity demand in that year. REM does, however, take into account some year-to-year effects in a simplified way in order to estimate the lifetime performance of a system, including degradation in solar panel and battery performance.

For a single run of REM, the user provides a description of the demand of a typical customer for a certain initial year and an annual rate of demand growth. The user also indicates how many years are between the initial year and the "final year" considered—this is the year in which project costs will be calculated and technical constraints must be met. REM considers that the user is making a decision in the initial year about what systems to build for the final year.

REM is a model under active development through the work of members of the Energy Access Research Group at MIT, under the supervision of Professor Ignacio Perez-Arriaga, and collaborators at Instituto de Investigación Tecnológica (IIT) at Comillas University in Spain. The ambition is that, through continued improvements to the model's core functionality and user interface, REM can become widely useful to the global energy access community. This document describes the model as of May 2015, and some recommendations for future improvements.

1.5 On the application of REM to Vaishali District

1.5.1 Introduction to the Vaishali Study

In the summer of 2014, members of the MIT Universal Energy Access research group, including me, visited Bihar, India to present work done on the REM model and to learn more about the energy access situation there. During that visit, the Bihar Energy Department and the Bihar State Power Holding Company agreed to provide assistance in applying REM to one district in Bihar. Due primarily to the availability of information, it was agreed to focus on Vaishali District. Vaishali is within the area of the North Bihar Power Distribution Company Limited (NBPDCL), one of the two state-owned electricity distribution companies in Bihar. NBPDCL and the staff of the Bihar Energy Department and Bihar State Power Holding Company in general were extremely helpful and generous with their time in providing information to support an application of REM to Vaishali.

After an initial round of data gathering in summer 2014 and a semester of work, our team returned to Bihar in January 2015 in order to present our progress on the Vaishali study and gather additional information. This thesis includes our best available information about Vaishali as of that trip.

1.5.2 India electricity access context

In order to put the Vaishali study into context, it is useful to understand some background information about the electricity access situation in India broadly.

Of the 1.3 billion people that were without access to electricity in 2011, about 300 million were in India, making it the single country with the most people lacking electricity access (International Energy Agency, 2013). The electricity access situation varies significantly across India, which is a very large and diverse country. In several states the rate of electricity connection is, at least officially, over 99%, according to the 2011 census. But in some states the situation is much worse. Bihar, the state which includes Vaishali, has 83 million residents, and only 16% use electricity as their primary source of light, also according to 2011 census data. Many of those with grid connection have suffered from poor quality of service, with electricity historically available maybe just a few hours per day and not necessarily every day or at the times when electricity is most needed, especially in rural areas (Santhakumar, 2008).

India has made significant efforts to improve the state of the power sector through major reforms via the Electricity Act, 2003 and subsequent supporting legislation and policy directives (Kumar & Chatterjee, 2012). Additionally, India has implemented several programs to support rural electrification, most notably the Rajiv Gandhi Grameen Vidyutikaran Yojana (RGGVY) program, which emphasizes grid extension projects. There are also active, but smaller, efforts in provision of off-grid electricity service through government programs such as the Decentralized Distributed Generation scheme and the activities of many private enterprises. While there is some coordination between government off-grid projects and grid extension plans, many commercial off-grid system providers are operating outside the scope of current regulation, and thus are not well integrated into planning processes. Despite some improvements resulting from these efforts, India has failed to achieve past goals of universal electrification, and hundreds of millions of Indians still lack reliable electricity access.

1.5.3 About Vaishali District

Vaishali is a district in the state of Bihar, India, which is located just across the Ganges River from the state capital of Patna. As mentioned above, Bihar is the state with the worst electrification rate, with only 16% of households reporting using electricity as a primary source of light.

As of the 2011 census, Vaishali has a population of about 3.5 million people, who are in about 626 thousand households. The population had grown about 29% between the 2001 census and the 2011 census. Vaishali has an area of about 2000 square kilometers and a population density of about 1700 people per square kilometer.

To put this in perspective, Vaishali has an area comparable to the smallest US state by area (Rhode Island) and a population comparable to the 29th largest US state by population (Connecticut), resulting in a population density more than three times greater than the most densely populated US state (New Jersey).

Despite this population density, about 93% of Vaishali's population lives in rural areas. Much of the district, like much of India, is characterized by densely packed villages or hamlets dispersed throughout farm land. There are also three urban areas: Hajipur (the largest), Lalganj, and Mahnar.

Most of the state is relatively flat and is generally accessible by road. The main accessibility challenges are found in the sub-district of Raghopur, which is located on an island in the middle of the Ganges River. For much of the year Raghopur is only accessible by boat and the island is prone to severe flooding.

Based on NBPDCL's customer count and the 2011 census count of number of households, about 20% of Vaishali households are electrified, making it a better than average district for Bihar, but still in need of massive expansion of the distribution infrastructure. There are active projects and proposals for the expansion of the grid in Vaishali, but a wide variety of opinions about when the grid will reach most customers.

2 DATA INPUTS

In order to receive quality results from REM, the data input must be of sufficient quality. This chapter describes the data inputs which are required to run REM, and some thoughts regarding how they can be obtained. In general, a pre-processing step is required in which the REM user transforms the available raw data (and assumptions, as necessary) into the data inputs in the appropriate formats for REM. The nature of this pre-processing step will vary from case to case, as the format and availability of data may vary widely, especially in developing world applications. As an example, the process of gathering and processing input data for Vaishali is described throughout the chapter.

2.1 ON THE ORGANIZATION OF DATA

Data inputs in REM are organized at several geographical levels. The table below provides some terminology related to various levels at which input data can be associated.

Data Level Name	Description
Study area	The entire area under consideration in a run of REM. Most of the input data is associated with the study area.
Sub-district	Geographical sub-divisions of the study area. Some input data is associated with the sub-district. Sub-district boundaries may be defined based on administrative boundaries, as this is a useful division for organizing input data and presenting results.
Analysis region	Geographical sub-divisions of sub-districts. Analysis regions are created for convenience or to manage computer resources, when solving an entire sub-district in one shot is not desirable or feasible. They inherit data from the study area and sub-district.
Customer	A single node representing a household or other customer.

Vaishali:

For the Vaishali study, the study area is the entire district of Vaishali. The sub-districts are defined by the administrative boundaries of the sub-districts of Vaishali. There are 16 sub-districts. For each REM scenario, the 16 sub-districts were divided into 30 analysis regions.

The administrative boundaries (defining the study area and sub-districts) were obtained by geo-referencing a map of Vaishali which had been recently produced by UNICEF[2]. This was the only source of data we were able to find regarding the 2011 census boundaries. An image of the map is provided below.

The analysis regions were defined by using the k-means algorithm[3] to divide the customers which are believed to be not currently grid-connected into groups with comparable numbers of customers, such that there were a total of 30 analysis regions.

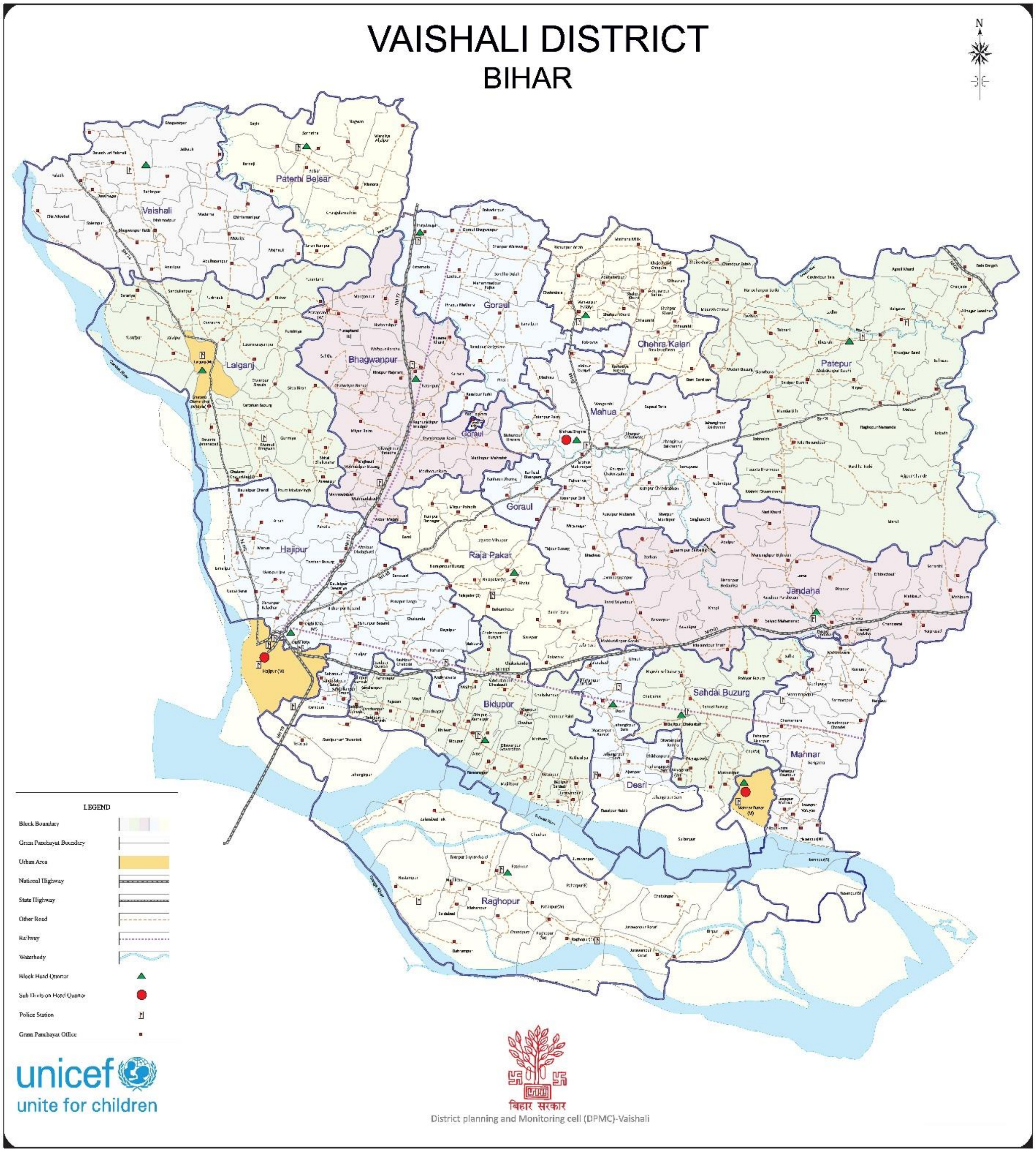
---

[2] Thank you to Tata Consultancy Services for support in this geo-referencing task.

[3] The k-means algorithm groups points by minimizing the average distance from points to the center of the group. More information is available at http://en.wikipedia.org/wiki/K-means_clustering.

The particular number of 30 for analysis regions was somewhat arbitrary, but here are some considerations regarding the selection of number of analysis regions:

- Larger analysis regions generally lead to better quality results because there are less artificial divisions that systems cannot cross.
- Analysis regions can be solved in parallel, so matching the number of analysis regions to the number of regions your computer can solve in parallel (considering constraints in memory and number of logical threads), can speed up model execution.



2.2 EXISTING GRID DATA

If extension of the existing grid is to be considered as an electrification mode, information about the existing grid is required. REM considers the distribution grid, including transmission substations, high voltage distribution lines, high-to-medium voltage transformers, medium voltage lines, medium-to-low voltage transformers, and low voltage lines. Several categories of information are needed: geographic and topological, electrical, economic, and service quality.

2.2.1 Geographic and topological

The geographical and topological description of the existing grid includes the specification of coordinates of grid components and the connections between components. This is done by associating point elements (substations and transformers) with nodes, and associating line elements (power line segments) with pairs of nodes. Elements associated with the same node are electrically connected. Each node has a specified location.

In the current version of REM, this detailed geographical and topological representation covers the transmission substations, high voltage distribution lines, high-to-medium voltage substations, and medium voltage lines. Detailed geographical and topological information about the existing medium-to-low voltage transformers and low voltage distribution lines are not included.

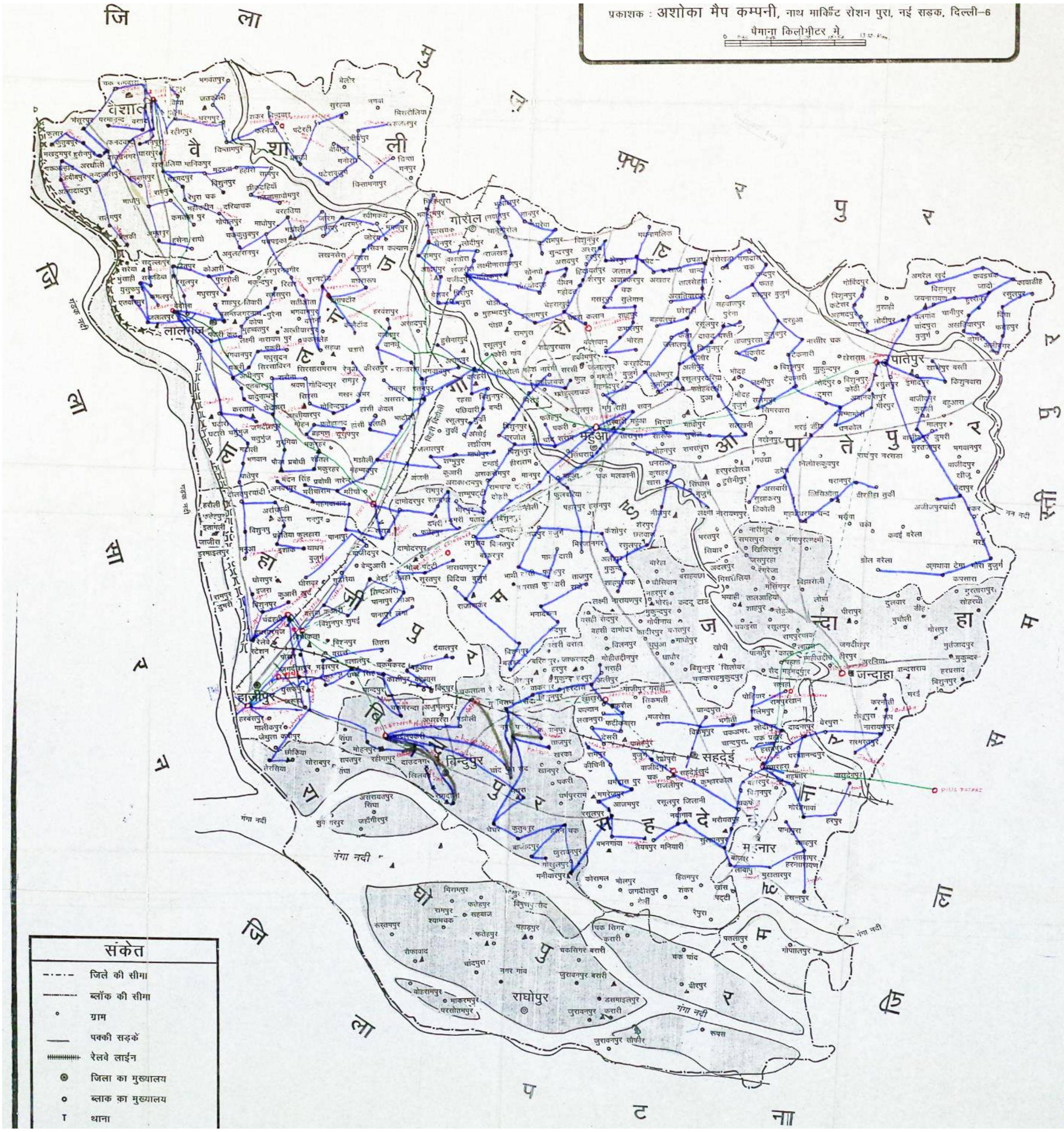
Each substation is associated with a name and each feeder emanating from a substation also gets a name. If a feeder branches, its path is described by multiple series of nodes—one for each branch. Each branch then gets a name, consisting of the feeder name followed by a number to differentiate the branches.

The locations of existing medium-to-low voltage transformers and low-voltage distribution lines are considered through the specification of customers which are connected to the low-voltage grid. In practice, high-quality information about low-voltage distribution may not be available, and so the path of the medium-voltage lines may be the best available information about locations of grid-connected customers.[4]

Vaishali:

For the Vaishali study, we were fortunate to find a map with geographical information about the paths of much of the distribution network. It appears that in most of Bihar there is not much documentation of the geographical layout of distribution infrastructure, so the existence of this map was one of the main reasons that we selected Vaishali for an initial study location. An image of this map appears below. It includes paths of high voltage lines (33 kV) in green, medium voltage lines (11 kV) in blue, and substation locations written in red. This map was later supplemented and modified with additional information about the locations of substations and power lines.

[4] See the section on customer data for additional information about how the locations of electrified customers are specified.



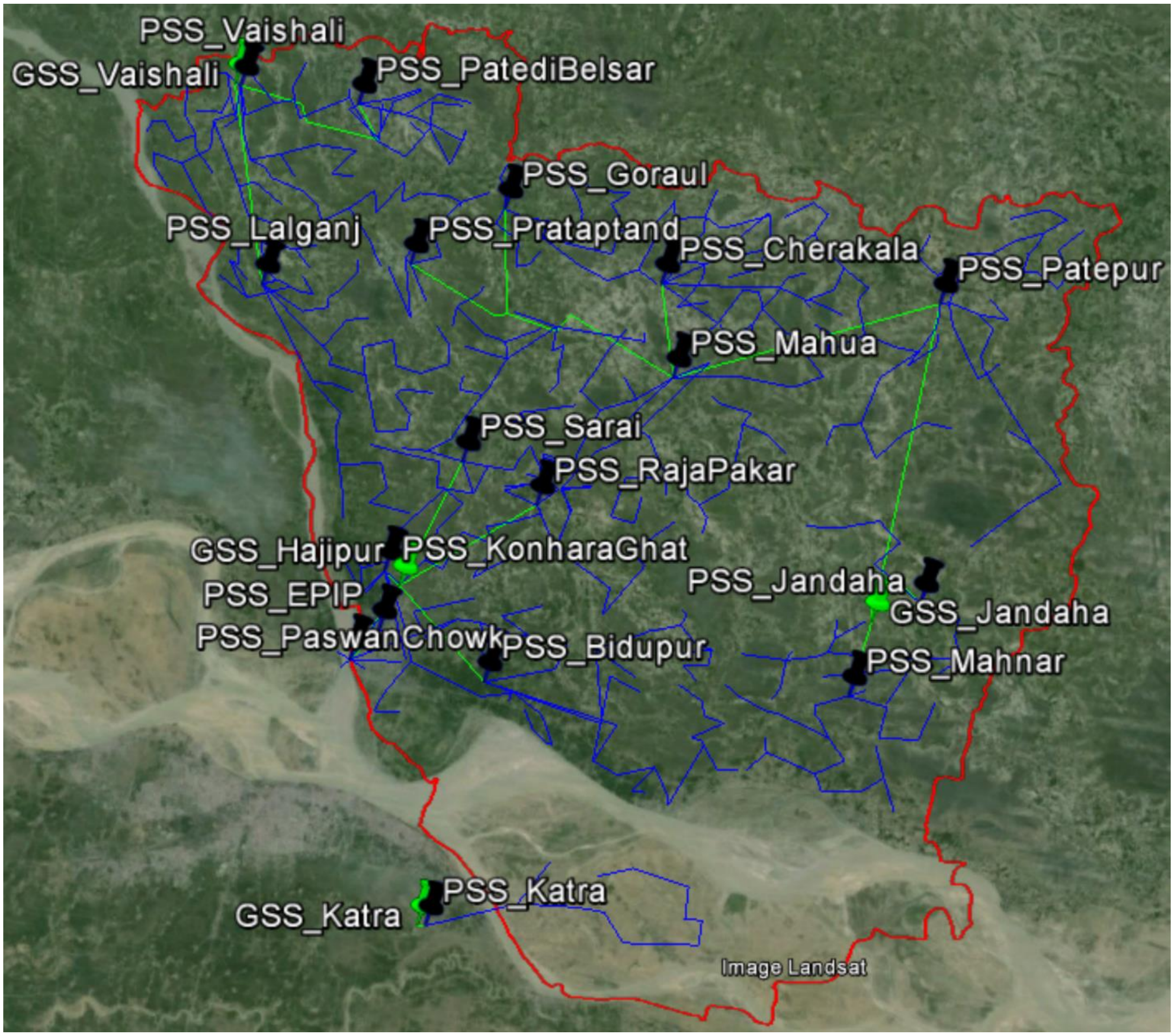
Here is a summary of the process that we used to turn this map into a digital representation of the topology and geographical locations of the distribution network in Vaishali:

1. We roughly oriented a digital image of the map over Vaishali district in Google Earth. At this point the features in the image were generally close to their true geographic location, but there were certainly significant errors, on the order of a few kilometers.
2. We created KML files in Google Earth containing point features for each of the transmission substations and high-to-medium voltage substations, and path features for the high voltage lines and medium voltage lines. On the original map, high voltage (33 kV) distribution lines are in green and medium voltage (11 kV) lines are in blue. The KML path objects representing the

medium voltage lines were organized by “feeders.” One feeder corresponds to all of the branches downstream of a single line which exits a high-to-medium voltage substation.

3. There were some ambiguities in the existing grid topology based on the information in the map, and we knew that the distribution company had done work on the grid subsequent to the creation of the map. To augment the map with the best-available current information, we visited the office of the chief engineer responsible for this part of the grid, who is located in the city of Hajipur in Vaishali. The chief engineer was able to provide us current information about recent projects which reconfigured the high-voltage connections of the high-to-medium voltage substations, created new sub-stations, and created new medium voltage feeders. Location information was provided to us in terms of sequences of villages, which we matched to village names listed on the map. The KML files described in step 2 above were updated to take account of this information. Since this updated information was provided to us verbally and not through validated documentation, it is possible that there are some errors in our digital representation.
4. The KML files were read into Matlab, where each node (corresponding to a substation or vertex of a distribution line) was given a reference number. Nodes which represented electrically-connected elements were joined into a single node (for example, a substation and the head of a feeder, or two branches of a feeder where they meet). A spreadsheet containing the name and coordinates of each node was saved, along with Matlab structures which described the topology of the grid in terms of sequences of nodes.
5. In order to improve the accuracy of the locations of nodes, we obtained the help of Tata Consultancy Services to identify the names of the villages corresponding to grid nodes on the map and find accurate coordinates of those villages. When they updated the spreadsheet with improved locations, we were able to incorporate this information. The updated locations have not yet been validated with on-the-ground information, and may still have significant errors.
6. The Matlab structures with the topological description of the grid and the updated spreadsheet with locations of nodes were used to create a structure with information about the medium voltage grid lines, which is used for producing grid extension designs. This structure represents the paths of the medium voltage grid as a series of line segments, each connecting a pair of nodes.

An image of the final existing grid data is shown below. High voltage lines (33 kV) are in green and medium voltage lines (11 kV) are in blue. The transmission substations are labeled with the prefix “GSS” for “grid substation” and the high-to-medium voltage substations are labeled with the prefix “PSS” for “power substation.”



Overall, this was a significantly labor-intensive process to get a reasonable picture of the topology of the high and medium voltage distribution network, and it is very likely that there are some errors in our description of the topology and locations. Also, due to ongoing projects, it is likely that the grid changed between the time this information was gathered and now. Thus, while this process was perhaps necessary and sufficient for an initial test case of REM, it would have been much better if the distribution company was able to directly provide validated digital GPS data about the existing grid. In Bihar, they were beginning a process of GPS mapping of their grid assets, so the availability and quality of information is likely to improve in the future.

2.2.2 Electrical

Additional information is added to the geographical and topological representation of the grid to enable the creation of an electrical representation of the grid. The electrical information about the grid is processed into a specific format so that it can be used by the brownfield version of the Reference Network Model. The brownfield RNM was developed at IIT Comillas University, and it is able to consider the electrical characteristics of the existing grid when evaluating grid extensions, and can check whether upstream upgrades to the existing grid are required. In the current implementation of REM, the brownfield model is not yet integrated into the logical flow and we instead use the greenfield RNM to design grid extensions. The greenfield RNM was also developed at IIT Comillas University (Peco González, 2001), and it can design distribution networks with a minimal representation and consideration of existing grid infrastructure. Despite the fact REM does not currently use the brownfield RNM, we have set up the data structure for the electrical information so that we are ready to integrate the brownfield RNM into REM in future work.

For the substations, each type of transformer present must be specified and the count of each of those transformer types. The transformer type is specified through the name of the transformer corresponding with the appropriate item in the Reference Network Model catalog (see the section on equipment catalogs below).

For the power lines, the conductor type must be specified. In the current implementation, each feeder is associated with one conductor type, where the conductor type corresponds to an item in the Reference Network Model catalog (see the section on equipment catalogs below). The conductor type, plus the lengths of feeder segments, can be used to calculate electrical parameters of the grid lines.

Information about the demand of current customers must also be included. In the current implementation, demand is associated with each medium voltage feeder. In a data processing step, this feeder demand is evenly dispersed among “imaginary” medium voltage customers located at the nodes of the medium voltage lines. The idea is that each of these imaginary customers represents a collection of real customers. This allows for a reasonable representation of existing demand on the medium and high voltage lines and transformers when finer resolution data is not available. Aggregating the real low voltage customers and low voltage lines into representative medium voltage customers also likely has the advantage of reducing the run time of the brownfield RNM.

The arrangement of electrical protection devices can be inferred from the topological description of the existing grid. A breaker is placed at the head of each high voltage and medium voltage feeder. A fuse is placed with medium-to-low voltage transformers.

Vaishali:

For the Vaishali study, information from NBPDCL was available with the number and size of transformers at each substation, the primary conductor types for each medium voltage feeder, and the peak and average demand on each feeder.

For detailed planning of required network upgrades, it would be important to know the type of conductor for each segment of each feeder, rather than just the primary conductor type for the feeder. This more detailed information was not readily available for Vaishali. If the information was available, the REM data files and structures would need to be modified to be able to account for this information.

2.2.3 Economic

The economic information associated with the existing grid entails the cost of procuring energy from the transmission grid plus the costs of distribution between the transmission substation and the start of the new grid extension. This is used to calculate the cost of energy supply to new grid extensions.

In the current implementation, this cost is represented by a volumetric charge in currency per kilowatt-hours. The cost is associated with medium voltage power lines, which are the lines to which grid extensions can connect in REM.

Vaishali:

The most recent tariff order for NBPDCL indicated that the average cost of supply for 11 kV services was estimated to be 6.39 INR per kWh for 2015-2016, or around $0.10 per kWh. This is the true cost to the distribution company, including purchasing power and all distribution and retail activities. However, these costs reflect the cost of supply to currently connected customers, who are disproportionately located in more densely populated areas relative to the currently non-connected customers. The cost of supply to remote rural areas is often dramatically higher than the average supply cost (Pérez-Arriaga, 2013). While a detailed study of the actual cost of supply to typical non-connected parts of Vaishali was not done, we used a value of $0.20/kWh to account for the fact that the cost of supply in these areas is likely higher than the reported average cost of supply.

In future work, the study would benefit from a more refined look at location-specific costs of supply from the existing grid.[5]

2.2.4 Service quality

Service quality is considered in terms of the availability of electricity supply from the grid. Availability information is associated with the medium voltage feeders. For each feeder, 24 availability numbers are given, representing the probability that grid power is available during the corresponding hour of the day. This allows the model to consider the effect of poorer reliability during certain times of day. Typically the availability of grid power will be poorest during peak hours, either due to supply limits or due to capacity constraints in transformers or conductors.

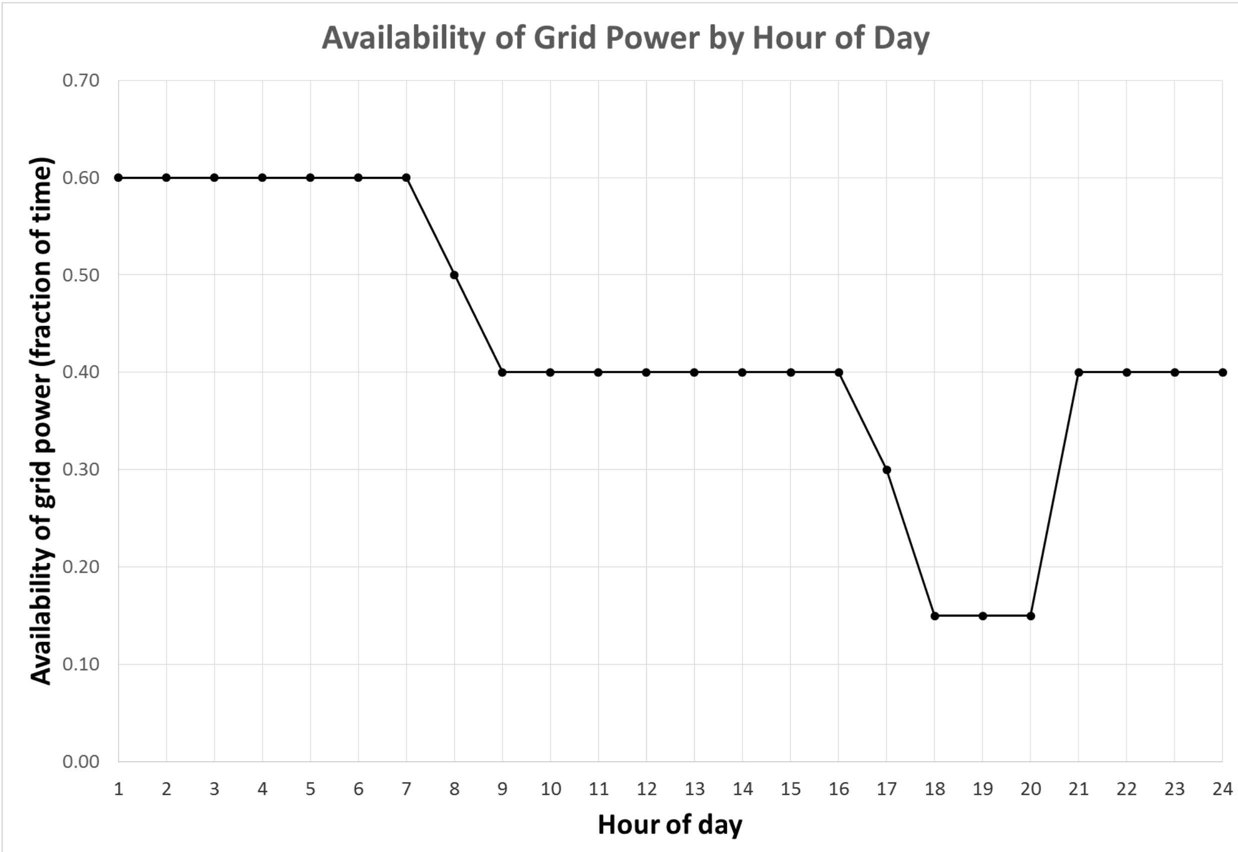
One limitation of this representation of availability is that it can only consider one typical or average day. In some cases, the total level of demand and the hours of peak demand may vary considerably over the course of a year.

