Essays on Infrastructure Design and Planning for Clean Energy Systems

Ayse Selin Kocaman

Submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2014

ABSTRACT

Essays on Infrastructure Design and Planning for Clean Energy Systems

Ayse Selin Kocaman

The International Energy Agency estimates that the number of people who do not have access to electricity is nearly 1.3 billion and a billion more have only unreliable and intermittent supply. Moreover, current supply for electricity generation mostly relies on fossil fuels, which are finite and one of the greatest threats to the environment. Rising population growth rates, depleting fuel sources, environmental issues and economic developments have increased the need for mathematical optimization to provide a formal framework that enables systematic and clear decision-making in energy operations. This thesis through its methodologies and algorithms enable tools for energy generation, transmission and distribution system design and help policy makers make cost assessments in energy infrastructure planning rapidly and accurately.

In Chapter 2, we focus on local-level power distribution systems planning for rural electrification using techniques from combinatorial optimization. We describe a heuristic algorithm that provides a quick solution for the partial electrification problem where the distribution network can only connect a pre-specified number of households with low voltage lines. The algorithm demonstrates the effect of household settlement patterns on the electrification cost. We also describe the first heuristic algorithm that selects the locations and service areas of transformers without requiring candidate solutions and simultaneously builds a two-level grid network in a green-field setting. The algorithms are applied to real world rural settings in Africa, where household locations digitized from satellite imagery are prescribed.

In Chapter 3 and 4, we focus on power generation and transmission using clean energy sources. Here, we imagine a country in the future where hydro and solar are the dominant sources and fossil fuels are only available in minimal form. We discuss the problem of modeling hydro and solar energy production and allocation, including long-term investments and storage, capturing the stochastic nature of hourly supply and demand data. We mathematically model two hybrid energy generation and allocation systems where time variability of energy sources and demand is balanced using the water stored in the reservoirs. In Chapter 3, we use conventional hydro power stations (incoming stream flows are stored in large dams and water release is deferred until it is needed) and in Chapter 4, we use pumped hydro stations (water is pumped from lower reservoir to upper reservoir during periods of low demand to be released for generation when demand is high). Aim of the models is to determine optimal sizing of infrastructure needed to match demand and supply in a most reliable and cost effective way.

An innovative contribution of this work is the establishment of a new perspective to energy modeling by including fine-grained sources of uncertainty such as stream flow and solar radiations in hourly level as well as spatial location of supply and demand and transmission network in national level. In addition, we compare the conventional and the pumped hydro power systems in terms of reliability and cost efficiency and quantitatively show the improvement provided by including pumped hydro storage. The model will be presented with a case study of India and helps to answer whether solar energy in addition to hydro power potential in Himalaya Mountains would be enough to meet growing electricity demand if fossil fuels could be almost completely phased out from electricity generation.

Table of Contents

Chapter 3

Chapter 4

Chapter 5

List of Tables

List of Figures

Figure 3.9 | Summarized results for *one-basin, one-demand point* case studies. Bhagirathi has 13.3 km^3 annual stream flow and Delhi has 7500 GWh annual demand. Chenab has 52 km^3 annual stream flow and Punjab has a 12000 GWh (quarter of actual demand) computed annual demand.

Figure 3.10 | Summarized results for *two basins-two demand points* case**.** Combined system generates 2% more hydro energy with much smaller reservoirs (78% and 58% decrease in reservoir size respectively). Solar energy generation increases almost 6% and the improvements in hydro and solar energy generation provide a 10% reduction in diesel usage. There is 13% more transmission cost to build extra transmission line but overall cost is 5% lower.

Figure 4.1 | A schematic illustration for hybrid system with pumped hydro storage. There are two-level of reservoirs and water can be pumped from lower reservoir to upper reservoir during periods of low demand to be released for generation when demand is high.

Figure 4.2 | Autocorrelation function for the stream flow data of Bhagirathi River (1970-2004). Time series of climate and meteorological variables are frequently serially dependent, reflecting the strong persistence of the meteorological and climate phenomena. For example, a rainfall effect can be seen in multiple consecutive periods.

Figure 4.3 | Decomposition of time series data. The seasonal fluctuations are roughly constant in size over time and do not seem to depend on the level of the time series. **a,** The components of the additive model for stream flow. **b,** The components of the additive model for solar radiation.

Figure 4.4 | Results for the first step of the heuristic algorithm (each scenario is run individually). **a**, Only the upper reservoir size varies significantly for each individual scenario with 0.62 coefficient of variation (CV) . **b-d** the other components of the power system especially solar panel area, lower reservoir size do and generator capacity not vary a lot between scenarios (Coefficient of variations: 0.08, 0.11 ,0.10 respectively) .