Vaishali:

In Vaishali, reliability information can be inferred from the logbooks at the substations. In these logbooks, the demand for each feeder for each hour is recorded, including any outages. This data is recorded with pen and paper and may only exist at the sub-station itself.

Our team was able to copy several months of logbook data for a rural medium voltage feeder of the Hajipur substation. Our team analyzed this data and produced representative availability percentages for each hour of the day. These numbers are shown in the figure below and were assigned to all medium voltage feeders for our study.

[5] There is also a more fundamental question of whether this cost should only include the cost of electricity supply, or the combined cost of electricity supply and distribution. The decision of whether or not to include distribution costs here should be synchronized with the consideration of distribution costs in other parts of the model (especially the consideration of any upgrades to the existing grid that may be required).



The representation of reliability of power supply from the grid that we used for this study does not account for all of the features of reliability that we qualitatively know to exist. Some of these missing features are:

- Reliability was worse in the summer months of data and better in the winter months in the data we gathered. This may be due to a combination of higher temperatures (which lower the effective capacity of transformers and conductors) and higher electricity demand in the summer (due to a combination of seasonal trends and demand growth—the summer data was most recent). We have ignored this seasonal variation and included a single representative day of reliability data for the sake of simplifying the model and reducing the volume of input data.
- We have heard that reliability tends to be worse on feeders that are more rural and remote, but we were unable to directly collect information to quantitatively estimate this effect.
- Reliability at the customer level may be significantly worse than the medium voltage feeder reliability in some cases. This would be primarily due to failures of equipment between the head of the medium voltage feeder and the customer. From what we hear, burnout of medium-to-low voltage transformers is a common issue, which may lead to worse availability of grid power to the customer.
- Investments are being made in the generation, transmission, and distribution systems which are leading to a trend of improving availability of power from the main grid. Additionally, programs

to segregate agricultural feeders from residential feeders may improve the availability of power to residential customers.[6]

2.3 Customer Data

From the perspective of REM, a customer is any load or collection of loads requiring electricity service. In the rural electrification context, the majority of customers will typically be residences. Other customers, however, can also be considered, such as commercial buildings, agricultural loads, or street lighting.

The information required for each customer includes the location, the customer type, and whether the customer is already connected to the electric grid.

2.3.1 Customer location

The location of each customer must be specified. REM represents customer locations in terms of UTM coordinates. If customer location information is already available, this can be input directly to the model. If pre-existing information is not available, survey work may be required to obtain high-quality information about the location of customers. This survey work would typically be a small fraction of total project costs including implementation, but still could be a significant barrier to preliminary studies. In the case where there is insufficient time or financial resources to conduct a survey of customer locations, satellite image processing algorithms can be used to approximate the locations of relevant features, such as buildings. To get the best results from satellite image processing it may be necessary to buy high-resolution satellite images and do on-the-ground building identification in sample locations to produce training sets. However, useful results can still often be obtained from processing of free images from sources such as Google Earth with training sets produced from human visual inspection of the images.

Vaishali:

We did not have access to any comprehensive pre-existing data sources about customer locations in Vaishali, and due to the total number of customers (around 650 thousand based on the number of households in from the 2011 census) it was not feasible to do manual surveys to identify household locations. Thus we used a satellite image processing algorithm to identify household locations, using free images from Google Earth. The image processing algorithm used is being developed and improved by other members of the MIT Universal Access Research Group team, but the version used for this study was a very early version. Thus, the algorithm identified buildings in a way that seems to give a general indication of regional building density, but had a high error rate in terms of missing buildings and finding false positives. In addition to improving the image processing algorithm, it is likely that the accuracy of building identification could be improved by using higher quality images.

The algorithm identified 250 thousand buildings, which was significantly less than the 650 thousand we expect to be in Vaishali based on the 2011 census count of households in Vaishali. To give a more realistic representation of customer density we artificially added more buildings in a way that preserved the variations in building density found via the satellite image processing. Thus the additional 400

---

[6] The immediate impact of feeder segregation in Vaishali may be marginal since, from what we have heard, diesel pumps are most common in this area.

thousand buildings were assigned to 50 meter by 50 meter cells in proportion to the density of identified buildings in those cells, and then were randomly placed within the cells.

For some runs of the model, we were interested in the impacts of increased building density on the REM results. To produce the building data set for these runs, we added an additional 13% of buildings (bringing the total to approximately 735 thousand) using the same procedure described above.

2.3.2 Customer electrification status

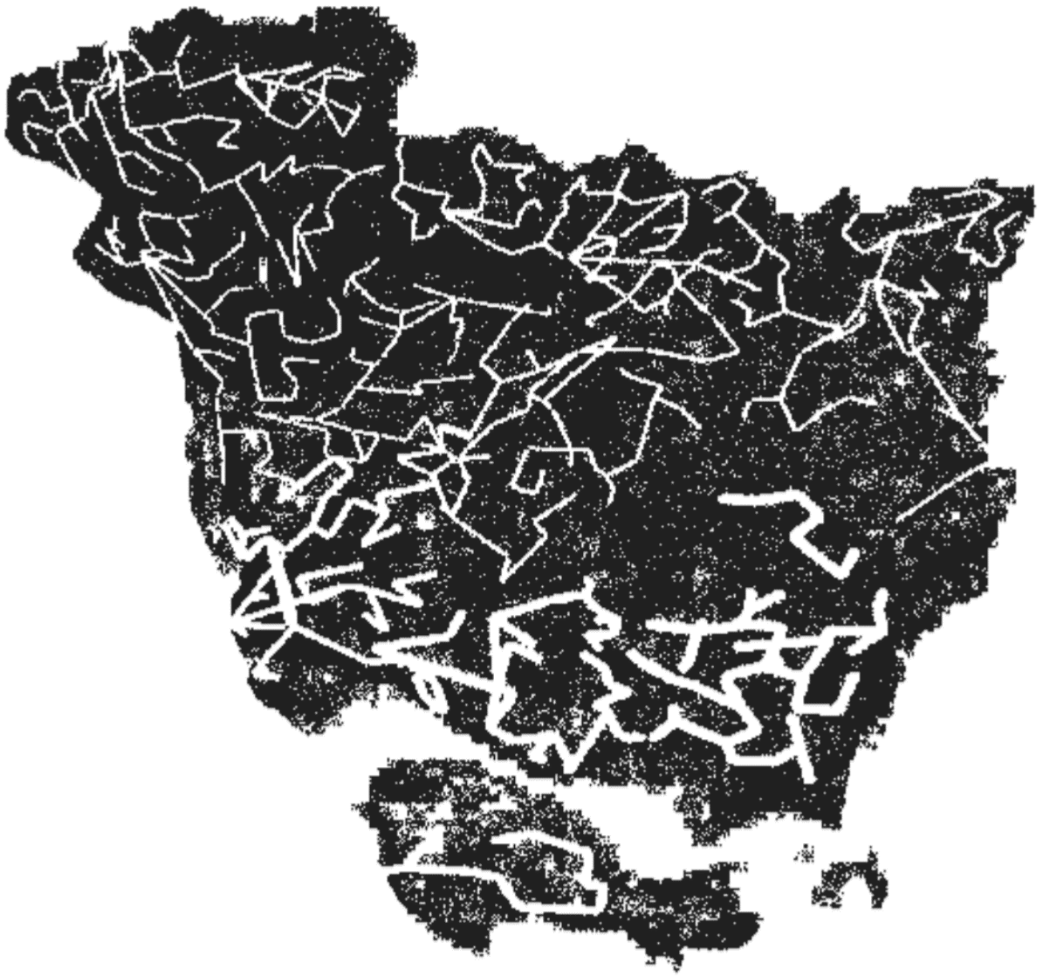
REM focuses on determining electrification modes and producing designs for customers which are currently not connected to the main electric grid.[7] Thus, it is necessary to designate which of the identified customers are already electrified or not.

If good information is available about which specific customers or which areas are electrified, this information can be directly used to designate which customers are electrified in a pre-processing step. If such information is not available, the location of customers which are already connected to the grid can be inferred from information about the existing grid lines. The paths of low-voltage lines would be best for this purpose, but even if this is not available, locations of medium-to-low voltage transformers or paths of medium voltage lines can be used to approximate where electrified customers are located.

Vaishali:

For Vaishali, we obtained information from NBPDCL about the number of current electricity customers by electrical subdivision. Each subdivision is an administrative unit of NBPDCL which is responsible for a subset of high-to-medium voltage substations, and the associated downstream infrastructure and customers. We used this information to produce buffers around the medium voltage feeders such that the number of buildings within the buffers corresponding to each subdivision matched the data. There were about 130 thousand current electrical customers, representing about a 20% electrification rate in Vaishali. The image below shows the remaining customers in black, after the customers within the buffers around the medium voltage lines were removed.

[7] Of course, many customers in the developing world who are nominally connected to the grid receive inadequate service. This is a very important issue, but one that REM does not address for now.



For the runs with additional customers added, the same buffers as the baseline run were used.

In reality, the medium voltage lines likely take less direct paths between the villages that they cover, and not all customers near to the lines will be connected. But given the availability of data, this approach gives a reasonable representation of where the customers connected to the grid are likely to be located.

2.3.3 Customer type

Each non-electrified customer is assigned a customer type, which identifies the customer with a description of its electricity demand.

Each customer type is associated with a function which can build a profile of 8760 hours (1 year) of electricity demand. For each hour of the year, the customer gets two demand numbers—one for critical demand and one for non-critical demand—each of which is associated with a different cost of non-served energy.[8] The functions that create demand profiles include some randomization in the production of profiles, so that successive profiles produced for individual customers with the same

[8] The cost of non-served energy notionally represents a cost incurred by the customer when electricity service is not available. This could represent a combination of financial costs to the customer and lost utility. From a practical perspective, the non-served energy cost drives balance between higher quality service and higher costs, as the model aims to minimize the total cost of service plus non-served energy cost.

customer type will be slightly different. This allows the model to consider the benefit of aggregating customers into larger systems—if demand profiles are varied, the peak power demand for a group of customers will generally be less than the sum of the individual customer peaks. More details on the specification and creation of demand profiles are provided in the section about creating demand profiles.

In the current implementation, only one customer type can be considered in each run of REM for a study area. In future developments, the model could be extended to consider multiple customer types. The majority of the modification would be in producing the lookup table of pre-solved local generation designs for off-grid systems, which is described in a later chapter.

Vaishali:

For this study, we assumed all non-electrified customers to be "typical" rural residential customers. The assumption that all customers are residential customers is not as great of a distortion as it might initially appear, because NBPDCL data shows that about 97% of their current customers in Vaishali are residential. If the reliability of the grid improves, it is likely that there would be more agricultural demand as farmers switch from diesel to electric pumps.

2.4 Energy resource and weather data

Information about local energy resources is required to produce designs and estimate costs of local generation for off-grid systems. In the current implementation of REM two types of energy sources are considered: solar photovoltaic (PV) and diesel generator sets.

2.4.1 Solar PV and weather data

For solar PV, data is required for the hourly DC output of a 1 kW PV array for one year. This solar resource data can be multiplied by the size of a solar array to give the hourly output of that array. PVWatts is a source managed by the United States' National Renewable Energy Laboratory (NREL) which can provide such data for many locations around the world. PVWatts provides solar irradiance and weather for a "typical meteorological year" and uses that information to calculate hourly output of a solar array. The output calculation is based on the irradiance and the calculated temperature of the PV cells. The weather data which comes with the PV output data—including irradiance and temperature—can also be used for demand modeling.

In the current model implementation, PV and weather data are associated with sub-districts. This allows for some regional variation in solar and weather patterns to be considered in the model.

Vaishali:

For the Vaishali study, one PVWatts file with hourly data was obtained for each of the 16 sub-districts, using a geographical location near the center of the sub-district. The standard installation and mounting assumptions in PVWatts were used.

2.4.2 Diesel fuel data

The cost of diesel fuel is an important input for calculating the cost of systems including diesel generators. In general, pump prices of diesel might vary over a study area, and additionally there could be significant transportation costs to get diesel fuel to more remote areas.

In the current implementation of the model, only a single diesel price is considered.

The MIT Universal Energy Access Research Group team has looked into adjusting local diesel prices based on estimated travel times from major towns or cities. However, the model run time became excessive when each off-grid system was designed with a separate local diesel fuel price, so we returned to using a single price. In the future, improvements could be made by factoring remoteness into the cost of diesel fuel, as well as considering remoteness in other system costs.

Since the cost of diesel is uncertain and variable over the lifetime of a project, some judgment should be used when selecting a diesel price and it may be beneficial to run scenarios with different diesel prices to get a sense of how sensitive the results are to this parameter.

Vaishali:

In the Vaishali study, we used a diesel fuel cost of $1/L. This cost was obtained as the local diesel fuel price for Patna, Bihar (the nearest large city to Vaishali) from late 2014 from the website mypetroprice.com. In recent months the price of diesel has dropped to around $0.8/L in response to the drop in global oil prices, but projections (for example EIA's Annual Energy Outlook, 2015) predict a rebound in oil prices in the coming years.

2.5 Equipment Catalogs

Information about costs and technical characteristics of available system components are important inputs for REM. This information is broken up into two catalog files: a local generation catalog and the Reference Network Model catalog.

2.5.1 Local generation catalog

The local generation catalog contains information about the electricity supply site for off-grid systems. The following table lists the items considered in the local generation catalog:

Item	**Notes**
Solar photovoltaic (PV) panels	Two sizes of solar panels can be specified as constituents of PV arrays. The smaller should be of a size appropriate for a single-customer system. The larger should be of a size appropriate for larger microgrids. For each panel, the following information is required: • Size (kW) • Cost ($) • Life (years) • Installation cost as a fraction of panel cost • Annual O&M as a fraction of panel cost • Annual O&M man-hours • Annual capacity loss (fraction)
Batteries	Two sizes of batteries can be specified as constituents of energy storage systems. The smaller should be of a size appropriate for a single-customer system. The larger should be of a size appropriate for larger microgrids. For each battery, the following information is required: • Cost ($)

	• Initial State of Charge (fraction) • Capacity at end of life (fraction) • Installation cost as a fraction of battery cost • Annual O&M as a fraction of battery cost • Annual O&M man-hours • Kinetic battery model parameters[9] • Lifetime energy throughput (kWh)
Diesel generators	A list of diesel generators of various sizes is included in the catalog. For each generator, the following parameters are required: • Generator size (kW) • Fuel consumption at ¼, ½, ¾, and full load (gal/hr) • Minimum power (kW) • Startup fuel (L) • Lifetime (h) • Cost ($) Additionally, the following parameters apply to all of the diesel generators: • Installation cost as a fraction of generator cost • Annual O&M as a fraction of generator cost • Annual O&M man-hours
Inverter/Rectifier	The following parameters are included for the inverter/rectifier: • A series of converter sizes (kW) paired with costs ($/kW) • Minimum size (kW) • Lifetime (years) • Inverter efficiency (fraction) • Rectifier efficiency (fraction) • Ratio of rectifier capacity to inverter capacity • Installation cost as a fraction of converter cost • Annual O&M as a fraction of converter cost • Annual O&M man-hours
Charge controller	The following parameters are included for the charge controller: • A series of charge controller sizes (kW) paired with costs ($/kW) • Minimum size (kW) • Lifetime (years) • Efficiency (fraction) • Installation cost as a fraction of charge controller cost • Annual O&M as a fraction of charge controller cost • Annual O&M man-hours
Other	These parameters are not associated with specific pieces of equipment: • O&M labor cost ($/hour) • Additional costs per system ($/year)

[9] The kinetic battery model is described in (Manwell & McGowan, 1993).

Vaishali:

The following tables list the parameters and values included in the generation catalog for the Vaishali study, along with some explanation of how the values were obtained.

- Other

O&M Labor cost ($/hr)	1.46
Cost Per System ($/yr)	0

These "other" parameters are numbers used in local generation design and cost calculation but are not associated with a particular system component. The O&M labor cost is based on 2010 hourly costs of manufacturing in India, as reported by the United States Bureau of Labor and Statistics in their International Comparisons of Hourly Compensation Costs in Manufacturing, 2012.[10]

No additional "per-system" cost was included in this study.

- Battery (1 of 2)

Battery Type	Cost ($)	SOC Init. (fraction)	Capacity at end of life (fraction of nameplate energy capacity)	Installation Costs as fraction of battery cost	Annual O&M as a fraction of battery cost	Annual O&M man-hours
TROJ_T105	150	0.3	0.8	0.2	0.02	5
VIS_CP12240D	60	0.4	0.8	0.2	0.02	5

- Battery (2 of 2) – kinetic battery model parameters

Battery Name	**energy**	**SOC min**	**SOC max**	**eff_batt _d_c**	**Nbatt _rt**	**alpha _c**	**Imax**	**Vnom**	**Lifetime Energy**	**c**	**k**
Units	**kWh**					**A/Ah**	**A**	**V**	**KWH**		**1/hr**
TROJ_ T105	1.38	0.3	1	0.92	0.85	1	11	6	845	0.28	1.85
VIS_ CP12240D	0.28	0.4	1	0.89	0.8	1	9.6	12	103	0.33	2.38

Two batteries were included in the battery catalog. Trojan T105 is a flooded lead-acid battery and the Vision CP12240D is a sealed lead-acid battery. The Vision battery was selected as a relatively cheap smaller battery that could be used in home systems, and the Trojan battery was selected as a battery

[10] Available at http://www.bls.gov/fls/ichcc.htm.

with low cost for larger systems. These batteries were also selected because the kinetic battery model (Manwell & McGowan, 1993) parameters for these batteries were available in the HOMER catalog.[11]

The unit costs of the batteries were obtained from online price quotes from retailers such as wholesalesolar.com. The initial state of charge was set to the minimum states of charge for the respective batteries.[12] The additional battery parameters were estimated based on the modeler's judgment, but could be improved if data were available.

- Solar

Size (kW)	Cost ($)	Life (years)	Installation Costs as fraction of panel cost	Annual O&M as a fraction of panel cost	Annual O&M man-hours	Annual capacity loss (fraction)
0.25	225	25	0.65	0.01	1	0.007
0.02	50	25	0.25	0.01	1	0.007

The 20W solar panel was selected as a relevant panel size for home systems and the 250W solar panel was selected for a relevant panel size for larger systems. The costs of these panels were obtained from online price quotes from retailers such as wholesalesolar.com. The 25 year lifetime was selected as a common warranty period for solar panels. The annual capacity loss was estimated based on the analysis in (Jordan & Kurtz, 2013). The solar installation costs were estimated based on the summary of a 2012 report on solar balance of system costs done by GTM Research and Solvida Energy Group, Inc.[13] The operations and maintenance numbers were estimated based on the modeler's judgment, but could be improved if data were available.

- AC-DC Converter

Costs ($/kW)	927	740	600	543	364	319	260	220	190	190
Sizes (kW)	0.15	0.2	0.25	0.3	1	1.5	5	6	10	11.4
Min Size (kW)	0.15									
Life (years)	15									
Inverter Efficiency (fraction)	0.95									
Rectifier Efficiency (fraction)	0.9									
Rectifier Capacity / Inverter Capacity Ratio	0.8									

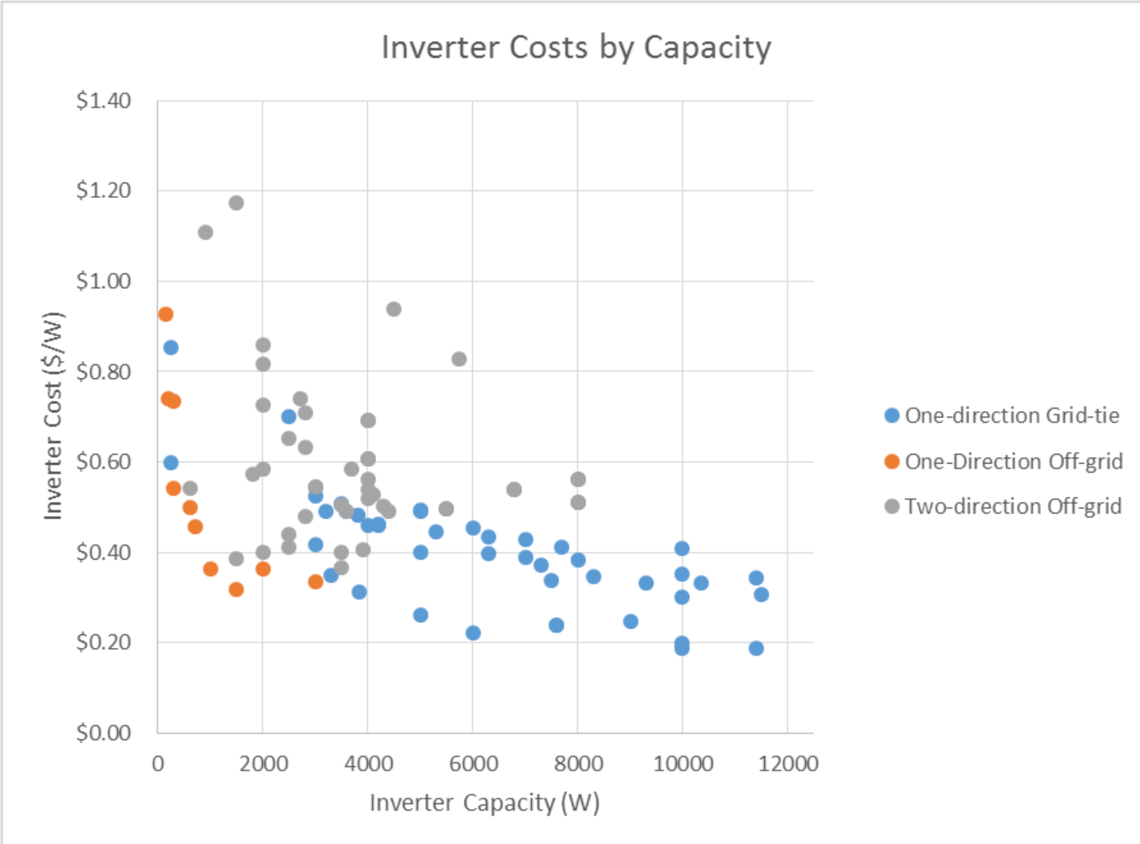
[11] More information about the HOMER model is available at www.homerenergy.com.
[12] The initial state of charge of the battery significantly influences the operation of the system in the first part of the simulation. If the total number of hours included in the simulation is long enough, this effect may not dramatically impact local generation designs or cost estimates. See the discussion of local generation design in the next chapter. It might be worthwhile in future work to modify the simulation so that it is less influenced by these first hours—perhaps by allowing the battery cycling to reach something like "steady state" before starting the "official" simulation.
[13] Available at http://www.greentechmedia.com/research/report/solar-pv-bos-2013.

Installation Costs as fraction of converter cost	0.1
Annual O&M as a fraction of converter cost	0.01
Annual O&M man-hours	2

The selection of data points for inverter sizes and costs were chosen based on data about available inverters from wholesalesolar.com in March 2015. The full set of costs and sizes of inverters that were available is shown in the figure below. The range of available prices and sizes of inverters varies substantially with the inverter features. The chart below indicates which inverters are one-direction grid-tie, one-direction off-grid, and two-direction off-grid. Additionally, other features like the range of control options, efficiency, warranty, and interoperability with other systems distinguish inverters. The guiding principle was to select prices for the lowest cost available inverters at a variety of sizes, but in practice, it may be beneficial to use more expensive inverters in some cases.

The inverter minimum size was selected as the smallest inverter available on wholesalesolar.com. The inverter life of 15 years was chosen based on discussions with solar system providers. However, warranties for inverters tend to be significantly shorter than 15 years.[14] Inverter efficiency, rectifier efficiency, and the rectifier to inverter capacity ratio were chosen as representative values based on looking at specification sheets for many inverters, but in practice these values vary. The parameters regarding installation costs and operations and maintenance were estimated based on the modeler's judgment, but could be improved if data were available.

[14] Some solar providers in India also repair inverters to extend their life.



- Charge controller

Costs ($/kW)	481	375	283	215	133	131
Sizes (kW)	0.054	0.12	0.24	1.44	3.84	4.128
Min Size (kW)	0.54					
Life (years)	15					
Efficiency (fraction)[15]	0.95					
Installation Costs as fraction of charge controller cost	0.1					
Annual O&M as a fraction of charge controller cost	0.01					
Annual O&M man-hours	2					

The charge controller sizes and costs were also selected based on available products at wholesalesolar.com. It is tricky to select a set of costs and sizes for two main reasons.

First, the prices and available capacities vary significantly with the capacities of the charge controller—the primary distinguishing feature being whether or not maximum power point tracking (MPPT) is

[15] Average efficiency

included. This feature impacts the amount of energy which can be obtained from the solar panels. In the current implementation, REM ignores this effect, and assumes that all charge controllers can maintain the panel at the maximum power point.

Secondly, charge controllers can often operate at multiple battery bank voltages, with power capacity being roughly proportional to the battery bank voltage for a given charge controller. Thus the “power capacity” of a charge controller is not well-defined unless it is paired with a specific battery voltage.

My general approach in choosing the data points for sizes and costs was to go for the lowest cost available at a variety of power capacities. In reality, the charge controller cost may imply a battery bank voltage that is not feasible for a given project, so the charge controller cost could be higher than estimated.

I did not have information about charge controller lifetime, so I assumed that is was similar to the inverter lifetime. The efficiency value is based on looking at charge controller spec sheets. In absence of better information, the installation costs and operation and maintenance cost parameters were chosen to match the values for the inverter.

- Diesel generator (1 of 2)

Generator Size (kW)	1/4 Load (l/kWh)	1/2 Load (l/kWh)	3/4 Load (l/kWh)	Full Load (l/kWh)	No-load (l/h)	Min power	Lifetime (h)	Cost (USD)	Startup fuel (l)
1	0.45	0.34	0.33	0.30	0.03	0.1	25000	200	0.0007
2	0.45	0.34	0.33	0.30	0.06	0.2	25000	400	0.0015
3	0.45	0.34	0.33	0.30	0.09	0.3	25000	600	0.0022
4	0.45	0.34	0.33	0.30	0.12	0.4	25000	800	0.0029
5	0.45	0.34	0.33	0.30	0.15	0.5	25000	1000	0.0036
6	0.45	0.34	0.33	0.30	0.18	0.6	25000	1200	0.0044
7	0.45	0.34	0.33	0.30	0.21	0.7	25000	1400	0.0051
8	0.45	0.34	0.33	0.30	0.24	0.8	25000	1600	0.0058
9	0.45	0.34	0.33	0.30	0.27	0.9	25000	1800	0.0065
10	0.45	0.34	0.33	0.30	0.30	1	25000	2000	0.0073
15	0.45	0.34	0.33	0.30	0.45	1.5	25000	3000	0.0109
20	0.45	0.34	0.33	0.30	0.61	2	25000	4000	0.0145
30	0.66	0.45	0.40	0.37	1.10	3	25000	6000	0.0595
40	0.61	0.44	0.40	0.38	1.51	4	25000	8000	0.1279
60	0.45	0.37	0.32	0.30	1.82	6	25000	12000	0.2085
75	0.48	0.34	0.31	0.31	2.31	7.5	25000	15000	0.3348
100	0.39	0.31	0.29	0.28	2.80	10	25000	20000	0.4909
125	0.38	0.30	0.29	0.28	3.44	12.5	25000	25000	0.7078
135	0.37	0.30	0.28	0.27	3.71	13.5	25000	27000	0.8745
150	0.36	0.30	0.28	0.28	4.13	15	25000	30000	1.0974
175	0.35	0.29	0.28	0.27	4.81	17.5	25000	35000	1.4241
200	0.36	0.29	0.28	0.27	5.45	20	25000	40000	1.7796
230	0.35	0.29	0.27	0.27	6.28	23	25000	46000	2.2415

250	0.35	0.29	0.27	0.27	6.81	25	25000	50000	2.6366
300	0.34	0.29	0.27	0.27	8.14	30	25000	60000	3.3955
350	0.34	0.28	0.27	0.27	9.50	35	25000	70000	4.2514
400	0.34	0.28	0.27	0.27	10.83	40	25000	80000	5.1717
500	0.33	0.28	0.27	0.27	13.51	50	25000	100000	6.8643
600	0.33	0.28	0.26	0.27	16.20	60	25000	120000	8.7195
750	0.33	0.28	0.26	0.27	20.21	75	25000	150000	11.4905
1000	0.33	0.28	0.26	0.27	26.91	100	25000	200000	16.1132
1250	0.33	0.27	0.26	0.27	33.61	125	25000	250000	21.1412
1500	0.33	0.27	0.26	0.27	40.31	150	25000	300000	26.5745
1750	0.32	0.27	0.26	0.27	47.01	175	25000	350000	32.4132
2000	0.32	0.27	0.26	0.27	53.71	200	25000	400000	38.6572
2250	0.32	0.27	0.26	0.27	60.41	225	25000	450000	45.3065

- Diesel generator (2 of 2)

Installation Costs as fraction of generator cost	Annual O&M as a fraction of generator cost	Annual O&M man-hours
0.8	0.05	730

The diesel generator data includes information about diesel generators at a range of sizes. The starting point for this table was previous student work, and values were updated based on a rough check of online diesel generator price quotes (for example from generatorsales.com). While the smallest generators available in the catalog are 1 kW, I was not able to find diesel generators for sale below 3 kW. However, gasoline generators are available in this size range (down to 1 kW). If the off-grid system requires heavy usage of the generator or remote start capability, these small gasoline or diesel generators may not be appropriate. Thus designs with small diesel generators should be treated with caution.

Information about generator installation cost was obtained from (Electric Power Research Institute, 2003).

2.5.2 Reference Network Model catalog

Information about distribution network components are included in the Reference Network Model (RNM) catalog. This catalog also includes parameters of the Reference Network Model algorithms. The catalog format and contents are described in detail in the Greenfield Reference Model User Manual. The following table lists the main categories of input data included in the RNM catalog file used by REM and some considerations regarding them for REM.

RNM catalog category	Notes
Conductors	A selection of high voltage, medium voltage, and low voltage conductors are included. Required data includes technical parameters, costs, and

	failure rates. This information must also be copied to the RNM catalog brief spreadsheet file, for use in other REM functions.
Transformers/Substations	A selection of transformers for high-to-medium voltage substations and for medium-to-low voltage transformation centers are included. Required data includes technical parameters, costs, and failure rates. This information must also be copied to the RNM catalog brief spreadsheet file, for use in other REM functions.
Other categories	Include: • Network parameters • Reliability requirements[16] • Reliability system costs • Capacitors • Voltage regulators • Costs related to slope, trenches, and supports • Parameters related to street mapping • Model configuration

RNM assumes balanced three phase power distribution, and thus the costs and technical parameters should refer to installation of three phase distribution equipment by default. It is also possible to include information representing single phase distribution, but in this case RNM may produce infeasible or unbalanced designs. The ability to better consider single phase distribution in RNM would be an important area for future work.

By default, RNM considers a separate set of conductor and transformer parameters for urban above ground, urban underground, and rural above ground. Given that REM is primarily focused on rural electrification, we have used the convention of providing the same set of conductor and transformer information for all of these categories, so that they are all effectively treated as rural above ground equipment. Model parameters regarding trenches and support infrastructure must be set in a particular way so that this interpretation is internally consistent.

REM uses just one catalog of distribution network equipment, which is used for both the grid extension and microgrid design cases. This is consistent with the perspective that basically the same equipment and installation practices would be used in both cases, so that the microgrid distribution infrastructure is matching the grid code. In future work, it would be interesting to explore the impact of providing a different set of distribution network options for microgrids, potentially representing lower-cost installation methods. A discussion of using lower cost distribution norms for rural electrification, which has been applied in Brazil, is contained in (Bornstein, 2007).

The ability to consider DC distribution in the catalog and model would also be a beneficial area for future work, especially since many deployed microgrids in the developing world use DC distribution.

[16] The RNM consideration of reliability here pertains to failure to serve demand due to failures of equipment. This is distinct from the REM consideration of reliability of power supply from the grid, which is often related to intentional curtailment rather than equipment failure.

The RNM catalog includes many parameters regarding considerations which are difficult to get information about in the developing world context and which may not neatly apply in the same way they do in the developed world context. In future work, it might be beneficial to create a version of RNM and the supporting catalog more catered to the developing world rural electrification context.

Vaishali:

As a starting point for the RNM catalog, we used a complete catalog employed for previous studies in Spain approximately 5-10 years ago. Values from that catalog were used as defaults where better information from Vaishali or India was not available.

From that starting point, significant changes were made in the listing of conductors and transformers, as well as in some parameters regarding how the model executes. Where available, information about types of components used and their costs from NBPDCL was used as the primary source. Information about project costs from Tata Power Delhi Distribution Limited—a distribution company in New Delhi—was used to supplement this as required. A lot of estimation, judgement, and assumption was required to complete the catalog entries, as there was not complete information available. The values used in these areas are presented below.

Low-voltage conductors:

Name	Resistance [ohm/km]	Reactance [ohm/km]	Rated current [A]	Overload [p.u.][17]	Min. failure rate [failures/(km*a)]	Max, failure rate [failures/(km*a)]	Av. failure rate [failures/(km*a)]	Overnight cost [$]	Predictive maintenance cost [$/(year*km]	Corrective maintenance cost [$/failure]
LV_weasel _singlephase	1.77	0.8	49	1.2	0.113	0.113	0.113	2227	2.8	427
LV_weasel _threephase	0.885	0.4	146	1.2	0.113	0.113	0.113	3341	2.8	427
50_mm2	0.783	0.4	154	1.2	0.133	0.133	0.133	10858	2.8	427
70_mm2	0.542	0.4	196	1.2	0.133	0.133	0.133	13820	2.8	427
95_mm2	0.392	0.4	242	1.2	0.133	0.133	0.133	17064	2.8	427
120_mm2	0.31	0.4	282	1.2	0.133	0.133	0.133	19884	2.8	427
150_mm2	0.253	0.4	322	1.2	0.133	0.133	0.133	22705	2.8	427
185_mm2	0.202	0.4	374	1.2	0.133	0.133	0.133	26371	2.8	427
240_mm2	0.155	0.4	446	1.2	0.133	0.133	0.133	31448	2.8	427
300_mm2	0.125	0.4	514	1.2	0.133	0.133	0.133	36243	2.8	427
400_mm2	0.0981	0.4	601	1.2	0.133	0.133	0.133	42378	2.8	427

[17] P.u. means "per unit." It is basically percent divided by 100.

Medium Voltage Conductors:

Name	Resistance [ohm/km]	Reactance [ohm/km]	Rated current [A]	Overload [p.u.]	Min. failure rate [failures/(km*a)]	Max, failure rate [failures/(km*a)]	Av. failure rate [failures/(km*a)]	Overnight cost [$]	Predictive maintenance cost [$/(year*km]	Corrective maintenance cost [$/failure]
MOLE	2.63	0.4	75	1.2	0.168	0.168	0.168	1648	700	900
SQUIRREL	1.34	0.4	114	1.2	0.168	0.168	0.168	2505	700	900
WEASEL	0.885	0.4	146	1.2	0.168	0.168	0.168	2975	700	900
RABBIT	0.529	0.4	197	1.2	0.168	0.168	0.168	4122	700	900
DOG	0.268	0.4	312	1.2	0.168	0.168	0.168	6855	700	900
WOLF	0.178	0.4	406	1.2	0.168	0.168	0.168	8920	700	900

Medium-to-low voltage transformers:

Name	Installed power capacity (kVA)	Guaranteed power capacity (kVA)	Voltage	No load losses (kW)	Resistance on the lower side of the transformer (ohms)	Maximum number of output ports	Investment cost per output port ($/port)	Failure rate min (failures/year)	Failure rate med (failures/year)	Failure rate max (failures/year)	Investment cost ($)	Preventative maintenance cost ($/year)	Corrective maintenance cost ($/failure)
CTI1	16	16	11	0.12	0.4	2	0	0.007	0.007	0.007	808	230	1500
CTI2	25	25	11	0.14	0.3	2	0	0.007	0.007	0.007	1263	230	1500
CTI3	40	40	11	0.15	0.08	2	0	0.007	0.007	0.007	2020	230	1500
CTI4	63	63	11	0.26	0.045	2	0	0.007	0.007	0.007	3030	230	1500
CTI5	100	100	11	0.38	0.035	2	0	0.007	0.007	0.007	3781	230	1500
CTI6	200	200	11	0.58	0.01	2	0	0.007	0.007	0.007	10101	230	1500

2.6 SETTINGS FILE

Some of the parameters discussed above and some additional parameters are input via a settings spreadsheet. The relevant fields in the settings spreadsheet are:

- Cost Energy Not Served - Normal ($/kWh)
- Cost Energy Not Served - Critical ($/kWh)
- Discount Rate (fraction)
- High Voltage Level (secondary of transmission SS) (kV)
- Medium Voltage Level (kV)
- Load Voltage Level (kV)
- Diesel Fuel Cost (Baseline) ($/L)
- Distribution system losses (fraction)
- Network Lifetime (years)
- Per Customer Cost Isolated ($)
- Per Customer Cost Microgrid ($)
- Per Customer Cost Grid Extension ($)
- Demand growth rate (fraction)
- Final year to consider for generation design
- Per Customer Investment Lifetime (years)
- Per Customer annual O&M Isolated ($/yr)
- Per Customer annual O&M Microgrid ($/yr)
- Per Customer annual O&M Grid Extension ($/yr)

Vaishali:

The following table list the settings values used for the Vaishali study in the baseline case. Additionally, a case was run with a higher annual demand growth rate (16% instead of 1%).

Cost Energy Not Served - Normal ($/kWh)	1.50
Cost Energy Not Served - Critical ($/kWh)	2.00
Discount Rate (fraction)	0.10
High Voltage Level (secondary of transmission SS) (kV)	33.00
Medium Voltage (kV)	11.00
Load Voltage (kV)	0.40
Diesel Fuel Cost (Baseline) ($/L)	1.00
Distribution system losses (fraction)	0.05
Network Lifetime (years)	40.00
Per Customer Cost Isolated ($)	50.00
Per Customer Cost Microgrid ($)	50.00
Per Customer Cost Grid Extension ($)	50.00
Demand growth rate (fraction)	0.01
Final year to consider for generation design / Number of years to consider for project	5.00
Per Customer Investment Lifetime (years)	40.00
Per Customer annual O&M Isolated ($/yr)	1.00
Per Customer annual O&M Microgrid ($/yr)	1.00
Per Customer annual O&M Grid Extension ($/yr)	1.00

2.7 REM EXECUTION SCRIPT

REM is called via a Matlab script. In this script, some additional model configuration information and parameters are specified.

First, the user must specify the locations of relevant folders, including the folder that contains the REM code and the folder which contains the REM data and outputs.

Next, the user must specify over which analysis regions the model will produce input structures or run the main REM module. For a given analysis region, it is necessary for the input structures to be created before the module which produces results can be run. It is possible to copy the input files for any analysis regions to another computer and run the main REM module for those regions on that computer.

The user can also specify whether the model should run in parallel mode, and if so the number of parallel processes. In parallel mode, some parts of the model execution are split up over the workers and run simultaneously. This can allow large study areas to be solved much more quickly. The number of workers is limited by the number of threads on the user's computer. Also, the user should be careful to check that the parallel operation does not overwhelm the computer's memory resources.

Finally, some model parameters can be set in the execution script. These are listed in the table below.

Model parameter	**Values**	**Notes**
config.setGridRel	Logical (true or false)	If config.setGridRel is set to true, all grid reliability data is overwritten with the value in config.gridRel.
config.gridRel	Number between 0 and 1	
pDuration	Integer	When sampling hours of the year for the local generation simulation, nPeriods number of blocks of pDuration number of hours are selected out of the 8760 hour year. The product of nPeriods and pDuration should be no greater than 8760.
nPeriods	Integer	
config.ntrialsGenStudy	Integer	The number of times to produce a local generation design for each example microgrid size in the generation pre-solving step. The set of results for each microgrid size are averaged.
config.nProfiles	Integer	The number of full demand profiles to produce for the demand library for each analysis region. The profiles for individual customers will be sampled from this library.
config.nMGSizesGenStudy	Integer	The number of microgrid sizes that will be included in the generation pre-solving step. The sizes selected are log distributed between 1 customer and the max number of customers in any analysis region. Some smaller size microgrids are automatically included also.

Vaishali:

For the Vaishali study, for each model case, all of the 30 analysis regions were solved in parallel. The following tables list the parameters set in the execution script by case. The only difference is that in the reliable grid case, the grid reliability data is overwritten with perfect reliability.

All scenarios except reliable grid:

Model parameter	**Values**
config.setGridRel	false
config.gridRel	Not applicable
pDuration	24*14 (i.e. 2 weeks)
nPeriods	6
config.ntrialsGenStudy	3
config.nProfiles	100
config.nMGSizesGenStudy	10

Reliable grid scenario:

Model parameter	**Values**
config.setGridRel	true
config.gridRel	1
pDuration	24*14 (i.e. 2 weeks)
nPeriods	6
config.ntrialsGenStudy	3
config.nProfiles	100
config.nMGSizesGenStudy	10

2.8 Reading source files into Matlab structures

Once all of the data is provided in the REM source data files, and the REM execution script is called, the first step of REM is to read the source data files and write the information to Matlab structures, which can be used for later REM design algorithms.

One folder of input Matlab structures is created for each analysis region, so that the analysis regions can be solved independently—either sequentially, or in parallel on one computer, or split up amongst several computers. In general, larger analysis regions will produce better quality results, since there would be less artificial barriers which systems cannot cross. However, the user of REM may find it desirable or necessary to divide sub-districts into smaller analysis regions to make the best use of available computing resources.

Each region's input folder contains all of the information necessary to run the REM algorithms on that region. That means that information which is common to the entire study area is included copied to each region. This common information includes the settings, the local generation catalog, the existing grid data, and the input parameters required to create demand profiles.

Other information is associated with sub-districts. This includes the PVWatts files with hourly solar and weather data for a typical meteorological year. Using this PVWatts data and other study area data, a demand profile library and a lookup table of pre-solved local generation designs are created for each sub-district. The processes of creating these items are described below. The PVWatts data, demand profile library, and pre-solved generation design information for each subdistrict are all copied into the input files for all analysis regions within the subdistrict.

Finally, some information is specific to each analysis region, and can directly be included in the corresponding input folder. This includes the location of each nonelectrified customer in the analysis region.

2.9 Creation of Demand Profile Library

For each sub-district, a sample of demand profiles is created and stored in a demand profile library. In the current implementation, all of the profiles correspond to a single customer type and differences between the profiles correspond to random variations. In the future, the capability to consider a variety of customer types in a single run could be added.

Each demand profile consists of a two series of 8760 numbers, corresponding to the quantity of critical and non-critical electricity demand for each hour of the year, in kilowatt-hours.[18] The total number of demand profiles to produce can be set in the execution script. A sufficient number of profiles should be saved to adequately represent the variability in the profiles. On the other hand, limiting the number of profiles reduces the memory requirements for the model and can allow the model to run faster.

The demand profile can be thought of as the amount of electricity the customer would consume in each hour of the year if the electricity service availability was not a limiting factor. However, the power systems designed in REM are not required to serve all of this demand. In the case of off-grid systems, depending on the costs associated with not serving critical and non-critical demand, a more cost effective solution can be to not serve some demand rather than increasing the size of system components. In the case of grid extensions, the upstream grid may not make power available to the connected customers all of the time.

To produce a demand profile the following information is needed:

- The data in the demand inputs spreadsheet, paired with a CreateDemandProfiles Matlab function,
- The PVWatts weather and solar data,
- The annual demand growth rate and design final year, from the settings.

The demand inputs spreadsheet contains the primary description of demand. The spreadsheet contains a series of rows, each corresponding with an electricity consuming activity. Each activity comes with the following information:

- Critical or non-critical – The cost of non-served energy that should be incurred if this demand cannot be served depends on this parameter.
- Energy – The amount of energy demand that should be added to an hour when this activity is occurring.
- Available hours – The hours of the day is it possible for this activity to happen.

[18] This can equivalently be thought of as the average demand in kilowatts for each hour.

- Availability restrictions – Any criteria that remove an hour from being available. For example, this could correspond to the level of solar irradiance (for lighting demand) or the temperature (for cooling demand).
- Average daily duration – The average number of hours that customers have this activity occurring each day.
- Average daily duration criteria – As an alternative to a flat average daily duration number, average daily duration can be calculated based on specified criteria for each day. For example, it could be calculated as the number of hours that temperature is above a certain level.
- Variability in daily duration – a parameter which describes the variation in number of hours the activity is occurring between customers on a given day.

Additionally, there is one parameter corresponding to the customer type, rather than a specific activity:

- Daily variability – A parameter similar to variability in daily duration, but corresponding to variations in the number of hours of all activities, rather than a particular one.

To produce a demand profile for a single customer for a single day, first the available hours for each activity are determined based on the availability parameters for that appliance and the PVWatts data. Then the target number of hours for each activity is calculated by adjusting the average daily duration for that appliance (which is calculated from the daily duration parameters and the PVWatts data) with a per-activity variability correction and a per-customer variability correction, as indicated below:

$$NHT_{c,a,d} = NHA_{a,d} \times \left(1 + V_a U([-1,1])_{a,d} + V_{ct} U([-1,1])_{c,d}\right),$$

Where:

$$NHT_{c,a,d} = target\ number\ of\ hours\ for\ a\ customer, activity, day$$

$$NHA_{a,d} = average\ number\ hours\ for\ an\ activity, day$$

$$V_a = variability\ parameter\ for\ an\ activity$$

$$V_{ct} = variability\ parameter\ for\ customer\ type$$

$$U([-1,1])_{a,d} = uniformly\ distributed\ random\ number\ for\ an\ appliance, day$$

$$U([-1,1])_{c,d} = uniformly\ distributed\ random\ number\ for\ a\ customer, day.$$

Then for each of the available hours for each appliance, demand is assigned to that hour with probability of the target number of hours divided by the number of available hours.

This procedure is repeated for all appliances for all days to produce a profile for a single customer. This is then repeated many times to produce a library of customer profiles for a given customer type. The library of profiles will be used throughout REM to determine the best power system design.

At the end, total demand is increased by a factor based on the annual electricity demand growth and the number of years for demand growth between the initial year and the final design year.

While this procedure for producing demand profiles is somewhat convoluted, it does have some useful features:

- Demand can vary with weather parameters such as temperature, and the weather parameters are synchronized with the solar resource in the PVWatts data. Thus seasonal variations can be accounted for.
- Giving a fixed energy demand per-activity and then having some variation regarding which hours that demand is activated mimics the reality of customers turning on and off a set of appliances. Having a realistic customer-level demand representation with variability allows REM to take into account the advantage of grouping customers together to create a more advantageous aggregate demand profile. (For example, a group of customers will tend to have a smaller ratio of peak to average demand compared to a single customer, and thus will require a less expensive system for the same quality of service.)

Vaishali:

For examples of demand profiles produced in the Vaishali study, see the chapter on REM outputs. The case of “typical” residential demand that was considered in this study includes lighting, television, and fan usage. The specific parameters used to describe the demand came from a combination of census data, studies, surveys, and demand data captured from Vaishali. For a fuller description of the process of determining the demand profiles, see the thesis of Yael Borofsky.

3 DESIGN OF LOCAL GENERATION FOR OFF-GRID SYSTEMS

3.1 PRIOR WORK IN DESIGN OF LOCAL GENERATION

3.1.1 Literature

A sub-problem within REM is the design of generation and storage for off-grid systems. A variety of methods have been demonstrated in the literature for the optimal selection of generation and storage in this kind of power system. A recent review of methods is provided in (Luna-Rubio, Trejo-Perea, Vargas-Vázquez, & Ríos-Moreno, 2012).

The general methodology for selecting a design for a local generation system includes searching the space of candidate designs, and evaluating their performance via one or more metrics based on some system simulation.

There are several categories of approaches for searching the space of candidate generation and storage options. These approaches include pre-selecting options, "gradient methods" such as linear programming (Tafreshi, Zamani, Ezzati, Vahedi, et al. 2011; Zamani et al. 2012; Erol-Kantarci et al. 2011), heuristic searches such as genetic algorithms (Khatib et al. 2012; Bala & Siddique 2009; Yang & Buxiang 2012; Tafreshi, Zamani, Ezzati & Vahedi 2011), and exhaustive searches. Among these methods, there is generally a tradeoff between accuracy in representation of the system and speed of the algorithm. Linear programming is among the fastest algorithms, but is not ideally-suited to handle non-linear effects, such as battery operation and degradation, or to handle non-continuous decision spaces, such as lumpy investment decisions. The REM utilizes a heuristic search through the design space with a simulation of each design option. This allows all relevant system dynamics to be included in the simulation, and for speed and accuracy to be flexibly traded by tuning the heuristic algorithm.

Within each of these search methods, there are various ways to represent system operation. System operation could be optimized (Hafez & Bhattacharya 2012; Bustos et al. 2012; Whitefoot et al. 2012) or based on heuristics (Tomoiaga et al. 2013), and can be represented with or without time-series constraints. While representations without time-series constraints (such as load block demand representations) enable faster-running solutions, they are not well suited for systems with large day-to-day variability, such as systems with high levels of intermittent renewables. Thus, a time-series representation of system operation has been chosen for the REM.

The literature includes microgrid design assessments based on many different metrics or objectives. Most of the metrics can be bucketed into the categories of cost (Erol-Kantarci, Kantarci, & Mouftah, 2011; Liu, Liu, Du, Ruiqi, & Huang, 2013; Tanrioven, 2005), environmental impact (Saito, Niimura, Koyanagi, & Yokoyama, 2009; Setiawan, Zhao, & Nayar, 2009), and reliability (Arefifar, Mohamed, & El-Fouly, 2013; Bahramirad, Reder, & Khodaei, 2012). Some evaluations also may use more technical metrics as objectives, such as ability to maintain stable frequency during grid events. Analyses may consider only a single objective factor, or multiple factors. While all of these factors are relevant to design of electrification systems, the REM focuses on costs because this is seen as the most pressing issue for deployment of these systems in the developing world rural electrification context. In REM costs can be considered from a business perspective (profit maximization) or from a social perspective (maximization of social welfare through minimization of system costs plus cost of non-served energy). For the Vaishali study, we have considered costs from a social perspective.

3.1.2 Existing solutions

Some existing software tools for the design of local generation exist. This are discussed below, and followed by a justification of why it was necessary to develop a new generation design module for REM.

HOMER

Hybrid Optimization of Multiple Energy Resources (HOMER) is a tool for selecting and sizing generation and storage devices for a microgrid system. HOMER allows the user to specify generation and storage options and then it simulates a year of system operation under each option in order to find the least-cost option. It allows the user to make limited choices regarding operation of the system. An overview of HOMER is provided in (T. Lambert, Gilman, & Lilienthal, 2006). HOMER is used or referenced in many papers, including (Agalgaonkar, Dobariya, Kanabar, Khaparde, & Kulkarni, 2006; Fang, Cai, Lin, & James, 2012; Hafez & Bhattacharya, 2012; Mohamed & Khatib, 2013; Sen & Bhattacharyya, 2014; Setiawan et al., 2009; Su, Yuan, & Chow, 2010; Tafreshi, Zamani, Ezzati, & Vahedi, 2011; Whitefoot, Mechtenberg, Peters, & Papalambros, 2012; Zamani, Ezzati, Farashah, Dahri, & Tafreshi, 2012; Zhu & Yang, 2012).

HOMER is a mature and well-maintained commercial software, and a legacy free version is also available. Advantages are ease of use and the ability to consider a large variety of generators. HOMER is limited in only representing one value of unserved energy, having limited system operation options, and not being easily integrated into distribution network design.

DER-CAM

Distributed Energy Resources Customer Adoption Model (DER-CAM) is a tool which selects generation technologies for a grid-connected home or microgrid with the objective of minimizing costs of utility bills. It also provides guidance for how the system should be operated to minimize utility costs. DER-CAM is described in (Bailey, Creighton, Firestone, Marnay, & Stadler, 2003).

DER-CAM has the advantage of better considering operations in the design of generation. However, it is limited for this application in that it is intended for grid-tied single-building systems, and thus does not consider siting of generation or network planning.

HOGA

Hybrid Optimization by Genetic Algorithms (HOGA) is a tool which performs integrated optimization of generation selection and operation for hybrid power systems via genetic algorithms.

HOGA allows for a very robust exploration of the generation design space in reasonable time, but it is limited by its lack of integration with network design.

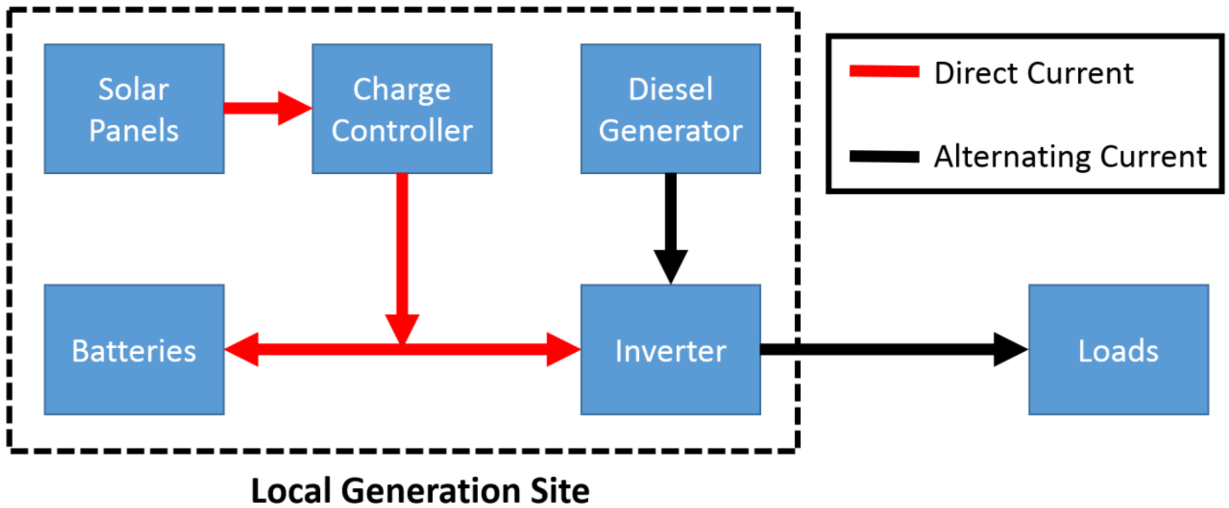
Relationship of REM to these solutions

There are two main differences between the REM approach and these existing solutions. First, is that REM is designed to automatically produce local generation designs for a large number of systems spread over a large area, and includes tradeoffs between detail and speed that are catered to this purpose. The second is that the REM local generation design approach is integrated into the broader algorithm for planning rural electrification systems (including clustering of customers and distribution network design).

3.2 Architecture of Local Generation Systems in REM

REM assumes that each off-grid system is supplied by its own single centralized local generation site, which supplies alternating current (AC) electricity to the local distribution network (for microgrids) or directly to the customer (for isolated systems). Sources at the local generation site can be solar photovoltaic (PV), a diesel generator, or both. Battery storage can also be included.

The figure below shows the general architecture assumed for local generation sites in REM.



The center of this architecture is a two-way inverter, which has three electrical ports: (1) a DC input/output for receiving power from solar and batteries and for charging batteries, (2) an AC input for receiving power from the diesel generator, and (3) an AC output for providing power to the loads. A charge controller is required between the solar panels and the inverter's DC port to convert the voltage to the battery bank voltage level. The charge controller may also have maximum power point tracking (MPPT) capabilities, which allows for maximum extraction of power from the solar panels (REM assumes it does).

The REM representation of this architecture is flexible in that not all of the components are required. Only one energy source is needed (solar panels or a diesel generator), batteries are optional, and REM will automatically include or not the charge controller and inverter as required.

There are of course many alternate architectures that could be considered for a microgrid or home system. This architecture was selected because it can be supported with commercially available off-the-shelf components and because it provides AC service to the customers, which allows a more straight-forward comparison with grid extension.

In practice, many home systems and small microgrids provide DC service, in part because it requires less and less expensive components. DC service can also be paired with highly efficient DC appliances. However this DC service is—at least today—seemingly less able to expand to meet customer aspirations in that there is only a limited set of DC appliances available. Also, components to support higher-voltage

DC distribution, which would be required in larger DC microgrids, are less readily available. It would be interesting in future work to do a more detailed study of in which cases a DC or an AC architecture provide lower system costs. But it is important to remember that the tradeoffs between DC and AC service depend on much more than cost, including issues such as:

- Customer preferences and perceptions;
- The ability to connect to the grid in the future;
- Availability of electrical components and appliances (and expertise in servicing them);
- Business models.

Other architectures that could be of interest in future work include:

- Generation assets distributed throughout a microgrid, to reduce network costs or support customer autonomy or new business models.
- More than one diesel generator, to allow more efficient diesel generator operation.
- Inclusion of other generation sources, such as micro-hydro.
- Grid-interactive architectures.

3.3 Local generation lookup table

REM produces a "lookup table" of pre-solved local generation designs for off-grid systems of various sizes, which is used in subsequent steps including clustering and selection of electrification mode. Since REM is designed to plan electrification for large areas, it would be very time consuming to separately plan a local generation site for every proposed microgrid. Instead, one lookup table is produced per sub-district based on a limited set of local generation designs. Since the PVWatts solar and weather data is associated with sub-districts, the lookup tables account for the associated variability in solar resource and demand between sub-districts. However, some variability which exists in reality is missed with this approach, including: (1) variability in distribution losses related to the specific layout of the microgrid customers, and (2) variability in location-specific costs, including diesel fuel. Improvements could be made in future work to consider these factors in a way which does not dramatically increase model run time.

In the current implementation, REM considers only one customer type per run, and so the lookup table can be simply organized based on system size, where size is defined by the number of customers in a system. For a sample of microgrid sizes, designs are saved including the following information:

- Size of solar array, battery bank, and diesel generator (not all are required)
- Annual financial cost
- Annual non-served energy cost
- Fraction of demand served

This information is used as a basis for quickly estimating local generation costs and other parameters for off-grid systems of all sizes in subsequent steps of REM.

The lookup table is created by producing local generation designs and calculating their costs for a sample of microgrid sizes, where several designs are produced for each size (the process of designing a microgrid for a given size is described in the following sections). The user specifies the number of microgrid sizes to be included in the lookup table, and REM automatically selects the sizes so that they are log distributed between a size of 1 customer and a size with the maximum number of customers in any analysis region (since that would be the largest possible microgrid for the run). Some small sized microgrids are automatically included also, since granular information about small microgrids is important to the clustering process described in the following chapter.

For a given trial of a microgrid with size n customers, the aggregate demand to be served is calculated by randomly selecting n profiles (with replacement) from the profile library and adding up their demand. Since there is some random variation between demand profiles, producing multiple designs and averaging the design results helps to make sure that the design produced was not an outlier. This is more important for the smaller microgrids, because at larger sizes the aggregate profile shape will approach the average demand profile. The user specifies how many trial designs to produce for each microgrid size in the REM execution script.

3.4 Pattern search of generation design space

For a given trial of a given microgrid size, REM must produce a microgrid design and calculate the associated costs of that design.

In principle, based on the REM local generation architecture and the local generation catalog, there is a significant number of design decisions, including:

- Type of solar panel (large or small)
- Size of solar array
- Type of battery (large or small)
- Size of battery bank
- Size of diesel generator
- Size of inverter
- Size of charge controller

REM simplifies this design problem into a 3-dimensional search over:

- Size of solar array
- Size of battery bank
- Size of diesel generator

The design variables of type of solar panel and type of battery are fixed before the search by estimating a likely final sizes of the battery bank and solar array (as proportions of the total energy demand), and determining which solar panel type and battery type are likely to be cheaper at that size. This is a crude approximation and likely could be improved in future work.

The design variables of inverter size and charge controller size are solved after simulating the microgrid operation. The sizes of these components are initially assumed to be infinite for the operation, and then they are set to be whatever size would be necessary to support the operation. In this way we lose the option of saving some cost by restricting the sizes of these components, but significantly reduce the size of the design space.

For the remaining design variables (sizes of solar array, battery bank, and diesel generator), REM seeks to find the optimal design choice, which is the design that minimizes the sum of annualized up-front costs, annual recurring costs, and annual non-served energy cost.

It would be difficult to calculate these costs accurately without actually simulating the operation of the microgrid. Many of the costs depend on complicated time-series considerations, such as the hourly solar resource and state-dependent operational rules. Thus this problem is not well-suited to methods that depend on calculating the gradient of the cost in the design space.

REM searches the design space using a method known as "pattern search" that does not require calculating a gradient (Hooke & Jeeves, 1961).[19] This pattern search includes the following steps:

- Define your space of candidate designs based on discrete options for the three design variables
- Select an initial set of candidate designs to evaluate. This set consists of a "center point" design, plus designs that are n design options away from the center in each dimension.
- Calculate the cost of each of these design options. (There is more information on how REM does this in the following sections.)
- The lowest cost design becomes the new center point. If it is different from the previous center point, select a new set of designs that are n design options away from the center in each dimension, calculate their costs, and set the center point to the lowest cost point again.
- Repeat this until the center point is the same for subsequent steps. When this happens, replace n by n/2 (rounded up to a whole number), and continue the process.
- When n has a value of 1 and all the immediately adjacent design options are higher cost than the center point, the center point design is a local minimum. This is selected as the design, and the costs and specifications of this design are returned.

This pattern search is guaranteed to find a local minimum in the defined design space, but may not necessarily find the lowest-cost of all possible designs. In practice though, this method has performed well for the cases I have studied.

3.5 Simulation of Operations

Many of the costs of interest of a given local generation design option depend on how the system operates. In REM, operation of the system dictates how much fuel is burned and how much demand is not served, how long batteries will last, and what sizes of power electronic components are required. Thus, some representation of system operation is necessary to produce a reasonable estimate of system cost and performance.

In REM, we consider system operation by simulating the hour-to-hour operation of the system over a sample of the hours of the year. The hours of the year to be considered can be specified by the user by designating a number of blocks of hours and the duration of each block. The blocks are automatically evenly spaced throughout the year, which allows the simulation to take into account seasonal effects. The parameters of selecting a sample of the year are included in the REM execution script.

[19] http://en.wikipedia.org/wiki/Pattern_search_(optimization) provides a helpful illustration of pattern search in two dimensions.

In this simulation the following are major inputs:

- **Demand (kWh):** Hourly demand is specified as the sum of the demand of each connected customer in the final year, assuming that the annual demand growth rate applies uniformly to every hour. In each hour the demand includes a critical and a non-critical component. The demand is considered for just the specified set of hours.
- **Solar resource (kWh/kW):** Hourly quantities for the DC energy production of a solar array per kW of solar array effective capacity. The solar resource is considered for just the specified set of hours.
- **PV nominal capacity (kW):** PV nominal capacity is part of what defines the candidate design. It is used to calculate PV investment and O&M costs.
- **PV effective capacity (kW):** The PV effective capacity accounts for degradation in PV capacity between initial installation and the final year. Since the installation of systems may be spread in time between the first and last year, REM considers a representative PV capacity degradation by applying the annual PV degradation for half the years between the initial and final years. It is assumed that a PV array with a certain effective capacity operates the same as a new array with that nominal capacity.
- **Battery nominal capacity (kWh):** Battery nominal capacity is part of what defines the candidate design. It is used to calculate battery investment and O&M costs.
- **Battery effective capacity (kWh):** The battery effective capacity accounts for degradation in battery capacity between initial installation and the final year. Since batteries are typically the component of the generation system with the shortest lifetime, the representative battery capacity is the midpoint between initial capacity and end-of-life capacity. It is assumed that a battery bank with a certain effective capacity operates the same as a new battery bank with that nominal capacity.
- **Diesel generator capacity (kW):** Diesel generator capacity is part of what defines the candidate design. It is used to calculate diesel generator investment and O&M costs. It is assumed that diesel generator capacity does not degrade over the lifetime of the diesel generator.
- **Relevant technical parameters and costs from the local generation catalog and settings**

The simulation represents hourly operation of the generation system for a selection of the final year. In each hour the demand plus any battery charging plus losses must be met by outputs of the PV panels, the diesel generator, and the batteries, plus any non-served demand (critical or non-critical). The simulation considers only the information available in the current hour, and does not include any forecasting.[20]

The simulation decides how to use the available resources in two steps:

1. The least cost set of resources are utilized to meet the demand
2. If justified, additional resources are utilized to charge the battery

---

[20] In future work, it would be interesting to explore what cost savings are achievable be incorporating forecast-based operation into the control logic. In practice, I have not yet seen developing world microgrid operators including forecast-based operations. For small microgrids, the additional cost and complexity of the control system would be challenging, and there could also be data connectivity issues. Still, it could make sense in some cases.

The basic principle behind the dispatch is to meet the demand (and possibly charge the battery) at least cost in each hour. The cost of using most of the resources in any hour is notionally known, but the decision to charge or discharge the battery includes an opportunity cost which ties the use of the battery to implications for other future hours.

For the battery, the opportunity cost of the battery's energy cannot actually be known to the model without simulating future hours (and dramatically increasing the time required for the simulation). To estimate this cost I created a simple heuristic. The basic idea of the heuristic is that the battery energy can substitute for another resource at a future time, so the value of the energy in the battery should be based on the avoided cost of not having to use that other resource. If the battery is being used optimally, it will substitute for the most expensive resources possible, so that total costs are minimized. If there is only a little bit of energy in the battery, this would likely be used to substitute for some of the most expensive resource over the coming hours. However, if the battery has more energy stored, it may be possible to completely eliminate use of the most expensive resource over some time, plus substitute for some of the second most expensive resource. In this way, when the battery is closer to fully charged, the average value of energy stored in the battery is reduced. This concept is implemented in REM by basing the battery energy value in each simulation hour on one of the other resources as a function of state of charge of the battery, such that the battery energy value grows as the battery is depleted (and vice versa).

To give a specific example, if there are three available non-battery resources to meet the demand, one of those three would be assigned as the replacement resource each hour. If the battery state of charge was in the top third of the effective range, the lowest-cost resource would be the replacement; middle third of state of charged would be matched with the middle cost resource, and bottom third of state of charge would be matched with the highest cost resource.

Once the battery is matched to a replacement resource, the value cost of energy in the battery is assigned the cost for that resource to serve the demand, with a correction for the losses between the energy stored in the battery and serving the demand (losses in the battery, the inverter, and the distribution wires), and with another correction for the cost of using up some of the battery's lifetime capacity.

While this heuristic for setting a value for stored energy in the battery bank certainly does not lead to the optimal possible operation of the system, it does lead to reasonable operation in most cases. For some examples of the resulting operation patterns see the chapter on REM outputs.

Step 1: Deciding how to meet the demand

For the first step of the simulation in each hour (deciding how to meet the demand), the following resources are available with the following costs:

Resource	Cost ($/kWh)
Solar	0
Generator	Cost of diesel fuel per kWh when the generator is at half load, with a correction for losses in the distribution grid.
Curtail non-critical demand	Cost of non-served energy for non-critical demand.
Curtail critical demand	Cost of non-served energy for critical demand.

Battery	Cost of the replacement resource, with a correction for losses in the battery, inverter, and distribution grid, plus a correction for using up battery lifetime.

Each of these resources also has a limited amount of energy it could provide in the hour. These limits are listed below:

Resource	Hourly energy limit
Solar	Array effective capacity (kW) multiplied by solar resource for the hour (kWh/kW).
Generator	The maximum power output of the generator multiplied by one hour.
Curtail non-critical demand	Quantity of non-critical demand in the hour.
Curtail critical demand	Quantity of critical demand in the hour.
Battery discharging	The remaining energy in the battery (above the minimum state of charge), or the limit determined by the kinetic battery model.[21]

These resources are used in order from least cost to highest cost until all there is a balanced generation equation (i.e. when supply equals demand plus non-served demand plus losses). Losses are induced in the battery, the charge controller, the inverter, and the distribution grid, where losses are calculated as the product of the energy entering each item and the efficiency of the item. A more realistic representation of the losses could be considered in future work.

Step 2: Deciding whether and how to charge the battery

For the second step of the simulation in each hour (deciding whether and how to charge the battery), the following resources are available with the following costs:

Resource	Cost ($/kWh)
Solar	The cost of using up battery lifetime.
Generator	Cost of diesel fuel per kWh when the generator is at half load, with an adjustment for losses in the rectifier and the battery, plus the cost of using up battery lifetime.
Curtail non-critical demand	Cost of non-served energy for non-critical demand, with an adjustment for losses in the distribution network, the rectifier, and the battery, plus the cost of using up battery lifetime.[22]

[21] For a description of the kinetic battery model, see (Manwell & McGowan, 1993). This model represents the battery as a two-tank system, in which energy in the battery is split between a chemically-bound tank and an immediately available tank. Time constants control the rate at which energy can transfer between the tanks. The model can be used to determine the maximum amount of energy the battery can absorb or release in a given time period.

[22] In the current implementation, the costs of using non-served demand as a “source” to charge the battery assumes that power flows from the demand to the battery and imposes losses. This should be corrected in future work, because curtailing demand will often actually reduce losses.

Curtail critical demand[23]	Cost of non-served energy for critical demand, with an adjustment for losses in the distribution network, the rectifier, and the battery, plus the cost of using up battery lifetime.

Each of these resources also has a limited amount of energy it could provide to charge the battery in the hour. The limit is the limit from step 1, minus the amount of the resource that was allocated for use during step 1.

If any of these resources have a cost below the opportunity cost of battery energy for that hour, they will be used to charge the battery, starting with the least cost option. Each resource with cost less than the battery energy cost will be used until all such resources are exhausted, the battery is fully charged, or the charging limit for the hour (calculated by the kinetic battery model) is reached.

At this point, the dispatch of all available resources has been determined for the hour. These dispatch values are saved, and the state variables are updated for the next hour (including battery state of charge, kinetic battery model charging and discharging limits, and battery energy value).

Once all of the hours are considered, the dispatch of all of the resources for each hour is saved, and is used to calculate metrics in later steps.

3.6 COST CALCULATION

Using the dispatch that results from the local generation simulation, several metrics can be calculated, which are then employed in the calculation of cost for the local generation site. The simulation produces the following major outputs:

- **Required capacities for the inverter and charge controller:** To reduce the dimensionality of the design space, inverter and charge controller capacity are set to the minimum required capacity to support the simulation. These capacities are used to calculate inverter and charge controller costs.
- **Battery lifetime:** The fraction of the battery's lifetime throughput that is used up in the final year is used to estimate the battery lifetime. This lifetime is utilized to calculate the annuity of the battery system investment.
- **Diesel generator lifetime:** The number of hours that the diesel generator is used to calculate the diesel generator lifetime. This is employed to calculate the annuity of the diesel generator investment.
- **Diesel fuel consumed:** The quantity of diesel fuel used in generating electricity plus the quantity used in starting up the diesel generator is calculated.
- **Non-served demand quantity:** The quantities of non-served critical and non-critical demand are used to calculate the non-served demand cost for the final year.

[23] In all or nearly all cases curtailing critical demand should not actually be selected as a resource for charging the battery, because it would not be cost-effective. Still, the option is included in the model.

These metrics are utilized to calculate a set of annual costs for the local generation system, including financial costs and the cost to the customers for non-served demand. Here are all of the costs included for a local generation system:

- Annuity of investment costs for PV, batteries, diesel generator, inverter, and charge controller (some of these may not be included in a given design)
- Annual O&M costs for PV, batteries, diesel generator, inverter, and charge controller
- Diesel fuel cost for the final year
- Annual per-system cost
- Non-served energy cost for the final year

For the calculation of annuities of investment costs the following information is needed:

- Investment cost
- Investment lifetime
- Discount rate

The investment lifetime and discount rate are used to calculate an annuity factor as follows:

$$Annuity\ Factor = \frac{dr}{1 - (1 + dr)^{-n}}$$

$$dr = discount\ rate$$

$$n = investment\ lifetime.$$

This annuity factor is then multiplied by the investment cost to get the annuity of the investment.

For most of the investments, the investment lifetime is specified as input data (e.g. solar panels, inverters, charge controllers).

For the battery, the lifetime is based on the lifetime energy throughput of the battery and the final year dispatch of the battery. The lifetime energy throughput is divided by the battery usage (charging plus discharging) to get the battery lifetime.

For the diesel generator, the lifetime is based on the lifetime operating hours and the final year dispatch of the generator. The lifetime is calculated as the lifetime operating hours divided by the hours of operation in the final year.

3.7 Local generation results for Vaishali

To see samples of local generation results for the Vaishali study, see the chapter on REM outputs.

4 CLUSTERING OF CUSTOMERS FOR OFF-GRID AND GRID EXTENSION DESIGNS

4.1 INTRODUCTION TO THE CLUSTERING PROBLEM

In this context, clustering refers to the grouping of customers into candidate off-grid systems and grid extension projects. The purpose of clustering is to define the sets of isolated system, microgrid, and grid extension designs that will be produced and compared in the part of REM that follows clustering, which is described in the following chapter.

Given an analysis region with a large number of customers, there is a huge number of ways to group those customers into potential off-grid systems or grid extensions. Ideally, we would like to know which of all these possible groupings leads to the least-cost design. In practice, however, it would take too much time to evaluate all of them. Thus, we need a strategy to reduce the number of possibilities so that there is a feasible number of decisions and designs to be produced.

In order to manage this decision space, REM organizes customers into a set of grid-extension clusters. Each grid-extension cluster represents a set of consumers which will be a part of a single grid extension candidate design and will together receive the same decision about whether to be grid-connected or receive off-grid systems.

Each grid extension cluster has an alternative off-grid clustering. The off-grid clustering defines the grouping of customers into microgrids and isolated systems which could be selected instead of the grid-extension design. Each off-grid cluster within the grid extension cluster represents a set of consumers which could be part of a single microgrid, or a single consumer which could be part of an isolated system.

This chapter discusses how these clusters are produced. The following chapter describes how the clustering result is used to design a set of systems, compare their costs, and decide on the final set of systems.

4.2 PRIOR WORK IN THE CLUSTERING OF CUSTOMERS

There is one existing tool which solves a problem very similar to the REM clustering problem, which is known as the Village Power Optimization model for Renewables, or ViPOR. ViPOR is a tool for the design of distribution networks for isolated power systems. ViPOR seeks to find the least-cost wiring design for an area and takes into account additional costs associated with geography. The algorithm used in ViPOR is described in (T. W. Lambert & Hittle, 2000). ViPOR is used in, for example, in (Rout & Parida, 2013).

ViPOR is based on a simulated annealing algorithm for design of minimum-cost distribution networks for isolated power systems. The general approach is to propose an initial set of connections among consumption and demand nodes and then iteratively remove or change wires until a near-optimal design is obtained. This method has the advantage of explicitly considering tradeoffs between connections of customers to a microgrid versus providing stand-alone home systems. However, it assumes generation sites and costs are known and takes a long time to run for cases with many demand nodes. ViPOR also does not include any consideration of technical constraints in its logic, such as voltage requirements.

ViPOR is not currently maintained, but a legacy version is available through HOMER Energy, the company that maintains HOMER. ViPOR's advantage is a relatively user-friendly interface, but it is not well-suited for the design of large networks. It also only allows very limited consideration of generation design in the network design process.

Additional insights on clustering approaches can be drawn from research in the design of traditional distribution networks. While these approaches do not consider off-grid systems as an option, they are especially relevant to the grid-extension option. (Gonzalez-Sotres, Mateo Domingo, Sanchez-Miralles, & Alvar Miro, 2013; Mateo Domingo, Gomez San Roman, Sanchez-Miralles, Peco Gonzalez, & Candela Martinez, 2011; Miguez, Cidras, Diaz-Dorado, & Garcia-Dornelas, 2002; Moreira, Miguez, Vilacha, & Otero, 2012) describe heuristic methods for the design of large distribution networks. Some of these methods are imbedded in RNM, so REM implicitly uses them during the network design step. In future work, there could be advantages to integrating REM's clustering process with the grid extension design clustering process.

4.3 Creation of off-grid clusters

As stated above, the clustering process divides the analysis region into a set of grid extension clusters, where each grid extension cluster has an alternate off-grid clustering. In order to achieve this result, the off-grid clusters are built first, and then they are joined together into the grid extension clusters in a second step. This section describes the creation of off-grid clusters.

The following are the steps to create off-grid clusters:

1. Build a minimum spanning tree (MST) that connects all of the customers.[24] The connections (or lines) in the minimum spanning tree define the off-grid clustering decisions that are considered. A connection can be "activated" to join the customers on either side of the connection into the same cluster, or not activated so that those customers are in separate clusters. Initially assume that all connections are not activated, so that each customer is in its own cluster of a single customer.[25]
2. Number the lines in the MST according to length, from shortest to longest. This defines the order in which the potential connections will be evaluated.
3. For each line, in the order defined above, evaluate whether to connect the customers on either side of the line and join them into a single cluster. This is done in two parts.
    a. Say that the customer on one side of this candidate connection (along with any other customers already connected to it) is in cluster 1. And the customers on the other side of the candidate connection are in cluster 2. Look at the local generation lookup table for the number of customers in cluster 1 and cluster 2. If these clusters would be given a generation site with no supply (i.e. the size of the generation components are zero),

---

[24] A minimum spanning tree is the set of connections between customers so that all customers are connected directly or through other customers and the sum of the lengths of the connections are minimized. To create the minimum spanning tree, we use the mst function in the MatlabBGL library, which was posted to the Matlab file exchange by David Gleich, Stanford University, 2006-2008.

[25] An alternative clustering logic could initially assume that all connections are activated, and then consider deactivating each connection. This alternative method was not selected because it was more difficult to implement and took more computer time to run.

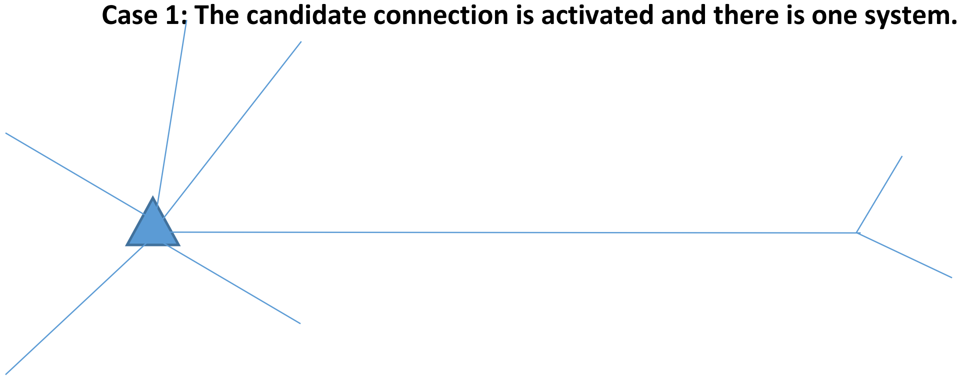
then join the clusters 1 and 2 into one cluster. This step helps to make sure that the clustering process does not get stuck in a local minimum where the microgrid size is too small and makes sure that REM is not designing microgrids with no generation site.

b. If the clusters were not already joined in step 3a, estimate the cost difference between building separate systems for clusters 1 and 2 and building one microgrid for all of those customers together. This cost estimation is described in more detail below. If it is economical, join the clusters 1 and 2 into the same cluster.

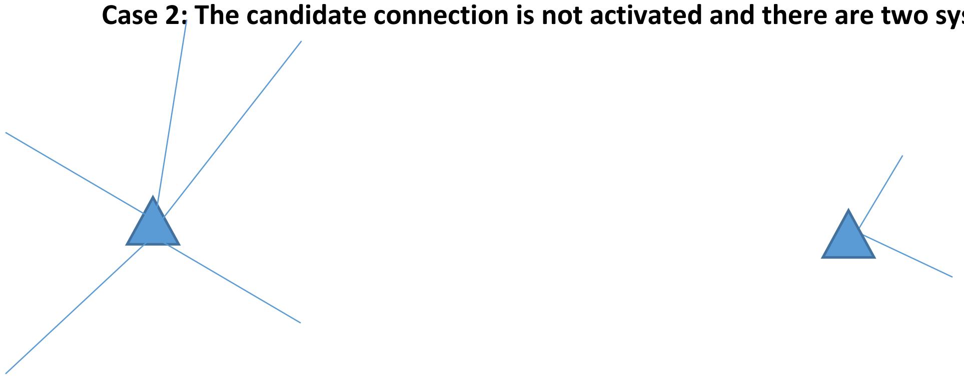
4. Once all lines have been considered, the sets of interconnected customers are the final off-grid clusters. Each of these clusters represents a candidate microgrid (if there is more than one customer in the cluster) or a candidate isolated system (if there is a single customer in the cluster).

Here is a more detailed description of step 3b, where we evaluate whether each connection is economical. The two situations which must be compared are shown in the diagrams below. Case 1 has the candidate connection activated so all of the customers are served by a single system. Case 2 has the connection under consideration removed, and two smaller local generation systems. The triangles represent the local generation sites and the lines represent connections between customers.

Case 1: The candidate connection is activated and there is one system.



Case 2: The candidate connection is not activated and there are two systems.



In the clustering process, REM only makes an estimate of the cost difference between these cases in order to decide whether to activate the connection of the MST that is under consideration. REM does not produce a full design for each case and do a detailed cost calculation because that would take too much time.

The cost estimate considers the differences in local generation costs and the differences in distribution network costs between the cases. The generation and network cost difference estimates are described in the following paragraphs.

The generation cost difference is estimated using the pre-solved local generation designs, which were described in the previous chapter. A local generation cost is extracted from the lookup table for the system in Case 1 and for each of the systems in Case 2. If the lookup table does not include an entry for the particular number of customers in one of the systems, the cost is estimated by a linear interpolation between the next larger and smaller entries in the lookup table.

The distribution network cost difference is estimated as the cost of a power line connecting the two clusters together. In the diagrams in the above figure, this line connecting the two clusters is the only difference between the network designs, but in reality we do not know how any of the network designs for the separate or combined clusters would look at this point. Thus this line cost is only a rough estimate of the network cost difference. The power line cost is estimated in a function that considers the length of the line (which is set as the length of the candidate MST connection) and the peak power flowing through the line (which is set as the annual peak demand of the smaller cluster). This simulates the situation of joining the clusters together and putting a generation site at the larger cluster.

The line cost function searches the RNM catalog for the least-cost low or medium voltage conductor that can carry the specified power over the specified distance, while not violating the thermal capacity of the line or causing a voltage drop of more than 10%.[26] The function returns the annualized investment cost of the line, plus annual maintenance costs. If there are no large enough lines, two lines can each carry half of the demand.

Finally these generation and network cost estimates for the two cases can be compared.

Case 1 includes the following costs:

- Generation cost for the combined cluster
- Cost of a line connecting the clusters

Case 2 includes the following costs:

- Generation costs for the two separate clusters

If the sum of the case 1 costs are lower than the case 2 costs, then the candidate MST connection is activated and the two clusters are joined into one cluster.

4.4 CREATION OF GRID-EXTENSION CLUSTERS

The off-grid clustering represents REM's best guess of how customers should be grouped into microgrids and isolated systems in order to minimize the cost of off-grid electricity access. The next step—building grid-extension clusters—groups off-grid clusters together into grid-extension clusters in order to define grid-extension design problems and comparisons between off-grid and grid-extension.

[26] The voltage drop tolerance can be changed in the line cost estimate function. The voltage drop is calculated for each conductor in the RNM catalog by using the MATPOWER Matlab toolbox (Zimmerman, Murillo-Sánchez, & Thomas, 2011).

Each final grid extension cluster defines a set of customers that are included in a single grid extension design problem that will be sent to the greenfield RNM. The process of actually producing this grid extension design, calculating its cost, and comparing it to the cost of off-grid electrification is described in the following chapter. It is important to remember that grid extension clusters just describe that a group of customers should be considered together when producing a candidate grid extension design—the grid extension clusters that this process produces are not yet associated with any particular connections to the existing grid and may not receive grid connections in the final design.

The process of building grid extension clusters is similar to the process of building off-grid clusters. The same basic paradigm of listing potential connections and then evaluating whether to activate each one is used. In the off-grid clustering, the basic question for each candidate connection was: “If the customers in these two clusters were to receive off-grid systems, would it be cheaper to build one system for them all together or two separate systems?” Alternatively, for grid-extension clustering, the question to consider for each candidate connection is: “If the customers in at least one of these two clusters will be connected to the grid, would it be better to connect them all to the grid together, or to give separate systems to the two clusters?”

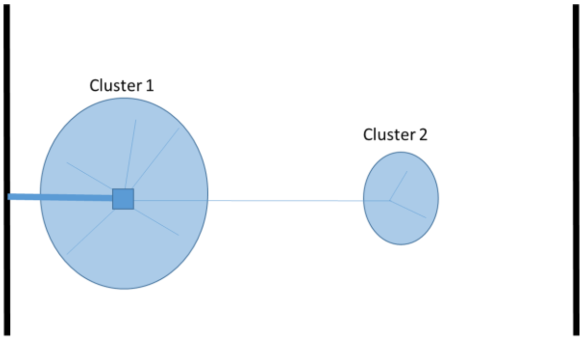
The following are the steps to create grid-extension clusters in REM:

1. The candidate connections between customers are defined by the same MST that was built in the off-grid clustering process. Assume initially that all of the connections that were activated in the off-grid clustering process are still activated, so that each off-grid cluster (microgrid or isolated system) is its own grid-extension cluster. The grid extension clustering starts with this off-grid clustering (as opposed to initially considering that every customer is separate) in order to ensure that a single off-grid cluster is not split over more than one grid extension cluster.
2. Number the remaining MST connections—those that are not activated—by length, from shortest to longest. This defines the order in which connections will be evaluated.
3. For each candidate connection, evaluate whether to activate it and joint the customers on either side of it into a single grid-extension cluster. This is done based on a cost estimate and comparison which is described in more detail below. This cost comparison is similar to the one done in the off-grid clustering process, but it includes estimates of the cost of connection to the grid.
4. Once all of the candidate connections have been considered, the sets of interconnected customers are the final grid-extension clusters. Each grid-extension cluster represents a set of customers which will receive a candidate grid-extension design in a single call to the greenfield RNM.

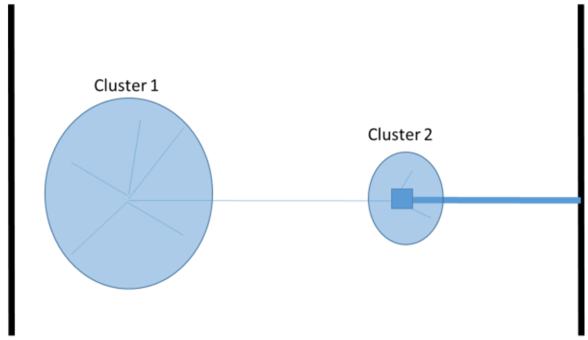
Here is a description of step 3, in which REM estimates the cost difference between joining two grid-extension clusters into a single cluster and keeping them separate. The basic idea is the same as the off-grid clustering comparison, where network and generation costs are estimated and compared between options. The main difference is that the scenarios considered assume that at least one of the clusters will be connected to the existing grid.

Following are diagrams of the scenarios considered and descriptions of estimated costs for the cases where the two clusters are joined into one. In these diagrams the heavy black lines represent existing medium voltage grid lines, the heavy blue lines represent extensions of the medium voltage grid, and the blue squares represent transformers. The thin blue lines represent connections between customers, and the shaded ovals represent the areas covered by each of the two clusters that could be joined.

Scenario 1: Two clusters connected together, and to the grid through cluster 1.



Scenario 2: Two clusters connected together, and to the grid through cluster 2.



Scenario 1 costs include:

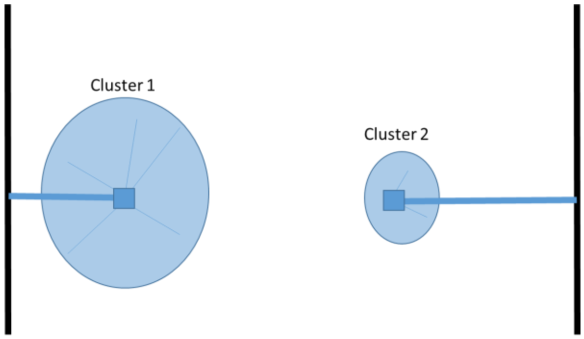
1. Procurement of grid energy, plus non-served energy cost for clusters 1 and 2.
2. Cost of a line from the center of cluster 1 to the nearest MV line, carrying power of the total peak demand of clusters 1 and 2 together.
3. Cost of a transformer with capacity of the total peak demand of clusters 1 and 2 together.
4. Cost of a line from the center of cluster 1 to the center of cluster 2, carrying the power of the peak demand of cluster 2.

Scenario 2 costs include:

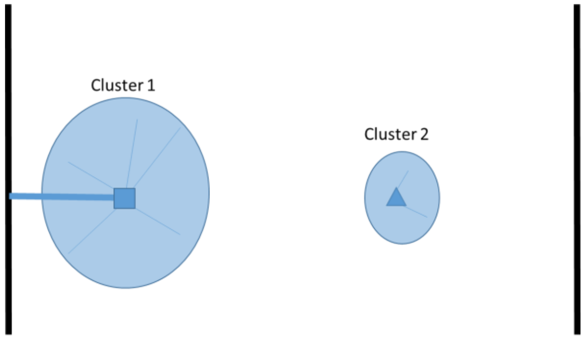
1. Procurement of grid energy, plus non-served energy cost for clusters 1 and 2.
2. Cost of a line from the center of cluster 2 to the nearest MV line, carrying power of the total peak demand of clusters 1 and 2 together.
3. Cost of a transformer with capacity of the total peak demand of clusters 1 and 2 together.
4. Cost of a line from the center of cluster 1 to the center of cluster 2, carrying the power of the peak demand of cluster 1.

Following are diagrams of the scenarios considered and descriptions of estimated costs for the cases where the two clusters kept separate. In these diagrams, as above, the heavy black lines represent existing medium voltage grid lines, the heavy blue lines represent extensions of the medium voltage grid, and the blue squares represent transformers. The thin blue lines represent connections between customers, and the shaded ovals represent the areas covered by each of the two clusters that could be joined. Additionally, the blue triangles represent local generation sites for off-grid systems.

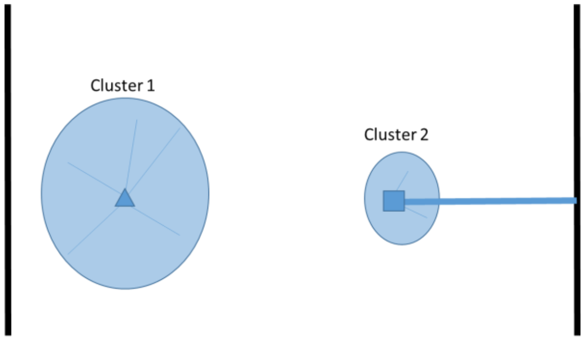
Scenario 3: Two clusters each connected to the grid separately.



Scenario 4: Cluster 1 connected to the grid, cluster 2 off-grid.



Scenario 5: Cluster 2 connected to the grid, cluster 1 off-grid.



Scenario 3 costs include:

1. Procurement of grid energy, plus non-served energy cost for cluster 1.
2. Procurement of grid energy, plus non-served energy cost for cluster 2.
3. Cost of a line from the center of cluster 1 to the nearest MV line, carrying power of the total peak demand of cluster 1.
4. Cost of a line from the center of cluster 2 to the nearest MV line, carrying power of the total peak demand of cluster 2.
5. Cost of a transformer with capacity of the total peak demand of cluster 1.
6. Cost of a transformer with capacity of the total peak demand of cluster 2.

Scenario 4 costs include:

1. Procurement of grid energy, plus non-served energy cost for cluster 1.
2. Cost of a line from the center of cluster 1 to the nearest MV line, carrying power of the total peak demand of cluster 1.
3. Cost of a transformer with capacity of the total peak demand of cluster 1.
4. Cost of off-grid energy for cluster 2.

Scenario 5 costs include:

1. Procurement of grid energy, plus non-served energy cost for cluster 2.

2. Cost of a line from the center of cluster 2 to the nearest MV line, carrying power of the total peak demand of cluster 2.
3. Cost of a transformer with capacity of the total peak demand of cluster 2.
4. Cost of off-grid energy for cluster 1.

There are two versions of the REM clustering logic. In the first, the lowest cost of the first two scenarios (i.e. the lowest cost with the candidate connection activated and the clusters joined) is compared to the lowest cost of scenarios 3-5 (i.e. the lowest cost with the candidate connection not activated then the clusters kept separate). If the cost of scenario 1 or 2 is lowest, then the connection is activated and the clusters are joined.

In the second version of the REM clustering logic, scenarios 1 and 2 are compared to scenario 3 only, with scenarios 4 and 5 not being considered (i.e. the scenarios with one cluster being grid connected and one receiving an off-grid system are not considered). While this second clustering logic considers less options and will perform worse than the first in some cases, in other cases this clustering performs better in practice. This may be because this second clustering logic avoids comparisons between ideally-sized microgrids and grid extensions that are too small early in the clustering process, which could result getting stuck in a local minimum of the clustering process.

Both of these two clustering logics were applied in the Vaishali study in various cases.

The off-grid energy costs and the line costs described above are calculated in the same way as they are in the off-grid clustering step. The only new types of costs here are the grid energy and the transformer cost estimate.

The grid energy cost has two components: the cost of procuring energy supply from the grid and the cost to consumers for non-served demand. First, the quantities of demand that will be served and not served are calculated using the total demand of the cluster(s) and the grid reliability data. Once this is determined, the non-served demand is imposed a cost based on the costs of non-served energy for critical and non-critical demand. For the part of the demand that is served, a cost is imposed for the volume of that demand plus losses in the distribution grid at a rate of the cost of grid energy.[27]

The transformer cost is estimated as the cost of the smallest medium-to-low voltage transformer in the RNM catalog with high enough capacity. If the total demand is greater than the capacity of the highest-capacity transformer, the required capacity is covered by one or more of the largest transformers, plus one smaller one (if needed).

4.5 VAISHALI CLUSTERING RESULTS

To see examples of clustering results from the Vaishali study, see the chapter on REM outputs.

[27] The cost of grid energy is described in the data chapter.

5 SELECTING THE FINAL POWER SYSTEM DESIGNS

This chapter describes the process of producing designs for various isolated systems, microgrid, and grid extensions, and comparing their costs to determine the final electrification mode for each customer and the final set of systems. The sets of systems to design and the cost comparisons to make were determined in the clustering process that is described in the previous chapter. While the clustering process relies on rough estimates of network costs, in this step more accurate network costs based on detailed network designs are used.

5.1 OVERVIEW OF COMPARISON OF ELECTRIFICATION MODES

When the clustering process is complete, all customers in the analysis region have been assigned to one of the grid extension clusters. Each grid-extension cluster is comprised of one or more off-grid clusters, which could be candidate microgrids or isolated systems. This clustering defines the designs to be produced and costs to be compared in order to decide the final set of systems. The steps in the process of producing system designs, calculating system costs, and comparing electrification modes to decide on the final designs are described in the following:

1. For each off-grid cluster that is a microgrid, design a local generation system and distribution network. Calculate the total cost of the microgrid.
2. Design an isolated system for the analysis region, and calculate its total cost. This design and cost will apply to any customers in the analysis region that receive isolated systems.
3. Compare each microgrid cost to the cost of replacing the microgrid with a set of isolated systems. Temporarily assign the relevant customers to whichever option (microgrid or isolated system) has a lower cost.
4. At this point the lowest-cost off-grid design has been produced.
5. For each grid extension cluster, estimate the costs that do not depend on the network design (grid energy, non-served demand, per-customer costs, and transformers of sufficient capacity). If these costs are greater than the total cost of off-grid systems for the customers in the grid-extension cluster, assign the customers to off-grid systems and skip steps 6 and 7. This is done to save the time of designing grid extensions in cases where they are clearly not the least cost option.
6. Design the distribution network for the grid extension cluster. Calculate the total cost of the grid extension.
7. Compare each grid extension cost to the total cost of off-grid systems for the relevant set of consumers. Assign the consumers to the option (grid-extension or off-grid) that has lower cost.
8. At this point, the lowest-cost set of systems has been selected for the analysis region.

The details of how the designs are produced and costs calculated for the three electrification modes (isolated system, microgrid, and grid-extension) are provided in the following sections.

5.2 DESIGN AND COST FOR ISOLATED SYSTEMS

The isolated system is the simplest mode to design for, because it has no distribution network. Thus, the design is defined entirely by the local generation system. The local generation system design is extracted from the local generation lookup table. This gives all isolated customers in a given analysis region identical isolated system designs.

With the design specified, costs for the isolated system can be calculated. The general idea is to estimate the annual costs for the “final year” in line with REM’s static model approach. Here are the included costs for isolated systems:

- Annuity of per-customer investment costs
- Annual per-customer operations and maintenance costs
- Annuity of local generation investment costs
- Annual local generation operations, maintenance, and fuel costs (for the last year)
- Annual per-system local generation costs
- Cost of demand not served (for the last year)

The first two of these costs are “per-customer costs” and refer notionally to equipment and maintenance done at the customer-level. This could include items such as meters and in-home wiring and activities such as meter reading and payment collection. These costs are calculated based on information in the settings. The annuities of investment costs are calculated using the provided information about investment cost and lifetime for isolated system per-customer costs, plus the discount rate.

The remaining costs fall out of the local generation design, and are already saved in the local generation lookup table.

5.3 Design and cost for microgrids

Microgrids are similar to isolated systems, except for having multiple customers and requiring a distribution network.

Microgrids include the following costs:

- Annuity of per-customer investment costs
- Annual per-customer operations and maintenance costs
- Annuity of distribution network investment costs
- Annual distribution network operations and maintenance costs
- Annuity of local generation investment costs
- Annual local generation operations, maintenance, and fuel costs (for the last year)
- Annual per-system local generation costs
- Cost of demand not served (for the last year)

The per-customer costs and the design and costs associated with the local generation system are produced in basically the same way as for the isolated system. There are only two differences in this part. First, the per-customer costs are multiplied by the number of customers. Second, it is possible that the local generation lookup table does not include a design for the exact size of this microgrid. In that case, the design and costs are estimated by linear interpolation between the next largest and next smallest microgrid sizes that are in the lookup table. This design by linear interpolation has some weaknesses that should be addressed in future work. This includes specifying infeasible designs in between lookup table entries with different generation mixes (e.g. in the transition between a solar-battery system and a solar-diesel system).

The design elements and costs that are new for microgrids are those associated with the distribution network. In order to produce distribution network designs and calculate their costs, REM takes advantage of another model called the Greenfield Reference Network Model, which is described in (Peco González, 2001). This model originally had the purpose of designing traditional distribution networks in order to estimate costs of the distribution company. For REM, we call RNM in a particular way and then post-process the results in order to be able to use them to make microgrid distribution networks.

By default, RNM assumes that the transmission substations are the primary source of power to the distribution grid, and power predominately flows from the transmission substation to high voltage lines, high-to-medium voltage transformers, medium voltage lines, medium-to-low voltage transformers, low voltage lines, and finally to the low voltage customers. The trick that REM uses to get RNM to produce microgrid distribution networks is to specify the transmission substation location as the geographic center of the microgrid.[28] RNM then designs the network from the bottom up, starting with the low voltage lines. If RNM places only one medium-to-low voltage transformer, in post-processing, REM replaces that transformer with a local generation system, and removes any higher voltage elements that RNM designed. Alternatively, if RNM requires multiple medium-to-low voltage transformers, and only one high-to-medium voltage transformer, the high-to-medium voltage transformer is replaced with a local generation site, and the higher voltage equipment designed by RNM is removed. If multiple high-to-medium voltage transformers are required, the current REM implementation will not produce sensible results. Thus far, this has not happened in practice.

In order to call RNM, a set of RNM input files must be built and then RNM is called with an executable file. RNM reads the input files, processes them, and then produces a set of output files. The output files include tables with cost and performance information about the design, plus a set of files in the shapefile format which specify the geographic locations and paths of distribution grid elements.

The following table lists the most significant RNM input files and how they are produced for the design of microgrid networks.

File	Notes
Low voltage customers	Information about the customers to be connected to the grid in low voltage. Includes: • Customer location – X and Y coordinates projected in meters for all the customers in the microgrid cluster. Z coordinate is set to 0 in current implementation.[29] • Demand – calculated as the peak hourly demand for the year for each customer. • Cost of non-served energy (low, medium, high) – low and high are the costs of non-served energy for non-critical and critical demand. Medium is a weighted average of the two.

[28] In the future it might make sense to have the center point weighted by demand. Since all customers are of the same type in the current implementation, this is a non-issue.

[29] If elevation information is available, RNM can take this into account when producing network designs.

Transmission substations	Information about any existing transmission substations. RNM assumes that a transmission substation is always the power source that feeds the distribution network. It is not relevant to the REM grid extension design, but it must be included for RNM to run. Includes: • Location – specified as center of the microgrid cluster.
Catalog	Information about available technical components (including conductors and transformers) and some model parameters. Most of the catalog is included in the case input data. Some catalog parameters are added during the execution of REM.

The following table lists the catalog parameters which are added during the execution of REM. These parameters are selected to reflect the “static model” approach of REM. Thus the design is produced to meet technical requirements (e.g. voltage) in the final design year, and the costs considered are the annuity of the investment costs and annual costs occurring in the final design year.[30]

RNM Input **(Spanish name in RNM documentation)**	**Value set in REM**
Discount rate (tasa de descuento)	Matches REM discount rate
Demand growth rate (tasa de crecimiento demanda)	0
Network equipment lifetime (años de vida útil)	Matches REM network equipment lifetime
Years of demand growth (años de crecimiento)	Matches REM network equipment lifetime
Years of demand growth for calculating losses (años de crecimiento solo pérdidas)	Matches REM network equipment lifetime
Fraction of initial demand to consider for calculating losses after years of demand growth for calculating losses (k demanda hasta fin vida útil)	Irrelevant (a small positive number—e.g. 0.01—will avoid RNM errors)
Cost of losses (Coste de las pérdidas (moneda/kWh))	The financial cost of the generation site per unit demand served (calculated as total annual financial costs divided by total annual demand served).

The cost of losses is the way for RNM to consider whether to invest in higher capacity conductors in order to reduce losses, and thus reduce the cost of electricity supply. It is necessary to consider this cost within RNM because the interaction between network and generation costs would otherwise be

[30] Within RNM there is a present value calculation of costs, including annual losses that could vary from year to year. However, the RNM settings used by REM make every year look identical so that this present value calculation does not influence the design.

ignored. When total systems costs are considered for the selection of electrification mode, however, the cost of losses is not included because the actual costs of electricity supply are directly included instead.

Once the design is produced, REM extracts the costs of conductors and transformers (including annuity of investment costs and annual maintenance costs) at the voltage levels which are required based on the post-processing logic described above.

Ideas for future improvements in design of microgrid networks:

This single step process to design a microgrid distribution that could be only low voltage, or a mix of medium and low voltage, is adulterated by the cost and capacity limits of the medium-to-low voltage transformers (which are placeholders for local generation sites in the former case and are "real" transformers in the latter case). It would likely be better to modify the RNM logic so that it is designed for the purpose of designing microgrids and this impact of the transformers is not an issue.

5.4 DESIGN AND COST FOR GRID EXTENSION

The grid extension design in REM includes all of the required equipment between a connection with existing medium voltage lines and the customer. This includes any new medium voltage lines, medium-to-low voltage transformers and low voltage lines.

Here are the cost which are considered for the grid extension option:

- Annuity of per-customer investment costs
- Annual per-customer operations and maintenance costs
- Annuity of distribution network investment costs
- Annual distribution network operations and maintenance costs
- Cost of energy purchased from the transmission grid (for the last year)
- Cost of demand not served (for the last year)

The per-customer costs are handled in the same way as they are for isolated systems and grid extensions. The costs of energy purchased from the grid and demand not served are calculated in the same way as they are in the clustering process.

The new piece is how to call the Greenfield RNM in order to produce the grid extension design. The general approach is similar to how RNM is used for producing microgrid designs, with a few significant differences in how the model is called and how the results are processed.

The following table lists the most significant RNM input files and how they are produced for the design of grid extension networks. The main difference between this and the microgrid network design is that there is an additional file for "imaginary" high-to-medium voltage transformers which actually represent the potential locations to connect to the existing medium voltage network. The secondary side of each of these transformers is at medium voltage, and RNM will consider them as potential medium voltage connection points for new grid extensions. In post-processing, REM will ignore these imaginary transformers and upstream infrastructure, and instead assume that the new medium voltage line connects to an existing medium voltage line at the same location as the imaginary transformer. A set of potential connection points is provided (10 of them) in case it makes sense to split up the grid extension cluster into parts that are fed from different connection points. This is especially important if the grid extension cluster is very large and could not be fed by a single connection point in medium voltage.[31]

File	Notes
Low voltage customers	Information about the customers to be connected to the grid in low voltage. Includes: • Customer location – X and Y coordinates projected in meters for all the customers in the grid-extension cluster. Z coordinate is set to 0 in current implementation. • Demand – calculated as the peak hourly demand for the year for each customer. • Cost of non-served energy (low, medium, high) – low and high are the costs of non-served energy for non-critical and critical demand. Medium is a weighted average of the two.
High-to-medium voltage transformers	Information about any existing high-to-medium transformers. From the REM perspective, these transformers actually represent the potential locations to connect to the existing medium voltage lines. Includes: • Location – locations are specified for connections points at the 10 nearest medium voltage line segments.
Transmission substations	Information about any existing transmission substations. RNM assumes that a transmission substation is always the power source that feeds the distribution network. It is not relevant to the REM grid extension design, but it must be included for RNM to run. Includes: • Location – specified as the location of the furthest away of the 10 medium voltage connection points from above.
Catalog	Information about available technical components (including conductors and transformers) and some model parameters. Most of the catalog is included in the case input data. Some catalog parameters are added during the execution of REM.

The following table lists the catalog parameters which are added during the execution of REM. These parameters are selected to reflect the “static model” approach of REM. Thus the design is produced to meet technical requirements (e.g. voltage) in the final design year, and the costs considered are the annuity of the investment costs and annual costs occurring in the final design year.[32]

[31] If the grid extension cluster is large enough, RNM may decide that a new high-to-medium voltage transformer is needed, beyond the set of those transformers that represent medium voltage connection points. If this happens, the final result will not be sensible. This error should be addressed in future work.
[32] Within RNM there is a present value calculation of costs, including annual losses that could vary from year to year. However, the RNM settings used by REM make every year look identical so that this present value calculation does not influence the design.

RNM Input (Spanish name in RNM documentation)	Value set in REM
Discount rate (tasa de descuento)	Matches REM discount rate
Demand growth rate (tasa de crecimiento demanda)	0
Network equipment lifetime (años de vida útil)	Matches REM network equipment lifetime
Years of demand growth (años de crecimiento)	Matches REM network equipment lifetime
Years of demand growth for calculating losses (años de crecimiento solo pérdidas)	Matches REM network equipment lifetime
Fraction of initial demand to consider for calculating losses after years of demand growth for calculating losses (k demanda hasta fin vida útil)	Irrelevant (a small positive number—e.g. 0.01—will avoid RNM errors)
Cost of losses (Coste de las pérdidas (moneda/kWh))	The cost of energy procured from the grid.[33]

Once the design is produced, REM extracts the costs of conductors and transformers (including annuity of investment costs and annual maintenance costs) for the new medium voltage lines, new medium-to-low voltage transformers, and new low voltage lines.

Ideas for future improvements in design of grid extensions:

The grid extension options considered and costs calculated in REM do not reflect all of the real options for gird extension or all of the true costs. For example, REM does not consider the options of connecting customers through extension of existing low-voltage or high-voltage lines. Also, REM does not plan for or consider costs of upgrades that might be required to the upstream network as a result connecting these additional customers.

One option for addressing these issues is to use the Brownfield Reference Network Model. This model considers a greater range of connection options and also considers impacts and required investments in the upstream grid. Unfortunately, there are some challenges associated with using the Brownfield RNM in the process of deciding which should be the final electrification mode. First, the Brownfield RNM takes much longer to execute than the Greenfield RNM, so it dramatically increases the time to solve a study area. Second, when there are multiple grid extensions being solved independently, it becomes complicated to determine required upstream investments and allocate their costs. Third, the Brownfield RNM does not deal well with systems that are undersized for peak demand, but undersized distribution infrastructure is very common in developing countries. These issues should be explored in future work.

[33] The current implementation assumes that all connection points face the same cost of energy procured from the grid. The model would need to be upgraded to allow consideration of varied energy costs from the various tapoffs.

One potential approach would be to use an estimate of upstream costs that can be quickly evaluated during the process of selecting electrification mode, and then using the Brownfield RNM model once at the very end of the process to design all of the grid extensions together.

Another limitation is that the selection of which medium voltage connection points to use within RNM is based on distance only, and does not consider the cost of energy available at that connection point or the reliability of energy from that connection point. The capabilities of RNM would need to be upgraded to be able to consider these factors.

6 REM OUTPUTS

This chapter demonstrates the outputs that can be obtained from REM by examining some of the outputs of the Vaishali study. Outputs discussed include demand profiles, local generation design and simulation, clustering, network design, and results summaries. Then an analysis of the results is provided.

Some different options were examined in several cases which were run for Vaishali. Most of these options were described throughout the chapters that describe the data and model logic. Here is a summary of the cases considered. The relevant features of the case are listed for the baseline scenario, and then deviations are listed for the other cases.

Case Name	**Notes**
Baseline – Clustering Logic 1	• Number of customers to match current Vaishali data • 1% annual demand growth • Grid reliability based on current data • Grid extension clustering includes microgrid comparison
Baseline – Clustering Logic 1 – Repeat	• Same as above
Baseline – Clustering Logic 1 – Reliable Grid	• Grid reliability set to perfect reliability
Demand Growth Scenario – Clustering Logic 1	• 16% annual demand growth
More Buildings Scenario – Clustering Logic 1	• 13% more customers added
Baseline – Clustering Logic 2	• Grid extension clustering does not include microgrid comparison
Demand Growth Scenario – Clustering Logic 2	• 16% annual demand growth • Grid extension clustering does not include microgrid comparison
More Buildings Scenario – Clustering Logic 2	• 13% more customers added • Grid extension clustering does not include microgrid comparison

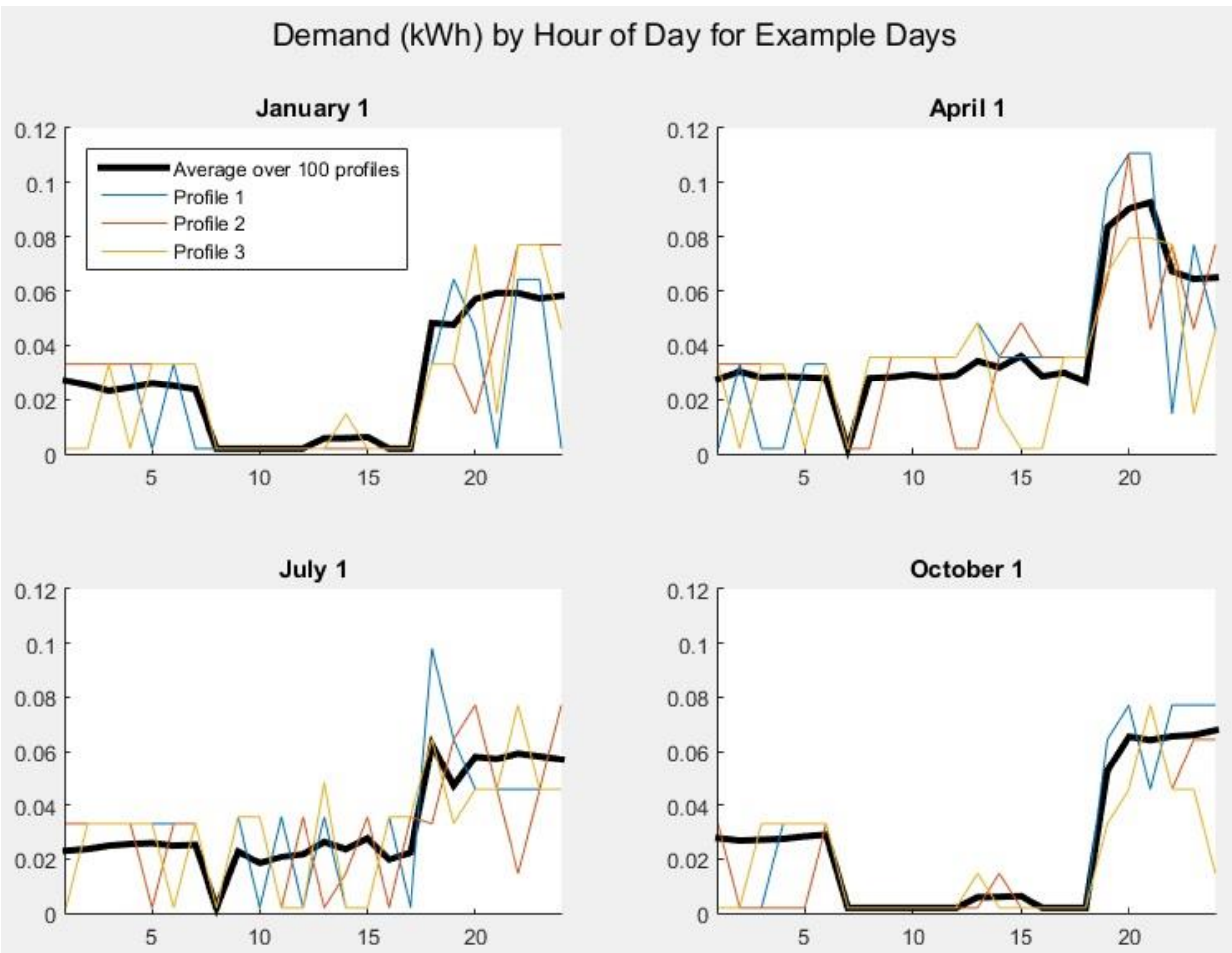
6.1 DEMAND PROFILES

The figure below gives a sense of how the customer demand profiles look for the Vaishali study. For each case, each analysis region got its own library 100 of demand profiles, including hourly values for critical and non-critical demand (as was described in the section creation of demand profiles in the data chapter). The figure below just gives some aggregate (critical plus non-critical) demand profiles for some sample days from analysis region 1 in the baseline scenario. The shape of the demand profiles will be generally similar across all analysis regions and cases, with some minor differences due to different weather data across analysis regions, and increased magnitudes of demand in the demand growth scenario.

In the figure, the black line represents the average customer profile over all 100 profiles in the profile library, and the colored lines represent single customer profiles for the first three profiles in the library. The average profile reveals daily and seasonal trends. In the cooler months, there is evening peak with lighting and television demand, nighttime lighting, and little demand during the day. In warmer months, the demand is significantly increased in the daytime and evening hours by fan usage.

The single customer profiles reveal some of the variation between profiles and the advantages of aggregating customers. Single customer profiles tend to have a greater peak demand than the average profile, which makes it more expensive to serve single customers.

Samples of demand profile library for baseline scenario region 1:



6.2 LOCAL GENERATION DESIGN AND SIMULATION

The tables below show some of the contents of the local generation lookup tables produced in some of the scenarios, for the first analysis region. The scenarios included are the baseline scenario, the repeat run of the baseline scenario, and the demand growth scenario. These were chosen because the set of number of customers in the lookup tables across these scenarios match and the locations of the first analysis region match, so it is easy to compare results.

Here are observations relevant to all three scenarios:

- In general, microgrid service improves (in terms of fraction of demand served) and levelized cost of energy (financial costs per unit energy provided) decreases as microgrid size increases. In the best cases, microgrids can serve nearly all demand at around $0.30/kWh.[34]
- There are four main "types" of designs produced: no system, solar only system, solar-battery system, and diesel-solar system. [35]
- For many single customer systems, no system was provided (i.e. the model chose a size of zero for all components).[36] This implies that the costs of the smallest equipment may be too expensive to support the AC architecture for very small systems (given the costs of non-served energy, demand profile, solar resource, etc.). In practice, most single customer systems offered in the market are DC systems.
- In some cases of small systems a solar only option, with no battery storage is produced. These designs are only able to cover about 17% of demand—essentially only daytime demand. In practice, small solar systems almost always are paired with batteries, so perhaps this reflects a lack of battery options in the catalog.
- Small to medium size microgrids receive solar-battery systems, which serve around 35% of demand.
- Larger microgrids receive diesel-solar systems and serve most demand. It appears that systems with diesel generators are not used in smaller microgrids because small enough generators are not available. In some cases the reliability dips, which is likely due to a lack of generator size options in the catalog. The demand growth scenario also had a dip in reliability for the largest microgrid size because the largest generator in the catalog could not meet the demand.

Here are observations from comparison of the baseline run and the repeat of the baseline run:

- Identical or nearly identical designs were produced across the baseline scenario and the repeat of the baseline scenario. This suggests that the generation design logic is somewhat robust to the random variations in demand profiles, and that the heuristic pattern search performs consistently, at least in this case.

Here are observations from comparison of the baseline runs and the demand growth scenario:

- The results across these scenarios are qualitatively similar, but with larger systems in the demand growth scenario (and diesel-solar systems being provided to microgrids with fewer customers). This should be expected because the demand in the demand growth scenario is just a proportional increase of the baseline scenario demand. If demand was to grow in a non-proportional way, the types of designs produced might look more significantly different. For example, if demand growth was concentrated in the daylight hours, it is likely that systems with more solar panels would be selected.

---

[34] This cost per unit energy only includes the cost associated with the local generation site. Costs of the distribution network and per-customer costs are not included.

[35] Other design types—such as solar-battery-diesel—are possible, but were not produced in these cases.

[36] In some analysis regions for the demand growth scenario, single customers were provided with solar systems.

Following the tables, are figures which show samples of how the dispatch of resources look in the REM simulations of solar-battery and diesel-solar microgrids.

For the solar-battery systems, the solar resource can cover daytime demand and charge the batteries during the day. The battery is used to cover most of the critical evening demand. Non-critical evening and nighttime demand is mostly not covered. During the sunny months, the battery may get fully charged during the middle of the day, and additional available solar resource is wasted. This inability to make full use of resources in off-grid systems is a large driver of higher energy costs for these systems, compared to reported costs of grid-tied solar energy.

For the diesel-solar systems, the solar panels cover daytime demand as available. The diesel generator covers nearly all of the critical and non-critical evening and nighttime demand. During peak hours of the year, the system may be capacity-constrained and not be able to serve all demand.

Baseline scenario local generation study results for region 1:

Number of Customers	Peak Demand (kW)	Average Demand (kW)	Solar Capacity (kW)	Battery Capacity (kWh)	Generator Capacity (kW)	Fraction of Demand Served	Financial Cost per Demand Served ($/kWh)
1	0.11	0.03	0	0	0	0	NA
2	0.22	0.06	0.25	0	0	0.17	0.75
3	0.33	0.09	0.25	0.46	0	0.21	0.60
4	0.44	0.12	0.5	2.76	0	0.39	0.71
10	1.10	0.30	1.17	5.52	0	0.37	0.62
31	3.31	0.94	2	0	2	0.94	0.52
96	10.05	2.93	6	0	7	0.97	0.43
301	31.18	9.17	21	0	20	0.95	0.39
942	97.13	28.69	63	0	100	1.00	0.39
2949	304.77	89.85	192.5	0	250	0.99	0.35
9231	952.81	281.23	516	0	750	0.98	0.33
28902	2982.67	880.48	1613.5	0	2250	0.98	0.33

Baseline scenario local generation study results for region 1 – repeat trial:

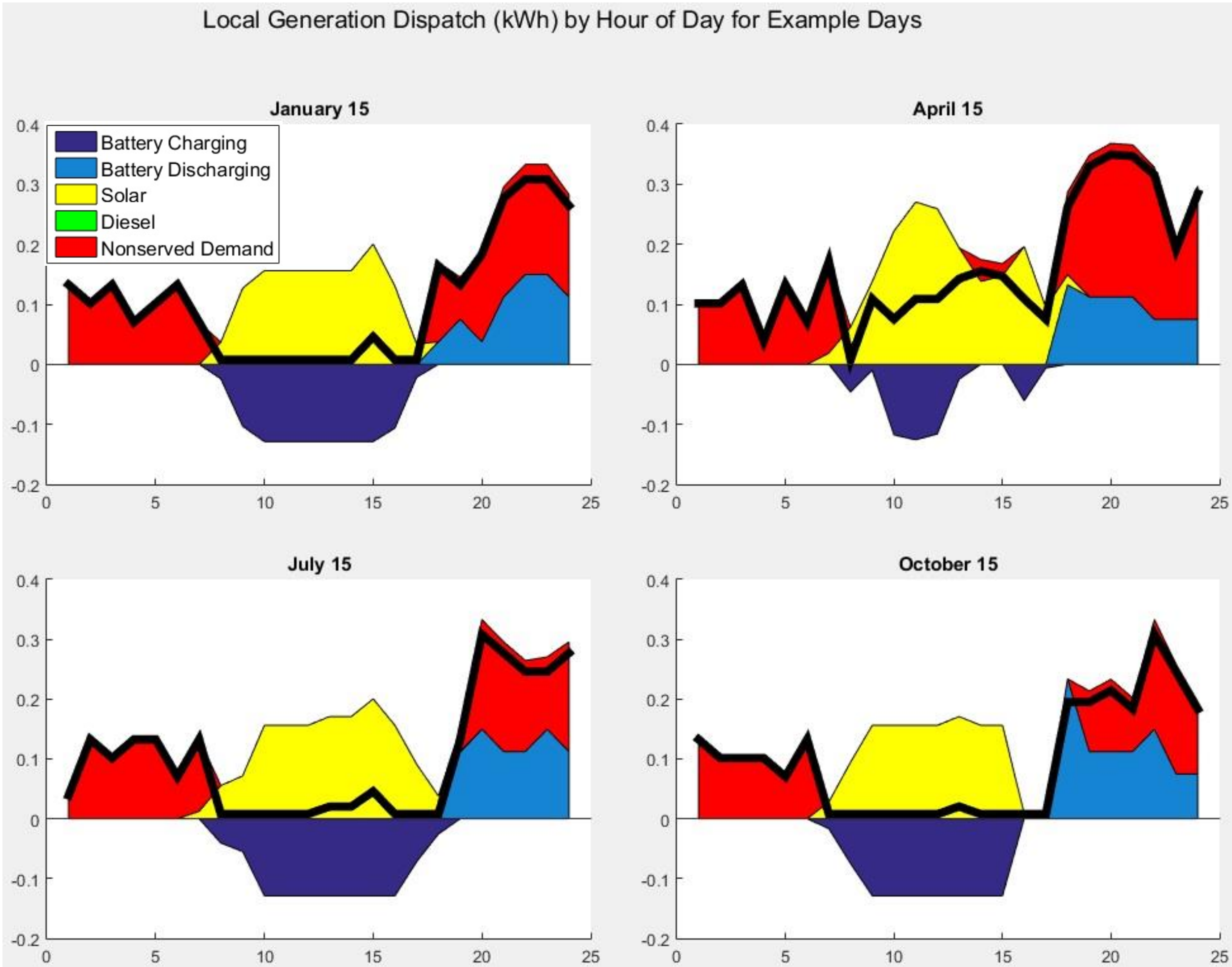
Number of Customers	Peak Demand (kW)	Average Demand (kW)	Solar Capacity (kW)	Battery Capacity (kWh)	Generator Capacity (kW)	Fraction of Demand Served	Financial Cost per Demand Served ($/kWh)
1	0.11	0.03	0	0	0	0	NA
2	0.22	0.06	0.25	0	0	0.17	0.75
3	0.33	0.09	0.25	0.46	0	0.20	0.61
4	0.44	0.12	0.5	2.76	0	0.39	0.71

10	1.10	0.30	1.08	5.06	0	0.35	0.61
31	3.28	0.95	2	0	2	0.94	0.51
96	9.99	2.92	6	0	7	0.97	0.43
301	31.11	9.17	21	0	20	0.95	0.39
942	97.08	28.70	63	0	100	1.00	0.39
2949	304.63	89.86	192.5	0	250	0.99	0.35
9231	950.55	281.27	516	0	750	0.98	0.33
28902	2979.95	880.71	1614	0	2250	0.98	0.33

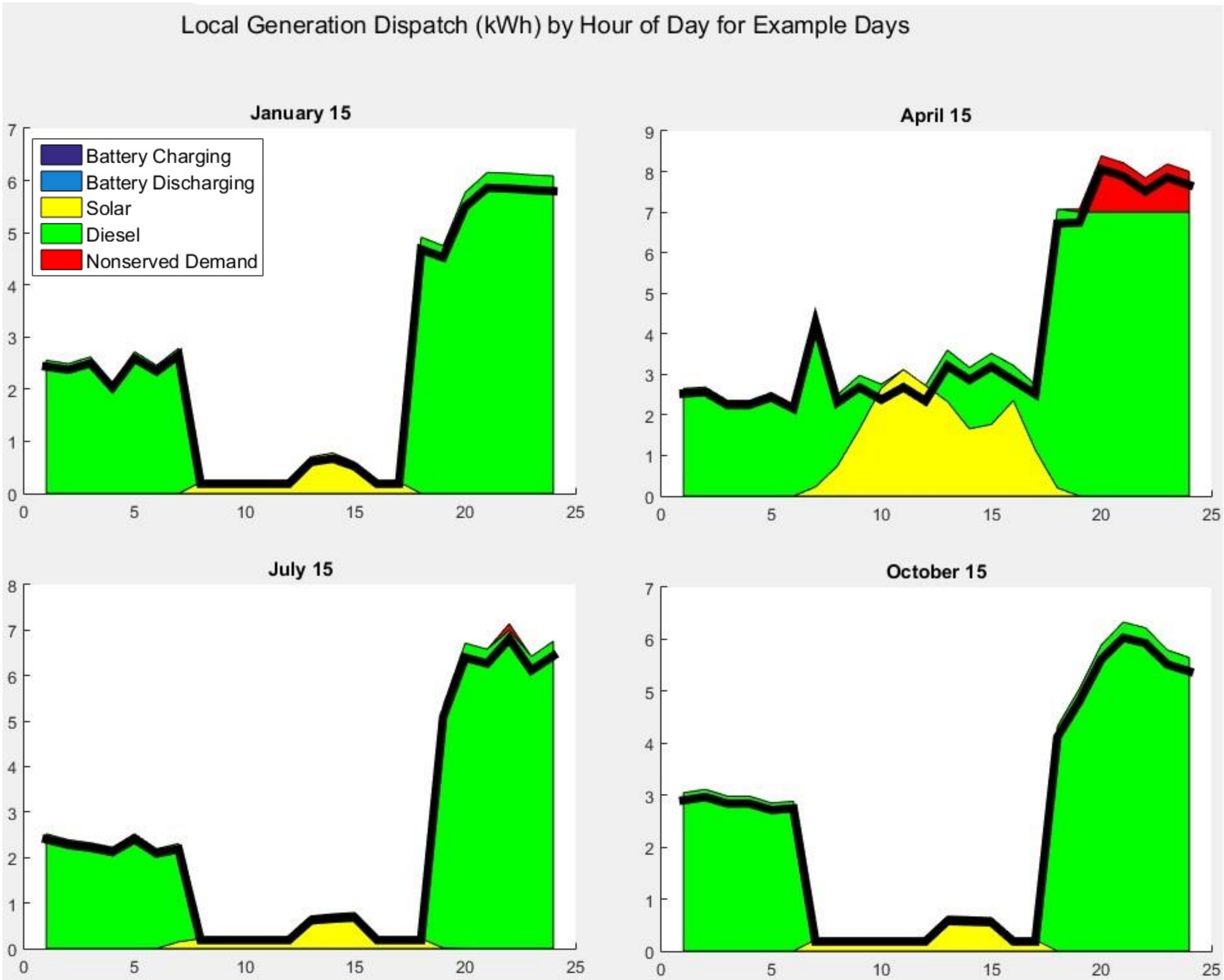
Demand growth scenario local generation study results for region 1:

Number of Customers	Peak Demand (kW)	Average Demand (kW)	Solar Capacity (kW)	Battery Capacity (kWh)	Generator Capacity (kW)	Fraction of Demand Served	Financial Cost per Demand Served ($/kWh)
1	0.19	0.05	0	0	0	0.00	NA
2	0.39	0.11	0.5	1.38	0	0.35	0.76
3	0.58	0.16	0.5	2.76	0	0.33	0.64
4	0.77	0.21	0.75	4.14	0	0.37	0.63
10	1.93	0.53	1	0	1	0.90	0.63
31	5.74	1.65	3.75	0	4	0.97	0.46
96	17.48	5.09	10.5	0	15	0.99	0.43
301	54.39	15.97	30	0	20	0.75	0.36
942	169.70	49.98	108.5	0	150	0.99	0.36
2949	530.00	156.49	288	0	400	0.98	0.33
9231	1656.81	489.92	898.5	0	1333	0.98	0.33
28902	5183.15	1533.74	936.5	0	2250	0.81	0.31

Operation simulation for solar-battery system (demand indicated by the black line):



Operation simulation for diesel-solar system (demand indicated by the black line):



6.3 CLUSTERING

The following tables show examples of clustering results for various scenarios. Only region 1 results are shown for scenarios in which region 1 refers to the same geographical area, just to give a sense of the nature of the results and some differences between cases that can be observed visually. Each analysis region has around 20 thousand customers, which are difficult to show in a single image. Thus, one should be careful about the inferences drawn from the images.

In the first table, off-grid clustering results are shown for the baseline and demand growth scenario. There is one cross for each customer, and customers in the same off-grid cluster have the same color (there are many more clusters than microgrids, so colors repeat). Qualitatively, the two results are similar, with many relatively small off-grid clusters. It is possible to get a deeper sense of the differences in the clustering results by looking at distributions in cluster sizes.

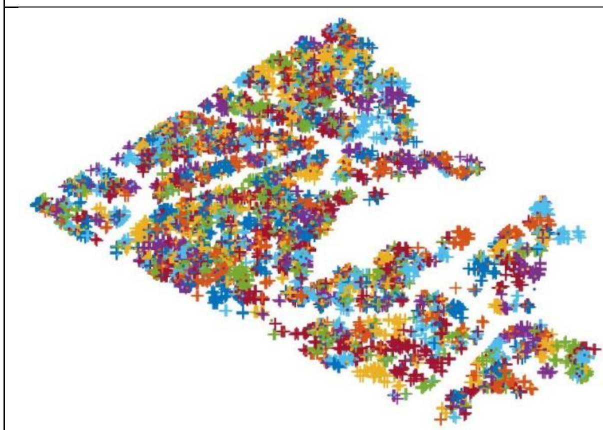
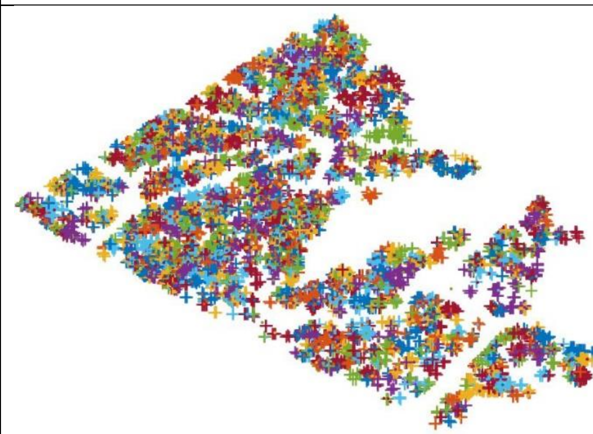
The second table shows grid extension clustering results. There is a dot for each customer, where customers in the same grid extension cluster have the same color.

The left column shows clustering results from runs in which the first clustering logic was used. In this clustering logic, the cost of a single grid extension is compared to the cost of two separate grid extensions, or to one grid extension and one microgrid. Under this clustering logic, very small grid extension clusters are produced when grid reliability is poor, and large grid extension clusters are produced for the perfect grid reliability case. This is likely driven by the trend that for small groups of customers, microgrid is typically more cost-effective than a grid extension with an unreliable grid.

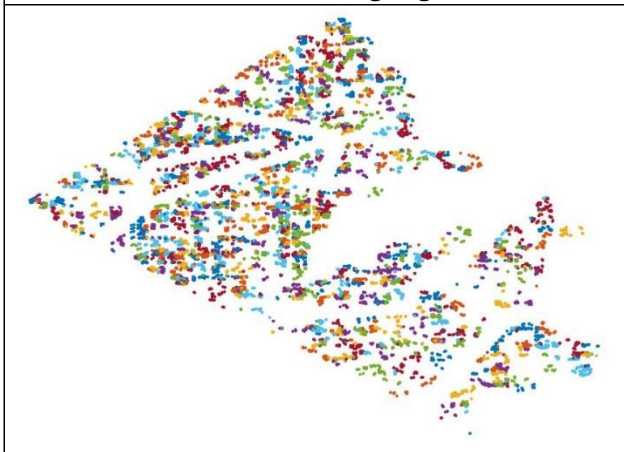
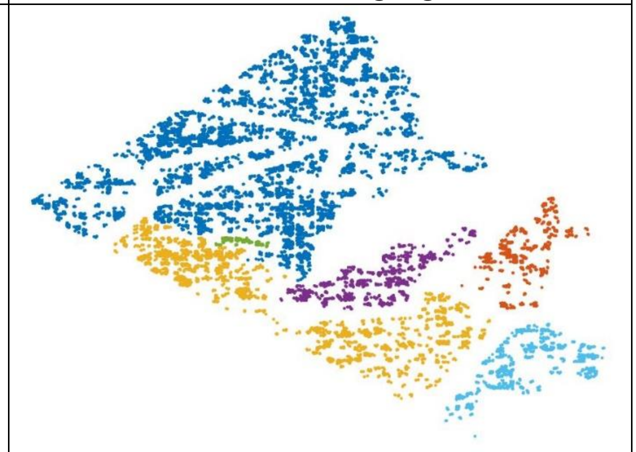
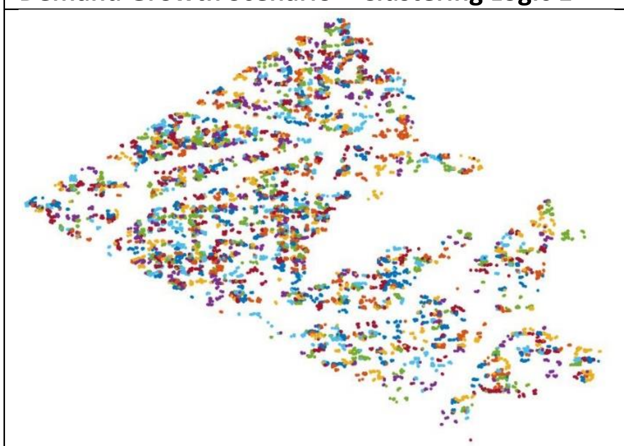
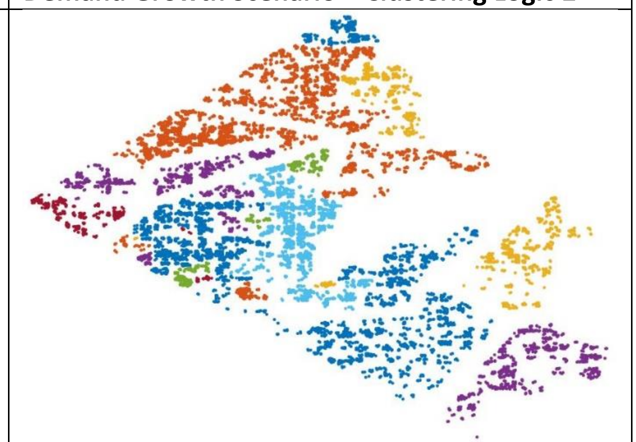
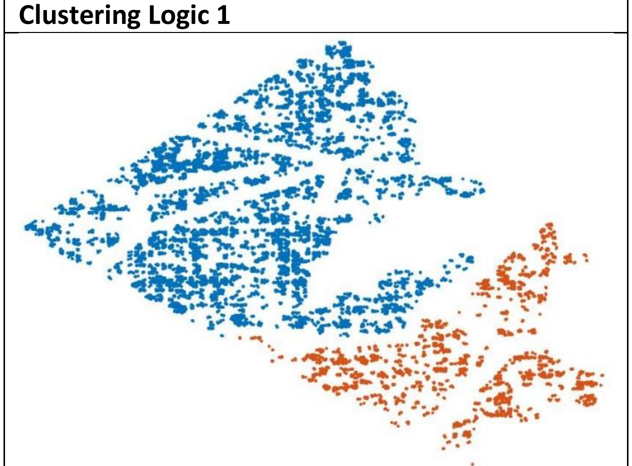
The right column shows results for cases with the second clustering logic, in which the cost of a single grid extension is only compared to the cost of two separate grid extensions—i.e. the option of one grid extension and one microgrid is not considered. In this case, much larger grid extension clusters are produced. There are more and smaller clusters in the demand growth scenario compared to the baseline scenario, perhaps reflecting that the higher demand per customer means that less customers can be served by a single medium voltage grid extension.

Overall it seems that the second clustering logic performs better for the Vaishali case, because the many small grid extension clusters produced in the first clustering logic stand no chance of being cost-effective grid extensions. However, it is possible to think up scenarios in which the first clustering logic performs better. Ultimately, each possible clustering logic may perform better or worse under certain circumstances. In future work, clustering options should be examined under a wider range of scenarios (e.g. real cases from other parts of the world) to get a better sense of which clustering logics perform well over a range of scenarios.

Off-grid clustering for region 1:

Baseline Scenario	**Demand Growth Scenario**
	

Grid-extension clustering for region 1:

Baseline Scenario – Clustering Logic 1	**Baseline Scenario – Clustering Logic 2**
	
Demand Growth Scenario – Clustering Logic 1	**Demand Growth Scenario – Clustering Logic 2**
	
Baseline Scenario + Reliable Grid – Clustering Logic 1	
	

6.4 Power System Designs

The following tables show images of the power system designs for the first analysis region from the same cases shown in the clustering results above.

Here is the key for the images:

- Black lines represent existing medium voltage lines
- Green circles represent microgrid generation sites. They are connected to red dots representing microgrid customers through blue lines representing low voltage distribution lines.
- Orange dots represent isolated customers.
- Red lines represent new medium voltage lines for grid extensions. They are connected to blue lines representing low voltage distribution lines for grid extensions.

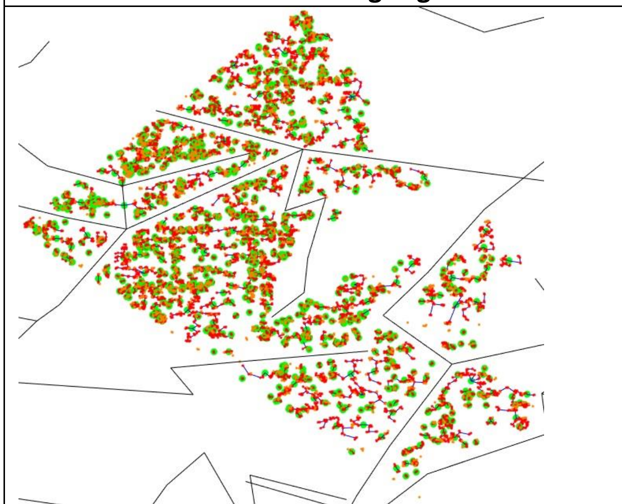
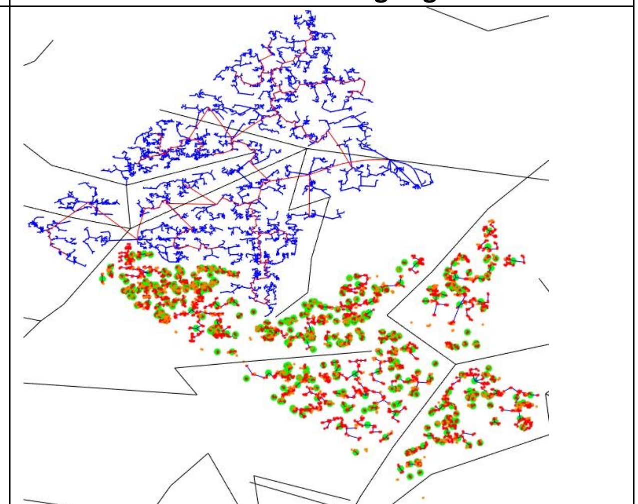
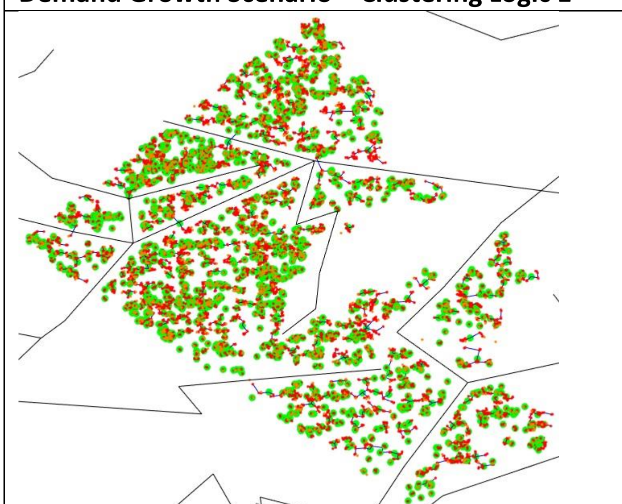
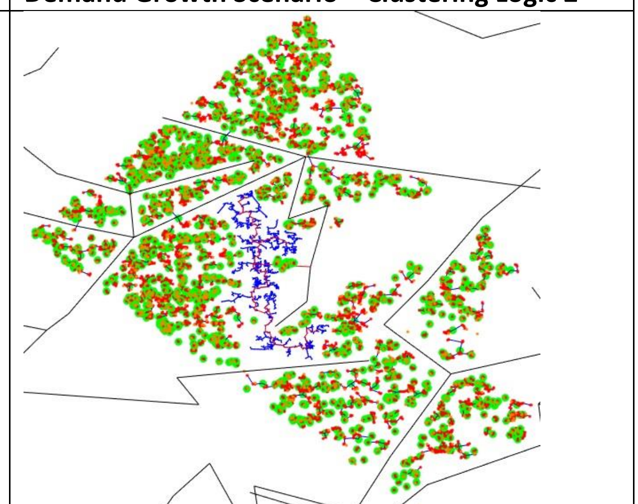
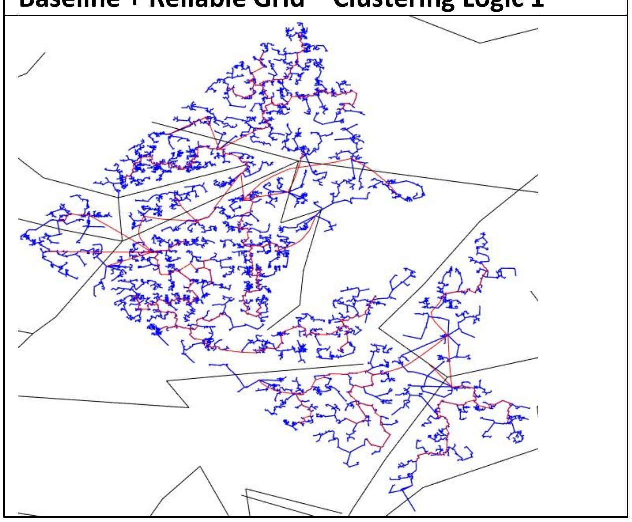
While it is difficult to see the details of the images given the large number of customers, it is possible to clearly see which parts of the region get off-grid systems or grid extensions.

As expected based on the clustering results which produces many small grid extension clusters, the baseline and demand growth scenarios with the first clustering logic produce no grid extensions and select the entire region for off-grid.

Alternatively, the second clustering logic for these scenarios leads to grid extensions for parts of the region. Interestingly, the parts of the region which get grid extension in the two cases are significantly different. I believe the basic intuition driving the decision between off-grid and grid extension is related to the idea that larger microgrids generally look better than large grid extensions (when grid reliability is poor), but grid extensions still look better than the smaller microgrids. Since the larger and smaller microgrids are all mixed up together in each grid extension cluster, then choice between grid and off-grid depends on the distribution of off-grid system sizes in that grid extension cluster. Obviously, this is a simplification of the situation, but I think the general intuition is valid.

For the case with perfect grid reliability, grid extension is selected for all customers. This is sensible, since Vaishali is a relatively densely populated place with decent coverage of medium voltage power lines.

Power system designs for region 1

Baseline Scenario – Clustering Logic 1	Baseline Scenario – Clustering Logic 2
	
Demand Growth Scenario – Clustering Logic 1	**Demand Growth Scenario – Clustering Logic 2**
	
Baseline + Reliable Grid – Clustering Logic 1	
	

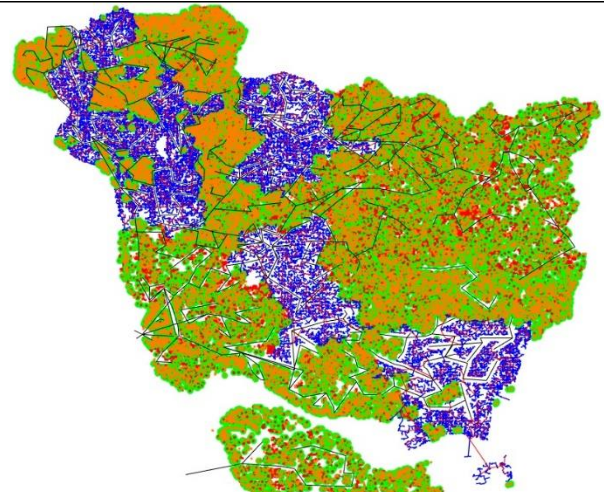
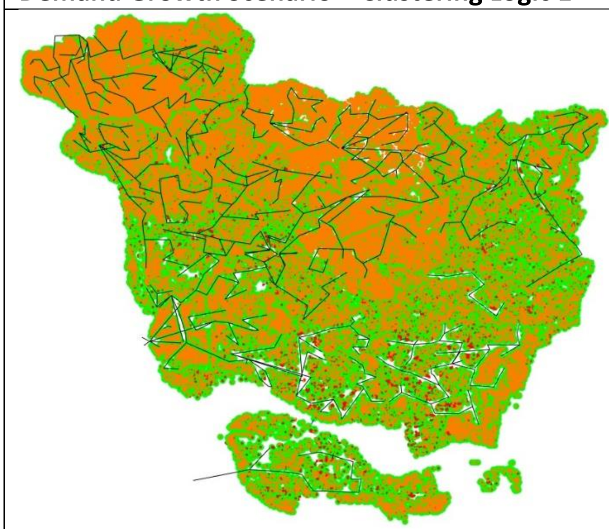
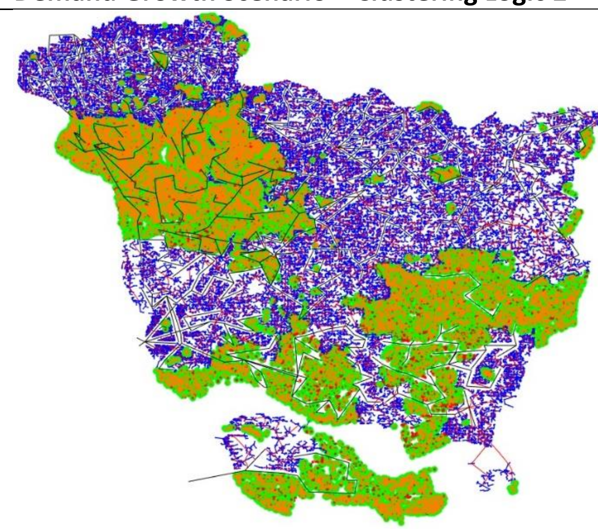
6.5 Vaishali District Results

The first table below shows images of the system design results for the entire district of Vaishali for many cases. In general, these images show the same trends as what was observed in the single analysis result images described in the previous section. The first clustering logic leads to all off-grid systems, unless there is a more reliable grid. The second clustering logic leads to a mix of off-grid and grid extensions. The differences between which areas get grid extension and off-grid varies, but not in a straightforward fashion.

The subsequent tables show some summary metrics of the results for the cases. Here are some observations based on looking at these results:

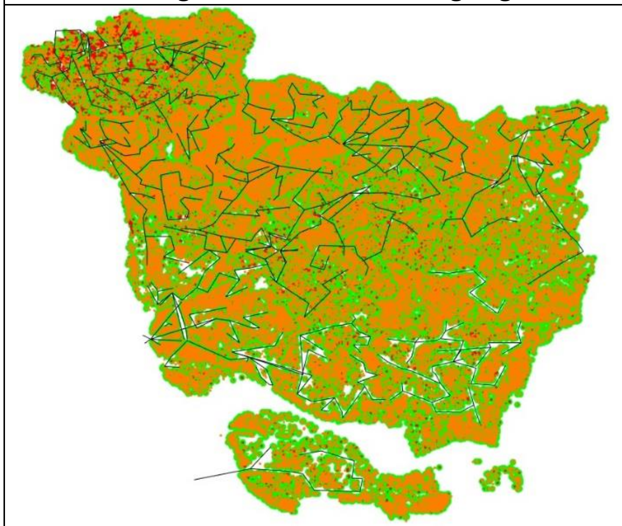
- Despite the similarity between the generation lookup table for the baseline run and the repeat of the baseline run, there are still significant differences in final results (with costs varying on the order of 10%). This may reflect that small differences in the generation lookup table have significant impacts to the clustering results. Alternatively, there could be other features of the logic that are driving this difference (perhaps variation in assignment of demand to customers in the clustering and design steps). This deserves further investigation in future work.
- Comparing the results of a given scenario with the two alternative clustering logics supports the intuition I described in the previous section about the choice between off-grid and grid extension—basically that grid extension is more likely to be preferred where more of the alternative off-grid systems are small. In going from clustering logic 1 to clustering logic 2 for all of the demand scenarios (baseline, demand growth, and more buildings) a greater proportion of isolated systems are converted to grid extensions than are microgrids. Thus it seems that grid extensions on average look more favorable in places that cannot support larger average off-grid system sizes.
- As demand increase, either through demand growth or adding more customers, a greater percentage of customers are selected for grid extension when the second clustering logic is used.
- Across solutions with comparable total social cost per customer (financial cost plus cost of non-served energy), there can be wide differences in the split of financial and non-served energy cost. Financial costs are the more important consideration for feasibility of implementation, since distribution companies and other system providers may have limited financial resources.

Images of final designs:

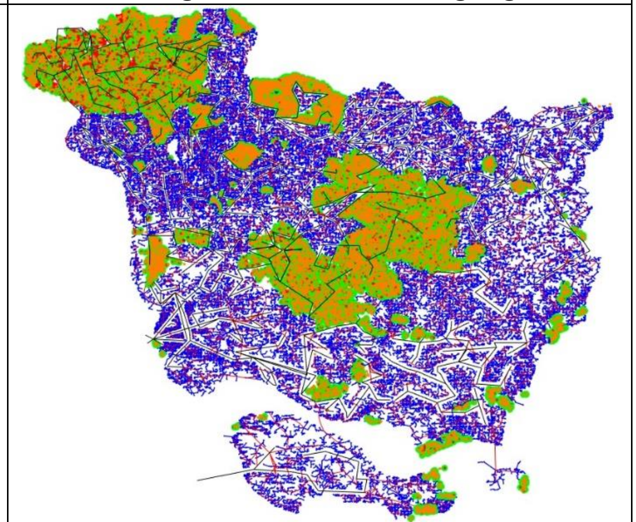
Baseline Scenario – Clustering Logic 1	**Baseline Scenario – Clustering Logic 2**
	
Demand Growth Scenario – Clustering Logic 1	**Demand Growth Scenario – Clustering Logic 2**
	

Images of final designs (continued):

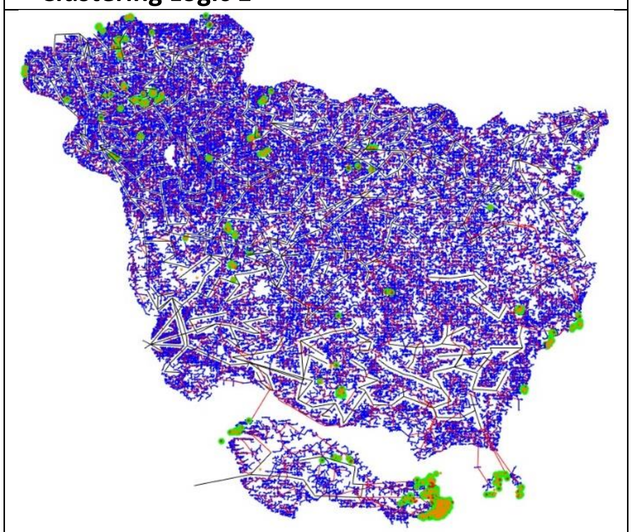
More Buildings Scenario – Clustering Logic 1



More Buildings Scenario – Clustering Logic 2



Baseline Scenario + Reliable Grid – Clustering Logic 1



Results Summary: Baseline Scenario – Clustering Logic 1:

	Microgrids	Isolated Systems	Grid Extensions	All
Number of Customers	423594	76328	0	499922
Fraction of Customers	0.85	0.15	0.00	1.00
Financial Cost Per Customer ($/yr)	107	0	NaN	91
Non-served Energy Cost Per Customer ($/yr)	60	213	NaN	84
Total Financial Cost ($/yr)	4.5E+07	0.0E+00	0.0E+00	4.5E+07
Total Financial and Non-served Energy Cost ($/yr)	7.1E+07	1.6E+07	0.0E+00	8.7E+07

Results Summary: Baseline Scenario – Clustering Logic 1 – Repeat Trial:

	Microgrids	Isolated Systems	Grid Extensions	All
Number of Customers	395984	103938	0	499922
Fraction of Customers	0.79	0.21	0.00	1.00
Financial Cost Per Customer ($/yr)	99	0	NaN	79
Non-served Energy Cost Per Customer ($/yr)	80	214	NaN	108
Total Financial Cost ($/yr)	3.9E+07	0.0E+00	0.0E+00	3.9E+07
Total Financial and Non-served Energy Cost ($/yr)	7.1E+07	2.2E+07	0.0E+00	9.3E+07

Results Summary: Baseline Scenario – Clustering Logic 2:

	Microgrids	Isolated Systems	Grid Extensions	All
Number of Customers	290852	35676	173394	499922
Fraction of Customers	0.58	0.07	0.35	1.00
Financial Cost Per Customer ($/yr)	113	0	43	81
Non-served Energy Cost Per Customer ($/yr)	47	213	130	88
Total Financial Cost ($/yr)	3.3E+07	0.0E+00	7.5E+06	4.0E+07
Total Financial and Non-served Energy Cost ($/yr)	4.6E+07	7.6E+06	3.0E+07	8.4E+07

Results Summary: Baseline Scenario – Clustering Logic 1 – Reliable Grid:

	Microgrids	Isolated Systems	Grid Extensions	All
Number of Customers	5054	1002	493866	499922
Fraction of Customers	0.01	0.00	0.99	1.00
Financial Cost Per Customer ($/yr)	100	0	79	80
Non-served Energy Cost Per Customer ($/yr)	77	214	0	1
Total Financial Cost ($/yr)	5.0E+05	0.0E+00	3.9E+07	4.0E+07
Total Financial and Non-served Energy Cost ($/yr)	8.9E+05	2.1E+05	3.9E+07	4.0E+07

Results Summary: Demand Growth Scenario – Clustering Logic 1:

	Microgrids	Isolated Systems	Grid Extensions	All
Number of Customers	359105	140817	0	499922
Fraction of Customers	0.72	0.28	0.00	1.00
Financial Cost Per Customer ($/yr)	200	20	NaN	150
Non-served Energy Cost Per Customer ($/yr)	94	359	NaN	169
Total Financial Cost ($/yr)	7.2E+07	2.8E+06	0.0E+00	7.5E+07
Total Financial and Non-served Energy Cost ($/yr)	1.1E+08	5.3E+07	0.0E+00	1.6E+08

Results Summary: Demand Growth Scenario – Clustering Logic 2:

	Microgrids	Isolated Systems	Grid Extensions	All
Number of Customers	178000	19614	302308	499922
Fraction of Customers	0.36	0.04	0.60	1.00
Financial Cost Per Customer ($/yr)	212	2	65	115
Non-served Energy Cost Per Customer ($/yr)	64	370	227	175
Total Financial Cost ($/yr)	3.8E+07	3.5E+04	2.0E+07	5.8E+07
Total Financial and Non-served Energy Cost ($/yr)	4.9E+07	7.3E+06	8.8E+07	1.4E+08

Results Summary: More Buildings Scenario – Clustering Logic 1:

	Microgrids	Isolated Systems	Grid Extensions	All
Number of Customers	419691	134389	0	554080
Fraction of Customers	0.76	0.24	0.00	1.00
Financial Cost Per Customer ($/yr)	97	0	NaN	74
Non-served Energy Cost Per Customer ($/yr)	82	214	NaN	114
Total Financial Cost ($/yr)	4.1E+07	0.0E+00	0.0E+00	4.1E+07
Total Financial and Non-served Energy Cost ($/yr)	7.5E+07	2.9E+07	0.0E+00	1.0E+08

Results Summary: More Buildings Scenario – Clustering Logic 2:

	Microgrids	Isolated Systems	Grid Extensions	All
Number of Customers	**134969**	**22096**	**397015**	**554080**
Fraction of Customers	**0.24**	**0.04**	**0.72**	**1.00**
Financial Cost Per Customer ($/yr)	**109**	**0**	**43**	**57**
Non-served Energy Cost Per Customer ($/yr)	**53**	**214**	**130**	**115**
Total Financial Cost ($/yr)	**1.5E+07**	**0.0E+00**	**1.7E+07**	**3.2E+07**
Total Financial and Non-served Energy Cost ($/yr)	**2.2E+07**	**4.7E+06**	**6.9E+07**	**9.5E+07**

6.6 Discussion of the Vaishali study

Overall, the experience with this electrification planning study has shown that it is possible to use a computer tool to automatically produce power system designs for a large area requiring electricity access. While the data gathering process was challenging and not perfect, in the end the model still produces sensible results.

By splitting up analysis regions and solving them in parallel, it was possible to produce a full set of results for the entire district in around 7 hours for each case. This makes it feasible to consider many cases in a planning exercise to explore alternative policy or technical options, or to explore data sensitivities. This opens up the possibility for much more rigorous and accurate planning of electrification policy and implementation than was previously available.

It is important to remember, however, that one should be cautious in interpreting the results of REM for Vaishali district, and in general. Some considerations related to interpretation of REM results and limitations of the model are listed below:

- Difference in the REM results from run to run could sometimes be based on idiosyncrasies in the model logic rather than “real” differences between the cases. The analyst should be careful to ensure the model is delivering consistent results before making major conclusions.
- In the Vaishali case, REM was only aiming to minimize total social cost, but was not considering many other important considerations than could impact the feasibility or desirability of solutions. Other considerations include access to financing or sufficient subsidy, ability of businesses to support installation and operation of the systems, ability of consumers to pay, and local preferences. For example, customers or businesses might prefer to consolidate small microgrids into larger ones, if the costs are not dramatically different, or they could prefer the opposite.
- While REM provides detailed designs down to the customer level, errors in local data and local implementation considerations will often necessitate modifications to the specific designs. Thus REM is best suited to producing aggregate metrics about the systems over larger areas, rather than information about very specific locations.

- REM considers demand as given, and thus does not consider reducing demand through more efficient appliances as an option. In practice, efficient appliances are especially relevant to remote areas where the cost per unit energy is very high.
- REM in the Vaishali considered only a single AC architecture for power systems. In practice, different architectures may be desirable in some cases, such as DC systems or centralized battery charging without distribution wires. Preferred architectures in practice depend on many more considerations than cost minimization.
- REM is a static model, yet it includes many parameters which are dynamic and uncertain. Special care should be given to such parameters including demand, reliability of the grid, and costs of grid power and diesel fuel.
- REM in the current implementation does not consider costs of upstream upgrades to the grid. This feature should be added in future work.
- Since many systems produced by REM have far from perfect reliability, the costs of non-served energy are very significant and should be treated with care. REM assumes non-served energy costs are discounted equivalently to financial costs, which is not necessarily always a justified assumption.
- In deciding between off-grid systems and grid extension for an area, there will often be winners and losers for either case, in terms of total social cost per customer and quality of electricity service. Thus some level of community engagement may be important to make the right decisions.
- Based on the clustering of systems, customers may be put into more or less expensive systems and better or worse performing systems more or less by chance. A policy solution may be important to ensure that decisions about service and tariffs are equitable to an appropriate extent.
- Even if a particular solar microgrid can serve less demand than the grid, the microgrid generally has the advantage of providing more predictable service. Most of the critical evening peak demand is covered on most days by the microgrid, whereas the grid is least reliable during the peak hours of critical demand and its availability may be subject to political interference.
- REM considers investments in a single step, but in reality multi-step planning is important. This should be considered in future work. An especially interesting angle would be to explore if it makes sense to build grid-compatible microgrids in the near-term, which can be connected to the grid at a later time. This, of course, would also depend on the development of technical and legal/business standards for the connection of microgrids to the grid.

7 FUTURE WORK ON REM

This chapter lists some of the priority areas for future work on the Reference Electrification Model. Many of these topics are discussed in more context and detail in the preceding chapters, but this list can serve as a convenient reference that consolidates ideas about future work.

General improvements:

1. REM is currently implemented as a static model (in that variation in parameters from year to year are not considered) with a single decision step (in that only one set of investment decisions is made). REM could be improved by considering year-to-year variations in parameters such as demand, degradation of components, reliability of power from the grid, and fuel costs. REM could also be improved by considering that investment decisions could be made in multiple stages (e.g. build microgrids in one year and then connect the microgrids to the grid in a later year).
2. It would be beneficial to run cases of REM with additional sensitivities to get a better sense of how different variables influence the final results. Some care should be taken to determine where these influences are "real" effects of the variable on the optimal design, versus idiosyncrasies or errors in the model logic.
3. For a given set of input data and parameters, it appears that REM can still have some significant variation in final results from run to run. Some care should be taken to determine the source of this variability, and to modify the model and/or produce guidelines for its use so that this variability is minimized, if possible.
4. Many of the heuristics contained within REM and potential alternatives will tend to work better or worse in different cases, depending on the input data. REM should be applied to a wider range of relevant cases in order to get a sense of which approaches perform well in which situations. The approaches that are successful over a broad range of scenarios should be used by default.

Customer data:

5. Improvements should be made to the method of identifying customers, including improving the building identification algorithms and adding the ability to include customers that are not associated with buildings.
6. If multiple types of customers are to be considered for a given run of REM, there will need to be a way to assign each customer to a customer type. Work could be done on assigning customer type based on extracted features from the satellite imagery, such as assigning larger demand to larger buildings.
7. Improvements should be made to the method of estimating the locations of customers that are already connected to the grid.

Demand:

8. The current representation of residential demand should be validated with real data. Additionally, additional demand representations for other customer types should be produced. For example this could include residential customers with various appliance sets, agricultural customers, and commercial customers.

9. In a dynamic or multi-step model, it would be interesting to look at changes in demand from year to year other than proportional demand growth. This could be represented through customers transitioning from one demand type to another.

Clustering:

10. Different clustering logics should be explored in a variety of scenarios to get a better sense of which perform well in which cases. Changes to the clustering logic could include starting with the assumption of all customers being connected (rather than starting from them being separate) and having tighter integration with the clustering logic that is internal to RNM for grid extension design.
11. The performance of the cost difference estimates in the clustering step should be explored in more detail.
12. If multiple customer types are used, the calculation of the location of the cluster's center should be weighted by demand.

Off-grid generation design:

13. The off-grid system costs, including the method of parametrizing the costs and the particular values, should be improved based on a study of actual costs of off-grid systems. So far, companies have been very protective of this information.
14. Alternate architectures for the local generation site should be considered. These include:
    a. Generation assets distributed throughout a microgrid, to reduce network costs or support customer autonomy or new business models.
    b. More than one diesel generator, to allow more efficient diesel generator operation.
    c. Inclusion of other generation sources, such as micro-hydro.
    d. Grid-interactive architectures.
15. The method of producing a generation lookup table should be modified to accommodate multiple customer types.
16. The method of selecting the solar panel and battery type should be improved.
17. The operation simulation could be modified to minimize the impact of the first hours of the simulation, before the battery state of charge reaches "steady state."
18. Alternative dispatch logics should be considered for the operation simulation, including considering forecasting, optimization, and matching operations with various system architectures.
19. In the current dispatch logic, the cost of curtailing demand as a resource to charge the battery should be corrected.
20. The representation of losses in the generation simulation should be improved, so that they are not just proportional to energy flow.
21. Improved representations of PV capacity degradation, battery capacity degradation, and other year to year effects should be explored.
22. Improvements should be made in using the lookup table to produce the final designs and costs for microgrids. This could include making adjustments based on local prices (such as a remoteness penalty), and based on the expected losses (which will vary with the size and shape of the microgrid). Additionally, the method of linear interpolation to determine the final design should be improved so that it does not produce infeasible designs.

Distribution network design:

23. The data about the locations and components of the existing grid should be validated and improved.
24. The costs of providing energy from the grid to various parts of the network should be validated. The ability to better account for different energy costs from different parts of the grid should be included.
25. The representation of and data about availability of power from the grid could be improved to account for seasonal effects and variations in availability in different parts of the network.
26. The network catalog should be simplified to reduce the data collection burden, especially by removing features that are less relevant for rural electrification.
27. The catalog values for the current case should be validated with more actual data from grid extension and microgrid projects.
28. Additional catalogs should be produced based on lower-cost distribution options.
29. The catalog and model should be expanded to better consider single phase and DC power distribution.
30. The consideration of the grid extension option should be improved through considering more options for grid extension (including extensions in low voltage and high voltage) and considering upstream costs that may be imposed by extensions. This could be accomplished by applying a version of the brownfield RNM.
31. New RNM modules should be developed that are catered to the REM needs in microgrid design, grid extension design, and network cost estimates. This could improve the quality of results, by avoiding clumsy post-processing and some errors that are caused in this step. Additionally, this could improve model execution by integrating the model into REM, rather than requiring the writing and reading of many files to pass data. One of the features of these new modules should be to consider networks that are not sized for peak demand, and are instead capacity restricted at some times, since this is a common feature of networks in the developing world.

8 PROPOSED REGULATORY REFORMS TO SUPPORT ELECTRICITY ACCESS IN INDIA AND THE ROLE OF REM

While REM is a useful and valuable tool in its ability to find cost-efficient technical system designs, there are many other systemic challenges that will still impede progress towards universal electricity access. Regulatory reform is one important avenue to address many of these challenges. Given that electricity access might be best provided through collective systems that may become natural monopolies, a layer of government oversight is often critical to support efficient and fair outcomes. Regulation serves as a link between electrification policy and implementation, and forces officials to make tough tradeoffs between conflicting objectives.

Regulatory reform has been a hot topic in India over the past several decades, and major reforms were introduced through the Electricity Act of 2003. While that regulation was primarily focused on improving the performance of the national grid, it also provided some support to the relative immature industry of off-grid system providers by allowing companies to distribute electricity in rural areas without a license. That provision was likely important to supporting the growth in off-grid system providers over the last decade.

Now with some more years of perspective and industry and technology development, it is becoming clear that India's regulation regarding electricity access is not sufficient—either for the provision of reliable electricity service through the existing grid, or for the adequate coverage of other customers through off-grid systems. Regulation of off-grid systems, in particular, has received attention of late. For example in 2012 India's Central Electricity Regulatory Commission sponsored the writing of draft regulation for off-grid electrification, which provides a pathway for off-grid system providers to become franchisees of the distribution company and eventually connect to the main grid.

This chapter discusses some of my ideas about potentially useful regulatory reforms to support universal electricity access in India. It also highlights some opportunities for computer models, such as REM, to support more effective regulation, especially in the harmonization of grid extension and off-grid system implementation.

8.1 OVERVIEW OF A NEW REGULATORY FRAMEWORK FOR ELECTRICITY ACCESS

8.1.1 Rethinking the universal service obligation

Under current regulation in India, electricity distribution companies (DISCOMs) (which typically manage distribution infrastructure and retailing of electricity) are legally sanctioned monopolies with universal service obligations for all customers within their service region. This means that they are obligated to provide electricity service to any customer who wants it and pays the appropriate tariffs. If this service obligation was upheld, electricity access would be a non-issue. DISCOMs, however, cannot meet this universal service obligation for two main reasons. Firstly, for many of those customers who are already connected, tariff collection falls far short of the cost of providing electricity service, so it is unaffordable to provide reliable service. Secondly, many of the customers that are not yet connected to the grid live in remote rural areas and are poor, meaning that the high cost of connection to the grid may not be recouped through the low volume of electricity sales. Under the current power sector structure, the universal service obligation is impossible to achieve, and thus it is largely ignored and not enforced.

Even if a DISCOM had the financial resources (e.g. via subsidy) to provide reliable grid service to all customers within a service area, this would not necessarily be a prudent use of resources. For some customers, especially those living in very remote areas and those who would demand small volumes of electricity, an off-grid system—such as a microgrid or home system—would be the least-cost approach to providing appropriate electricity service. While DISCOMs could hypothetically provide or facilitate off-grid systems in order to meet a universal service obligation, the business of managing off-grid power systems is quite different from the usual business of a DISCOM in a centralized power system. Additionally, the best methods (in terms of technologies and business models) for implementing and managing off-grid power systems are not yet well-understood, so many new ideas will need to be tried, which could more easily come from many competing entities rather than a few DISCOMs.

While a DISCOM, as a sanctioned monopoly, ought to have some service obligation, a service obligation for customers which the DISCOM is not equipped—financially or functionally—to serve does not make sense. Thus, the universal service obligation of the DISCOMs should be modified so that they are obligated to provide a reasonable quality of service to all those customers whose least-cost electrification mode is the grid. Concurrent with this, other regulatory changes must be made to enable that this level of service is actually provided. Section 2 of this chapter discusses changes to the current regulation that should be made to enable acceptable electricity service for grid-connected consumers.

For those customers that would be better served by off-grid systems, a totally new regulatory approach is required. This approach should accommodate the fuzzy interface between the areas appropriate for on-grid and off-grid electrification and distinguish between various potential business and technical approaches. Recommendations for the regulation of off-grid systems are described in section 3.

8.1.2 How to determine appropriate electrification modes

If different regulatory approaches are to be used for different modes of electricity access (grid, collective off-grid, and individual off-grid) it is necessary to determine which modes are appropriate for which regions.

Regulators should make this determination considering the goal of maximizing social welfare, via the minimization of the cost of electricity service and the social costs of non-served electricity demand and lack of electricity access. Other relevant social and practical considerations should also factor into these determinations.

Estimates of least cost electrification modes for various areas could be estimated based on judgment and experience or rough indicators, such as population density, distance from the grid, and income level. However, since profit margins and subsidy dollars tend to be tight in the world of rural electrification, a more sophisticated approach may be warranted to support efficient use of financial resources. This is a place where computer models such as REM can provide significant value.

8.2 Changes to Grid Regulation

For those customers in India who would be most efficiently and effectively provided electricity access through the main grid—presumably including those in areas which already have grid connections—significant regulatory changes are necessary to enable true electricity access. Some recommendations for those regulatory changes are described and justified below.

8.2.1 Addressing incentives of the DISCOMs

A first regulatory change which would support reliable electricity access in grid-connected areas is the restructuring of the entities responsible for distribution and retail. Currently, the activities of managing the electricity distribution infrastructure and selling electricity to end consumers are typically undertaken by a single entity, known as a distribution company, or DISCOM. These DISCOMs are usually state-owned companies which were established around the time the Electricity Act, 2003 required unbundling of the State Electricity Boards into separate generation, transmission, and distribution companies. The State Electricity Boards were elements of the state governments that managed the vertically integrated and state-owned power system in each state. In the spirit of power system liberalization activities around the world, the Electricity Act was intended to promote efficiency and competition in order to reduce costs, improve service, and encourage private investment. In order to ensure that these benefits were achieved, independent regulatory agencies were implemented to oversee the power sector.

In reality, however, regulators have had limited success in driving DISCOMs towards efficiency and competition. DISCOMs are often driven instead by short-term political goals, due to the close connection between DISCOMs and the state governments (Kumar & Chatterjee, 2012). Activities of DISCOMs in line with these short-term political considerations include:

- Allowing tariffs to be below the cost of service and allowing low payment collection rates and rampant theft, in order to generate votes
- Procuring supply through politically favored companies or state-owned entities (often under cost-plus contracts) rather than via competition to ensure least cost
- Inefficient operations through maintaining excessive staffs (potentially related to providing jobs as political favors)
- Being slow to enable open-access to the distribution system, in order to reduce competition and maintain political power (Kumar & Chatterjee, 2012).

As a result of these activities, DISCOM costs are high and revenues are low, meaning that they are constantly losing money. As a state-owned company, this revenue inadequacy is mitigated in the short-term by state subsidies and state debt. But over time, the burden of subsidies and debt have become too large, which has created additional problems. DISCOMs are unable to adequately invest in distribution infrastructure, which means they cannot afford adequate grid extension (directly impacting energy access). They cannot afford required maintenance and upgrades for existing infrastructure, which leads to high losses (further increasing cost) and poor service. Their poor financial state make DISCOMs risky business partners, and they are not able to attract sufficient investments in generation (Kumar & Chatterjee, 2012).

Eventually, due to lack of capacity and high costs, DISCOMs are forced to resort to massive load-shedding, resulting in frequent blackouts and unpredictable service. The costs are highest and revenues lowest in the rural areas, leading to especially poor service in those areas. In the cities the frequent blackouts are a significant frustration and economic burden, often requiring households and businesses to invest in expensive backup systems. In many rural areas, however, the service is so poor that it is debatable whether it should count as even basic electricity access. For example, a 2006 survey found that households in Bihar faced blackouts nearly every day and the average outage was for 16 hours (Santhakumar, 2008).

8.2.1.1 Improving strength and independence of regulators

The myopic and politically-driven activities of many of the DISCOMs in India are certainly not in the spirit of the reforms of the Electricity Act, 2003. But the blame should not be placed solely on them. Under the Electricity Act, it is the responsibility of the regulators to oversee the activities of the DISCOMs and other utilities to enforce compliance with relevant laws and policies. While the DISCOMs have incentive to act in their own self-interest (typically reflecting the interests of politicians who are in power), the regulators are supposed to be independent from politics and act in the long-term interest of the people and investors. Unfortunately, state regulators in India often lack the independence and strength to act effectively.

Regarding independence, several factors keep the regulators interests in line with the interests of state politicians and the DISCOMs. These include:

- Lack of financial independence, due to dependence on the state governments for grants (Kumar & Chatterjee, 2012)
- Selection of regulators based on considerations of political patronage, rather than regulatory experience (Bhattacharyya & Palit, 2013).

Regarding strength, regulators have limited instruments to enforce compliance with regulations. Typically, regulators may use financial incentives to encourage good behavior by regulated entities. But when the regulated entity is a state-owned company, profits are secondary to political objectives and fines can be absorbed by state budgets. Thus, even an independent regulator with the best intentions would have limited ability to enforce that the state-owned DISCOMs follow the regulation. In practice, most regulators do not even attempt to impose any penalties (Bhattacharyya & Palit, 2013).

The following changes are recommended to improve the ability of regulators to effectively enforce regulation that supports current and emerging laws and policies:

- Guarantee regulatory agency funding outside of the annual state budgeting process, in order to support financial independence
- Require that at least one member of the regulatory commission be a non-Indian with regulatory experience, in order to encourage competence and impartiality of the commission
- Allow the regulator to impose personal fines against management of utilities and to revoke the license of utilities in extreme cases, in order to give effective enforcement tools for regulation of state-owned entities.
- Mandate regulators to make public detailed statistics of quality of service of the DISCOMs that they regulate and to establish comparison with the data from other states and other parts of the world.

8.2.1.2 Restructuring of distribution and retail

Currently distribution and retail are regulated as a single activity, which makes regulation challenging and creates some perverse incentives.

One issue is that there is no recognition of retailing as an authorized separate activity in the Electricity Act, which makes retail competition practically nonexistent. This means that a competing retailer would need to also be a distributor and thus would need a redundant distribution infrastructure. Since distribution is a natural monopoly, this leads either to unnecessary costs or a lack of retail competition.

In practice, there have been few cases of more than one co-located distribution company (Kumar & Chatterjee, 2012). Distribution and retail should be recognized as separate activities in the Electricity Act so that it is at least possible for retail competition to exist on a single distribution infrastructure.

Secondly, since distribution and retail are performed together, it is difficult to apply best practice regulation regarding allowed revenues and incentives to each activity. There should be at least accounting unbundling between distribution and retailing to allow separate and appropriate economic regulation to be used for each activity. This accounting unbundling is also recommended in (Kumar & Chatterjee, 2012).

These two changes—recognizing distribution and retail as separate activities and requiring at least accounting unbundling between distribution and retail—would enable targeted regulation to be put in place to help improve the distribution and retail functions.

8.2.1.3 Incentive regulation of the distribution activity

The distribution activity should be regulated with incentives that encourage critical performance parameters for energy access—specifically providing physical grid connections and maintaining technical losses at an acceptable level.

The distribution activity should be remunerated based on a revenue cap—so that distributors are not incentivized to encourage excess consumption—and a methodology to encourage efficiency, such as RPI-X (Pérez-Arriaga, 2013). This should be augmented with incentives for improvements in quality of service up to an optimal level, improvements in losses up to an optimal level, and providing connections to more customers. The revenue cap and incentive parameters could be determined via the use of a reference network model (Pérez-Arriaga, 2013).

In this case, the use of a reference network model is important in order to estimate and set targets for technical losses, because without such a technical model it would be difficult to distinguish technical losses (which should be the responsibility of the distribution activity) from non-technical losses (which should be the responsibility of the retail activity). Thus, this is another area where computer models like REM can support regulation to support electricity access.

8.2.1.4 Incentives and requirements for retail

The retail activity in the context of electricity access is a complicated matter. Retail responsibilities should include contracting adequate supply to meet customers' desired level of service at minimal cost, and minimizing non-technical losses (including theft and non-collection of payment). Considering these responsibilities, it is not obvious to what extent retail competition is desirable and how it should be implemented.

On the one hand, open retail competition could allow more innovation and greater diversity with respect to the service provided to customers. Different customer classes may desire different service levels and features and be willing to pay different rates for the service. For example, wealthy households may be willing to pay more for non-interrupted service, while poorer households may be willing to accept lower service levels at lower rates.

On the other hand, retail competition creates an issue with respect to theft. If retailers can choose their customers, who will be responsible for those illegal grid users that are not customers? Additionally, retail may be a natural monopoly in some remote areas, since the default payment method is door-to-door collection and the cost of getting to the village to collect may be high compared to the potential revenue from electricity sales.

As a compromise to these considerations, I propose a system of retail competition at the community level. In general, it would make sense for "community" boundaries to be aligned with the electric grid topology, such that different service levels could be provided to each community. For example, the community could be all customers downstream of a certain substation. Each community would hold periodic (maybe every 5-10 years) auctions to select a retailer. Retailers would come to the community with various packages of service levels and tariff structures, and the community could pick the one that best serves their needs. A community board or village leader could represent the consumers in negotiations with the retailers. The chosen retailer would then have responsibility for providing adequate service to all connected customers and for managing theft by households within the community.

This method would allow for a great diversity of retail models to be implemented, in order to address the variety of desired service characteristics and the need for innovation in retail.

In order to protect consumers and reduce transaction costs of this scheme, regulators should establish standard minimum retailer requirements and templates for standard contract terms. Retailer requirements should include the obligation to procure supply competitively, rather than via cost-plus arrangements with single preferred parties. This would mitigate the issue with current DISCOMs in which they overpay for supply from politically-favored private companies or state-owned generation (Kumar & Chatterjee, 2012). Retailers should also be required to project demand and contract adequate supply for any expected growth, which has historically not been done by DISCOMs (Kumar & Chatterjee, 2012).

Customers above a certain size (certainly at least those connecting to the grid at medium or high voltage) should be given full open access to the network, meaning that they can use the network to purchase power from any supplier. Open-access serves as another mechanism of retail competition to encourage efficient and high quality service (Kumar & Chatterjee, 2012). Efforts should be made to allow nondiscriminatory network access, such as isolating transmission utilities and system operators from financial interests in incumbent retailers (Kumar & Chatterjee, 2012).

8.2.1.5 Privatization of distribution and retail

Given the dramatic changes required in distribution and retail, and the track record of state-owned companies in India acting on short-term political objectives, it seems that greater involvement of the private sector in distribution and retail could be beneficial. Additionally, the profit motive of private sector actors improve the effectiveness of regulators and their financial incentives.

In practice, there have been positive experiences with distribution and retail privatization in India, although there have certainly been mistakes to learn from.

One model—the replacement of a state-owned DISCOM with a private company—has been implemented in Delhi and Orissa. In Delhi this transition was largely successful—for example, New Delhi Power Limited dramatically reduced total technical and nontechnical losses from 53 percent in 2002 to 15 percent in 2010 and managed to reduce costs below the tariff rate (Kumar & Chatterjee, 2012). In Orissa, however, the private utility contributed to a less significant reduction in losses from 44 percent in 2005-6 to 39 percent in 2008-9 (Kumar & Chatterjee, 2012). Challenges in Orissa have been tied to lack of support from the state in the transition to privatization (Kumar & Chatterjee, 2012). These examples show that there is great potential in privatization, but the transition must be taken with care.

In another model, known as the franchisee model, a state-owned distribution company may contract out part of their responsibilities to a private company. This model also has examples of success. In one case, a distribution franchisee in Bhiwandi reduced losses from 40% to 18% and improved payment collection from 61% to 99% between 2006 and 2009 (Kumar & Chatterjee, 2012).

Given the promise of privatization in distribution and retail and the great challenges regulators have in controlling state-owned companies, regulators should be given the power to initiate privatization of elements of distribution and retail for chronically underperforming state-owned companies. It is worth noting, though, that despite the strong performance of some private utilities, the Indian public still feels significant uncertainty about the potential for improvements via utility privatization (Santhakumar, 2008).

8.2.2 Sending efficient signals to consumers

In order for utilities to provide electricity access, consumers must be charged tariffs that allow utilities to cover expenses. Additionally, tariffs should send signals to consumers that reflect the costs that they impose on the system, in order to encourage efficient use of system resources. These best practices become complicated, however, in the electricity access context. Revenue adequacy comes into conflict with the social obligation to provide service to below poverty line households or political obligations to provide cheap electricity to favored groups. Correctly allocating costs becomes complicated when sophisticated meters are not affordable and a large fraction of consumption is unmetered (legally and illegally). This section discusses some recommendations regarding how to set tariffs and send efficient signals to customers in the context of electricity access.

8.2.2.1 Revenue adequacy

It is critical that utilities are adequately remunerated to cover their costs and support required investments in the power system. The changes in distribution and retailing recommended above should help to significantly reduce costs, and thus financial losses, for utilities. But changes in the tariff structure will likely still be necessary.

With increased privatization of distribution and retail, increased strength and independence of regulators, and increased community choice of service level, appropriate tariffs should naturally result for most consumers. Retailers can decide, based on the nature of their service area, whether to offer simple tariffs or more advanced schemes. I think there would be interesting opportunities for retailers to innovate in low-cost methods for providing time-dependent pricing signals to consumers, without depending on expensive sophisticated meters for every customer.

There are still, however, some items which could lead to revenue inadequacy which must be addressed. These include customers who cannot afford the minimum socially guaranteed level of service, the policy-based subsidies to certain customer classes (especially agriculture), and theft or non-payment. These issues are addressed below.

8.2.2.2 Subsidies for the poor

To the extent that electricity is seen as a basic social right, or a critical enabler of development, the government may choose to establish a basic universal service guarantee.

No matter the level of basic service, there will be some customers who cannot afford to pay for the full cost of this level of service. The total quantity of this revenue gap will depend on the level of service. For a very minimal level of service, for example power for a few lights and a phone charger for 4-6 hours per day, most customers should be able to pay. Experience has shown that much of the rural poor are already paying for lighting (often from candles or kerosene) and phone charging (from local entrepreneurs), and rates they pay are far above the costs of grid electricity. If a higher level of service was guaranteed—for example service deemed sufficient for productive uses to spur development—the induced revenue gap would likely be larger.

In either case, some mechanism must be used to allow that this service is provided and to address the revenue gap. One strategy is to charge the poor lower tariffs, and recoup the missing revenue via a government subsidy to utilities. A second is to charge the poor lower tariffs, and recoup the missing revenue via a cross-subsidy, where the utility charges another class of customers more than their costs. A third strategy is to charge the poor tariffs that reflect the cost of service, and provide a direct subsidy to the consumers to allow them to afford the minimum guaranteed level of consumption.

Historically, India has favored a combination of the first two strategies. They charge poor consumers lower rates, and attempt to make up for the revenue deficit via direct government subsidy and a cross-subsidy in which industrial and commercial consumers are charged higher rates (Kumar & Chatterjee, 2012). This has led to several issues. Since the poor do not see the true marginal cost of their consumption, they are encouraged to waste energy and not invest in higher-efficiency appliances with higher upfront cost. This increases the amount of missing revenue that must be found by utilities. When commercial and industrial customers are charged rates that are higher than costs, many choose to find an alternate supplier or depend on on-site generation. This effect combined with the first, leads to increasing government subsidy requirements in order to cover the costs of utility operation. Eventually the governments cannot afford these subsidies, utilities cannot afford to invest, and system quality deteriorates. The result is that we aspired to give the poor affordable electricity access, but instead gave them almost none.

In response to these experiences, and basic economic principles, the third strategy is recommended, in which poor consumers still see the appropriate cost signals, but receive a direct subsidy to support basic consumption. This strategy would likely also have some complications. For example, the process of distributing the subsidy to consumers could be challenging and consumers might use the subsidy for things besides electricity. One potential approach to address these concerns would be to allow retailers to manage the subsidy for consumers, to simplify logistics and tie the subsidy directly to electricity use. In this case, appropriate protections would need to be put in place prevent abuse by the retailers.

8.2.2.3 Subsidies to agriculture

Subsidies to agriculture have been a prominent policy feature in India. In part, these subsides have come through dramatically reduced electricity rates, resulting in essentially free and unmetered electricity for farmers (Kumar & Chatterjee, 2012). Since electricity is free and unmetered, there are incentives for farmers to over-consume electricity, which makes the required cost of subsidy unaffordable. A few measures should be put in place to make any preferential electricity rates to farmers affordable.

First, subsidized farming consumption should be metered where possible. Currently, consumption of farmers are estimated based on the horsepower of their allowed irrigation pumps (Kumar & Chatterjee, 2012). But there are strong incentives for farmers to attach additional illegal pumps or use the power supply for non-approved purposes. Since there is no metering, and technical and non-technical losses on the network are so high, it is difficult to know how much extra subsidized electricity is being consumed and by whom.

Second, regulators should implement a similar tariff scheme to the one recommended for the poor, in which farmers see the full cost of their consumption in the tariff and receive a separate subsidy. This will reduce the risk that this subsidy becomes unaffordable and will allow farmers to respond to efficient economic signals.

Finally, with more metering and stricter enforcement of subsidized consumption limits, farmers might complain that they can no longer afford to irrigate their land. This is likely related to the use of inefficient pumps and inefficient flood-irrigation techniques. Regulators should encourage energy service companies to facilitate installation of energy and water efficient irrigation systems in order to yield utility bill savings. Metering and stricter enforcement should be rolled in gradually after a warning period, to provide farmers and energy service companies time to make required investments.

8.2.2.4 Theft and non-payment

So-called “non-technical losses” including theft and non-payment are a serious barriers to the financial viability of utilities in India. Hence, there are also barriers to utilities being able to provide universal access to electricity. Some measures to reduce non-technical losses have proven effective and should be replicated.

The first, and probably best strategy, to reduce non-technical losses is to provide acceptable quality of service. If customers are pleased with their service, they will value it and be more likely to be willing to pay.

Secondly, community and social pressures can be leveraged to prevent theft and encourage payment. Both positive and negative pressure are potential tools. As an example of positive social pressure, Tata Power Delhi Distribution Limited has actively engaged in community development and education, in parallel with improvements in service, in order to encourage residents of urban slums to pay for electricity service. As an example of negative social pressure, Maharashtra has tied load shedding for various regions to the level of non-technical losses (Kumar & Chatterjee, 2012). In this way, reduced theft is rewarded with better service, and vice versa.

Finally, the strong profit motives that come with privatization seem to naturally lead to reduced theft and increased payment collection. While state-owned utilities may prefer to accept losses rather than facing public pressure against cracking down on payment collection, private for-profit entities do not have this level of mixed incentives. Some examples of dramatic reductions in losses through privatization were mentioned above.

8.3 RECOMMENDATIONS FOR OFF-GRID REGULATION

Appropriate regulation for areas that can be well-served by the main grid will go a long way towards enabling universal electricity access. But it will take some time for the grid to reach everywhere it is currently the least-cost electrification mode, and many places would currently be more efficiently served by off-grid systems.

Off-grid electricity systems fall on a spectrum, including devices, single home systems, and microgrids. Along this spectrum, different business models are possible, including system sale, system leasing, and providing electricity as a service. It is likely that a single regulatory approach will not be appropriate for all systems and business models.

Additionally, off-grid electrification activities will need to be coordinated with on-grid activities, in order to ensure that both activities can be viable and to minimize wasteful duplication of effort.

Finally, regulation of off-grid electricity systems should pay particular attention to broader considerations.

Recommendations regarding these issues are discussed in the rest of this section.

8.3.1 General principles for regulation of off-grid systems

Reiche, et al. outline two "golden rules" of electrification that should be carefully considered when crafting regulation for off-grid electrification:

> "Rule 1—Regulation is a means to an end. What ultimately matters are the outcomes (such as sustainable electrification)—not regulatory rules.
>
> Rule 2—The benefits of regulation must exceed the costs of regulation." (Reiche, 2006).

While common regulatory paradigms have been developed for centralized electric grids, it is crucial to not attempt to blindly apply those to off-grid systems without consideration of these golden rules.

The conditions that make it challenging to provide centralized electricity service in certain areas—especially remoteness and poverty—also make it difficult to implement enforceable and affordable regulation. Carelessly implemented requirements would be ignored or make it so that business is impossible. Off-grid systems have different cost structures and enable different business models, which may not fit well under standard remuneration and tariff-setting procedures. Based on these considerations, regulation of many off-grid systems ought to be significantly different in nature and less in quantity compared to regulation of the main electric grid.

However, there are still strong justifications for why some regulation is appropriate. Some forms of regulation can protect consumers from systems that are of poor quality or unsafe and protect them from companies exerting monopoly power. Regulation can also help to reduce business risk for companies, so that providing electricity access becomes a profitable business in more places. Regulation can encourage conformity to standards that enable interoperability and expandability, so that investments in initial off-grid systems do not all become stranded assets as demand changes and the main grid reaches the area.

Thus a balanced approach is required, where simple and enforceable regulations are implemented in the cases where they are justified.

8.3.2 Current situation in India

Currently, India has taken a mostly hands-off approach to the regulation of off-grid power systems. The Electricity Act, 2003 allowed for the distribution of electricity in rural areas without a license. This has allowed for many small companies and non-governmental organizations to engage in the provision of electricity through off-grid systems, without government involvement. These groups have implemented systems with a wide variety of technical and business models. In parallel, several government agencies—including the Ministry of Power, the Ministry of Rural Development, the Ministry of New and Renewable Energy, and state agencies—are supporting additional off-grid electrification projects.

While this hands-off regulatory approach has probably encouraged more activity, it has also created some confusion. It is not clear if anyone is tracking all of these projects, how many of them meet the government's expectations for basic service, and who is responsible for those who have not been served (Bhattacharyya & Palit, 2013). Additionally, there is significant confusion regarding the relationship of these off-grid projects to grid extension activities—including the technical, legal, and financial conditions for eventual connection of off-grid systems to the main grid.

8.3.3 A recommended approach for off-grid electrification planning

The initial hands-off approach of the Indian government has facilitated many off-grid electrification projects, with at least hundreds of thousands of Indians now receiving some electricity service through off-grid systems. This has been great, in that it has allowed significant learning to be done about which technical and business models are feasible and it has allowed for some capacity building in off-grid electrification. On the other hand, though, the present impact of off-grid systems is just a drop in the bucket compared to the hundreds of millions of Indians who still lack electricity access. A more hands-on may be appropriate to transform what has been learned into a process with the ability to achieve universal access.

I recommend that off-grid electrification should be planned at a high-level in coordination with the process of planning on-grid electrification. Since the government believes that basic electricity service should be guaranteed to all Indians, all Indians should be included in the plan. A proposal for a basic outline of a holistic electrification plan follows:

1. Nationally or state-by-state, acceptable minimum standards of electricity access should be established. These standards should consider both the provision of basic services and the level of access which enables economic development through the productive use of electricity.
2. Each state should identify all areas which have adequate electricity service from the grid. These areas have electricity access, and do not need to be addressed by the electrification plan.[37]
3. Each state should identify all areas which are grid connected, but do not have acceptable service. The on-grid electrification reforms, like those described in section 2 of this chapter, should be implemented so that service quality can be improved to an acceptable level. If acceptable access will not be provided rapidly, there is a business opportunity for single-user systems to provide supplemental basic service.

[37] This assumes that the quality of grid service will not deteriorate in these areas. Smart planning, investment, and operation is required to ensure that deterioration of service does not happen.

4. Each state should identify areas that are not currently grid connected, but could be served at least cost by grid connection. The state should establish an order of priority and timeline for connection of these areas.
    a. For areas which will receive grid service soon, single user systems can provide basic service in the near term. These areas should fall under the universal service obligation of a distribution company.
    b. For areas which will not receive grid-service soon, licenses should be awarded for implementation and operation of grid-compatible microgrids which will be connected to the grid eventually. The license should give the microgrid company a universal service obligation for the area. Arrangements should be made for the microgrid company to eventually sell the microgrid assets to a DISCOM or to operate as a distribution franchisee or licensee.
5. Each state should identify the remaining areas, which do not currently have grid service and would be served at least cost by off-grid systems. A certain agency (perhaps the Ministry of Rural Development) should be assigned responsibility for tracking which households in these areas have off-grid systems that provide an acceptable level of electricity access. It should also have the responsibility of facilitating businesses to provide off-grid systems to those households that have not been reached, including providing information about which technical and business models are most likely to be successful. The Ministry of Rural Development would be well suited to be a lead agency for off-grid electrification, since social concerns rather than technical ones tend to be primary in this space (Bhattacharyya & Palit, 2013).
    a. Collective systems, such as microgrids should be managed by light-handed regulation. This could consist of, for example, an obligation to file some basic information about the system, standard service provider obligations, and standard consumer protection mechanisms (Bhattacharyya, 2013). The terms of this light-handed regulation should depend on the nature of the entity providing electricity service (Reiche, 2006). For example, greater protection against monopoly abuse is required for a private entity compared to a coop.
    b. Incentives should be established that encourage investments in grid-compatible systems in the regions that are closest to being ready for grid connections. These incentives should target covering the incremental capital cost of the grid-compatible infrastructure, in cases where operating revenues will cover operating costs, so that the system is not dependent on perpetual subsidy (Bhattacharyya & Palit, 2013). The Ministry of Power and traditional electricity regulatory agencies should have primary responsibility over these systems which are slated for eventual connection to the grid.
    c. Companies that can profitably provide individual systems should not require significant economic regulation, but other consumer protections may still be required (Bhattacharyya & Palit, 2013). The government should facilitate establishment or awareness of product safety and quality standards to provide consumer protection.
    d. For customers that cannot afford basic off-grid service, the government should have subsidy programs to cover a percentage of initial capital costs, but avoid implementing recurring subsidy of operating costs. In exchange for this subsidy, the system providers (of collective or individual systems) should fall under a light-handed regulation where basic service responsibilities are outlined, including an obligation to provide service for a

certain period of time. An approach similar to this was implemented by Bolivia for solar home systems (Reiche, 2006). This program should be managed by the Ministry of Rural Development, which can act as an intermediary between system providers and communities.

Computer models such as REM can support this plan in two important ways. First, they can help planners identify which areas are best suited for grid extension now or in the near term, and which areas are best served with off-grid systems. Second, they can help to estimated costs of these systems in order to support setting appropriate tariffs and plan for required subsidies.

8.3.4 Capacity Building

This new regulatory approach would require significant investment of financial and human capital into an ecosystem that can support universal electricity access. Especially important considerations include:

- Improved capacity of regulatory bodies to track a large number of decentralized projects
- Improved access to financing for off-grid systems
- Training and education for designers, builders, and servicers of off-grid systems
- Training and education of villagers to encourage efficient use of energy and the productive use of energy to encourage economic development.

8.4 SUMMARY OF REGULATORY REFORM RECOMMENDATIONS

In order to achieve a meaningful level of universal electricity access in India, a holistic regulatory approach is required.

First, regulators should support regional electrification planning including the designation of which areas should fall under the traditional on-grid regulation versus a new off-grid regulation. This planning should ideally use technical models such as REM to find least-cost electrification plans, but also take into account social and practical considerations.

Second, in order enable true access through the on-grid electrification mode, significant reforms to India's on-grid electricity sector are required. These reforms are mostly related to the insulation of power-system decisions from myopic and short-term political motivations.

Finally, a significant new regulatory capacity must be developed to manage a light-handed regulatory approach for off-grid electricity access. This regulatory capacity would be best managed by an organization like the Ministry of Rural Development, which has experience working with isolated rural communities.

9 CONCLUSION

In this thesis I have described a computer model for the planning of electricity access systems called the Reference Electrification Model (REM), which was developed by me with the support of other members of MIT's Universal Electricity Access Research Group. The application of REM to the district of Vaishali in Bihar, India is described, in order demonstrate the feasibility of using computer models to support large-scale electrification planning. The thesis ends with a discussion of recommended regulatory reforms to support electricity access in India, including considerations regarding how models like REM can support effective regulation.

Overall, I have shown that even in its current premature state, REM is a powerful and useful tool which can be practically applied to a large region in a developing country. However, the data collection process is challenging and the model requires an expert user and careful interpretation of results. Also, given the current regulatory approaches in India, some of the potential values of REM might not materialize.

With some additional work to reduce the burden of data gathering and improve the consistency of results, REM could be made accessible to many organizations involved in planning or studying electricity access and have a major impact. With or without REM, regulatory reform is essential in order to support effective and affordable provision of universal electricity access. Regulatory reform that takes advantage of advanced planning tools such as REM may have the best prospect of harmonizing activities between the diverse set of technologies and business models in the electricity access space.

10 REFERENCES

Agalgaonkar, A. P., Dobariya, C. V., Kanabar, M. G., Khaparde, S. A., & Kulkarni, S. V. (2006). *Optimal sizing of distributed generators in MicroGrid*.

Arefifar, S., Mohamed, Y., & El-Fouly, T. (2013). Optimum Microgrid Design for Enhancing Reliability and Supply-Security. *IEEE TRANSACTIONS ON SMART GRID*, *4*(3), 1567–1575.

Bahramirad, S., Reder, W., & Khodaei, A. (2012). Reliability-Constrained Optimal Sizing of Energy Storage System in a Microgrid. *Ieee Transactions on Smart Grid*, *3*(4), 2056–2062. http://doi.org/10.1109/TSG.2012.2217991

Bailey, O., Creighton, C., Firestone, R., Marnay, C., & Stadler, M. (2003). Distributed energy resources in practice: A case study analysis and validation of LBNL's customer adoption model. *Lawrence Berkeley National Laboratory*. Retrieved from http://escholarship.org/uc/item/1dp0j5z6

Bhattacharyya, S. C. (2013). To regulate or not to regulate off-grid electricity access in developing countries. *Energy Policy*, *63*, 494–503. http://doi.org/10.1016/j.enpol.2013.08.028

Bhattacharyya, S. C., & Palit, D. (2013). *Rural Electrification Through Decentralised Off-grid Systems in Developing Countries*. (S. Bhattacharyya, Ed.)*Green Energy and Technology* (Vol. 116). Springer London. Retrieved from http://www.scopus.com/inward/record.url?eid=2-s2.0-84883183826&partnerID=tZOtx3y1

Bornstein, D. (2007). *How to change the world: Social entrepreneurs and the power of new ideas*. Oxford University Press. Retrieved from http://books.google.com/books?hl=en&lr=&id=P_g8gVyuuEgC&oi=fnd&pg=PR7&dq=how+to+change+the+world,+by+David+Bornstein,+Oxford+University+Press&ots=BCO_WxvSCC&sig=r4ug7nTxapVYR9ScMQXSA9N9SZo

Chakrabarti, S., & Chakrabarti, S. (2002). Rural electrification programme with solar energy in remote region–a case study in an island. *Energy Policy*, *30*(1), 33–42. http://doi.org/10.1016/S0301-4215(01)00057-X

Chaurey, A., & Kandpal, T. C. (2010). A techno-economic comparison of rural electrification based on solar home systems and PV microgrids. *Energy Policy*, *38*(6), 3118–3129. http://doi.org/10.1016/j.enpol.2010.01.052

Electric Power Research Institute. (2003). *Costs of Utility Distributed Generators, 1-10 MW*.

Erol-Kantarci, M., Kantarci, B., & Mouftah, H. T. (2011). Cost-Aware Smart Microgrid Network Design for a Sustainable Smart Grid. *2011 Ieee Globecom Workshops (gc Wkshps)*, 1178–1182.

Fang, L., Cai, J., Lin, H., & James, G. (2012). Economic Analysis of Two Microgrid Prototypes. In C. S. Zhang (Ed.), (Vol. 433–440, pp. 2411–2416).

Gonzalez-Sotres, L., Mateo Domingo, C., Sanchez-Miralles, A., & Alvar Miro, M. (2013). Large-Scale MV/LV Transformer Substation Planning Considering Network Costs and Flexible Area Decomposition. *IEEE Transactions on Power Delivery*, *28*(4), 2245–2253. http://doi.org/10.1109/TPWRD.2013.2258944

Hafez, O., & Bhattacharya, K. (2012). Optimal planning and design of a renewable energy based supply system for microgrids. *Renewable Energy*, *45*, 7–15. http://doi.org/10.1016/j.renene.2012.01.087

Hooke, R., & Jeeves, T. A. (1961). " Direct Search" Solution of Numerical and Statistical Problems. *J. ACM*, *8*(2), 212–229. http://doi.org/10.1145/321062.321069

International Energy Agency. (2013). World Energy Outlook 2013. Retrieved from http://www.deltalinqsenergyforum.nl/documents/2014/IEA presentatie World Energy Outlook 2013-2035.pdf

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

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

Kumar, A., & Chatterjee, S. K. (2012). *Electricity Sector in India: Policy and Regulation*. Oxford University Press.

Lambert, T., Gilman, P., & Lilienthal, P. (2006). Micropower System Modeling with HOMER. In F. A. Farret & M. G. Simoes (Eds.), *Integration of Alternative Sources of Energy* (pp. 379–418). John Wiley & Sons, Inc.

Lambert, T. W., & Hittle, D. C. (2000). Optimization of autonomous village electrification systems by simulated annealing. *Solar Energy*, *68*(1), 121–132. http://doi.org/10.1016/S0038-092X(99)00040-7

Liu, Y., Liu, Y., Du, Y., Ruiqi, W., & Huang, W. (2013). Optimal Allocation of Distributed Generation in Micro-grid Based on the Theory of Life Cycle Cost. In M. Sun, G. Yan, & Y. Zhang (Eds.), (Vol. 614–615, pp. 1903–1907).

Luna-Rubio, R., Trejo-Perea, M., Vargas-Vázquez, D., & Ríos-Moreno, G. J. (2012). Optimal sizing of renewable hybrids energy systems: A review of methodologies. *Solar Energy*, *86*(4), 1077–1088. http://doi.org/10.1016/j.solener.2011.10.016

Manwell, J. F., & McGowan, J. G. (1993). Lead acid battery storage model for hybrid energy systems. *Solar Energy*, *50*(5), 399–405. http://doi.org/10.1016/0038-092X(93)90060-2

Mateo Domingo, C., Gomez San Roman, T., Sanchez-Miralles, Á., Peco Gonzalez, J. P., & Candela Martinez, A. (2011). A Reference Network Model for Large-Scale Distribution Planning With Automatic Street Map Generation. *IEEE Transactions on Power Systems*, *26*(1), 190–197. http://doi.org/10.1109/TPWRS.2010.2052077

Miguez, E., Cidras, J., Diaz-Dorado, E., & Garcia-Dornelas, J. L. (2002). An improved branch-exchange algorithm for large-scale distribution network planning. *IEEE Transactions on Power Systems*, *17*(4), 931–936. http://doi.org/10.1109/TPWRS.2002.804998

Millinger, M., Marlind, T., & Ahlgren, E. O. (2012). Evaluation of Indian rural solar electrification: A case study in Chhattisgarh. *Energy for Sustainable Development*, *16*(4), 486–492. http://doi.org/10.1016/j.esd.2012.08.005

Mohamed, A., & Khatib, T. (2013). *Optimal Sizing of a PV/Wind/Diesel Hybrid Energy System for Malaysia*.

Moreira, J. C., Miguez, E., Vilacha, C., & Otero, A. F. (2012). Large-Scale Network Layout Optimization for Radial Distribution Networks by Parallel Computing: Implementation and Numerical Results. *IEEE Transactions on Power Delivery*, *27*(3), 1468–1476. http://doi.org/10.1109/TPWRD.2012.2190305

Oda, H., & Tsujita, Y. (2011). The determinants of rural electrification: The case of Bihar, India. *Energy Policy*, *39*(6), 3086–3095. http://doi.org/10.1016/j.enpol.2011.02.014

Peco González, J. P. (2001). *Modelo De Cobertura Geográfica de una Red de Distribución de Energía Eléctrica*. Universidad Pontificia Comillas.

Pérez-Arriaga, I. J. (2013). *Regulation of the Power Sector*. (I. J. Pérez-Arriaga, Ed.). London: Springer London. http://doi.org/10.1007/978-1-4471-5034-3

Reiche, K. (2006). Electrification and regulation: principles and a model law. *Energy and Mining …*. Retrieved from http://www-wds.worldbank.org/external/default/WDSContentServer/WDSP/IB/2007/01/04/000310607_20070104171918/Rendered/PDF/383280EMSBDP0EnergyPaper18.pdf

Rout, A., & Parida, M. (2013). Design and Analysis of SPV-Diesel Hybrid System for Rural Electrification in Odisha. *International Journal of Scientific & Engineering Research*, *4*(12). Retrieved from http://www.ijser.org/researchpaper\Design-and-Analysis-of-SPV-Diesel-Hybrid-System-for-Rural-Electrification-in-Odisha.pdf

Saito, N., Niimura, T., Koyanagi, K., & Yokoyama, R. (2009). *Trade-off Analysis of Autonomous Microgrid Sizing with PV, Diesel, and Battery Storage*.

Santhakumar, V. (2008). *Analysing Social Opposition to Reforms: The Electricity Sector in India*. SAGE Publications.

Sen, R., & Bhattacharyya, S. C. (2014). Off-Grid Electricity Generation with Renewable Energy Technologies in India: An Application of HOMER. *Renewable Energy*, *62*, 388–398. http://doi.org/10.1016/j.renene.2013.07.028

Setiawan, A. A., Zhao, Y., & Nayar, C. V. (2009). Design, economic analysis and environmental considerations of mini-grid hybrid power system with reverse osmosis desalination plant for remote areas. *Renewable Energy*, *34*(2), 374–383. http://doi.org/10.1016/j.renene.2008.05.014

Su, W., Yuan, Z., & Chow, M.-Y. (2010). Microgrid Planning and Operation: Solar Energy and Wind Energy. *Ieee Power and Energy Society General Meeting 2010*.

Tafreshi, S. M. M., Zamani, H. A., Ezzati, S. M., & Vahedi, H. (2011). Optimal Unit Sizing of Distributed Energy Resources in MicroGrid Based on Mixed-Integer Bacterial Foraging Algorithm. *International Review of Electrical Engineering-Iree*, *6*(3), 1297–1307.

Tanrioven, M. (2005). Reliability and cost-benefits of adding alternate power sources to an independent micro-grid community. *Journal of Power Sources*, *150*, 136–149. http://doi.org/10.1016/j.jpowsour.2005.02.071

Whitefoot, J. W., Mechtenberg, A. R., Peters, D. L., & Papalambros, P. Y. (2012). *Optimal Component Sizing and Forward-Looking Dispatch of an Electrical Microgrid for Energy Storage Planning*.

World Bank, & International Energy Agency. (2014). *Sustainable Energy for All 2013-2014: Global Tracking Framework*. World Bank Publications. Retrieved from http://books.google.com/books?hl=en&lr=&id=85zN24_pi7UC&oi=fnd&pg=PR5&dq="work+is+available+under+the+Creative+Commons+Attribution+3.0+Unported+license+(CC+BY+3.0)"+"Under+the+Creative+Commons+Attribution+license,+you+are+free+to+copy,+distribute,+transmit,+and"+&ots=A0EoMZc7G2&sig=bmI-PqOb2GKb3JzOA6jMokeGSTM

Zamani, H. A., Ezzati, S. M., Farashah, M. D., Dahri, E., & Tafreshi, S. M. M. (2012). Linear Programming for Optimal Sizing of DGs in MicroGrid Considering Loss of Power Supply Probability Technology. *International Review of Electrical Engineering-Iree*, *7*(1), 3470–3477.

Zhu, L., & Yang, X. (2012). Design and Simulation for Microgrid System based on Homer Software. In Q. J. Xu, H. H. Ge, & J. X. Zhang (Eds.), *Natural Resources and Sustainable Development, Pts 1-3* (Vol. 361–363, pp. 1874–1877).

Zimmerman, R. D., Murillo-Sánchez, C. E., & Thomas, R. J. (2011). MATPOWER: Steady-State Operations, Planning, and Analysis Tools for Power Systems Research and Education. *IEEE Transactions on Power Systems*, *26*(1), 12–19. http://doi.org/10.1109/TPWRS.2010.2051168