Figure 4.5 | Superimposed results for 3hourly and daily run for a scenario**.** Once the "solar energy related" components of the model are estimated, the resolution of our problem can be changed (one day or one week) to estimate the size of the upper reservoir. **a**, 3hourly and daily individual run of year 1974. **b**, 3hourly and daily run of year 1981.

Figure 4.6 | The data for each year is used to run the model fixing the investment decision variables to the values in the optimal solution to observe how the diesel usage changes over different years. Red dots are the years chosen as scenarios and we observe that the unit cost of the system varies between the bounds determined by the years chosen as scenarios. This result supports the fact that scenarios are chosen to represent extreme event years such as most variables, most dry, wettest etc.

Figure 4.7 | Input data used in the model **-** Data used in Chapter 3 for illustration of single basin, single demand point case is also used here in order to be able to show the benefits of pumped hydro model.

Figure 4.8 | Water stored in the upper and lower reservoirs – The upper reservoir is assumed to be full at the start and the end of the cycle as the planning horizon starts in September which is close to end of the Monsoon season. There is no water stored in the lower reservoir in the Monsoon season since there is no need to pump water to upper reservoir. (0.01 km3 ~ 2.4 GWh)

Figure 4.9 | Flows from/to the upper reservoir level - Water is stored in dry season (November to March) and is utilized starting from March. As there is high solar radiation in spring-summer period that can be used to pump more water, lower reservoir utilization is increased until Monsoon seasons starts again. (0.01 km^3 ~ 2.4 GWh)

Figure 4.10 | One week operation balance in reservoir and demand points – **a,** Solar energy is enough to meet the demand during the day and hydro becomes effective at night. Deficit is zero during all week. **b,** Due to the inflow observed in the first day of the week, extra solar energy generated in the demand point is not necessary to pump to upper reservoir in the first two days. No spill is observed during all week. (0.01 km^3~ 2.4 GWh)

Figure 4.11 | Alternative Technologies - Compared to other hybrid systems, the system with pumped hydro storage is the most cost efficient design with much lower unit cost. Having two-level reservoir with a pump system reduces the intermittency effect of renewable sources and the system can utilize more clean energy. In particular, total hydro production increases with even smaller reservoir size compared to conventional system. Shaded area in the solar power bar is transferred to hydro power.6% is lost due to generator and pump efficiencies.

Figure 4.12 | Comparisons of solar production profile of one day in November for conventional **(a)** and pumped hydro systems **(b) -** Total production is scaled up by increased the solar panel area as the solar radiation curve is same. The role of solar energy in pumped system is twofold: internal solar and pumped solar. Solar energy used internally for the 4^{th} and 5^{th} time periods are the same between two systems as extra solar energy is spent for pumping, however shaded area in 3^{rd} and 6^{th} time periods in (**b)** represents extra solar internal.

Figure 4.13 | Comparisons of internally used solar energy for one week in November. The area under curves explains the difference between solar proportions to meet the demand in Table 4.7.

Figure 4.14 | Multi-basins, Multi-demand Points Cases – a, is the system discussed in Section 5.5.1. b, introduces an additional basin (Pinder River with 6.2 km3 annual inflow) to the system and reduces the unit cost. c, Higher flexibility from the variety of resources is observed when third basin (Chenab River with 45.5 km^3) is included. d, Another demand point (Punjab -

constraint, system increases the size of the upper reservoir. Up to 4%, system increases the size of solar panels, upper and lower reservoirs; however after 4% system only increases the upper reservoir size. **b**, Duel price of the constraint also presents that increasing diesel by 1 GWh reduces the objective function by $140000 between 0% and 4%. The effect of relaxing the constraint decreases around optimal solution (8%).

Figure 5.1 | Match between proposed network and existing grid. Proposed transformers, low voltage (LV) and medium voltage (MV) networks compared to partial existing grid. Algorithm outputs 90 transformers for 1175 demand points.

Figure 5.2 | Network Sensitivity Analysis. a, The sudden change in the number of transformers as the ratio between C_{MV} (unit cost for medium voltage line) and C_{LV} (unit cost for low voltage line) cost parameters increases. **b,** Change in total MV and LV lengths from the 1000 transformer case. **c,** The change in the number of transformers for different *q* ratios as *p* ratio increases.(See Section 5.1.2 for q). **d**, The difference between the curves in b with MV weighted with different *p* ratios.

Figure 5.3 | Basins and demand points determined in India for analysis. Data is collected from CEA (Central Electricity Authority, Power Ministry of India) and other official websites to accurately estimate the 3-hourly demand load profile of each state for one year. If there is missing data for some days or hours within a day, interpolation/extrapolation methods are performed for projection.

Figure 5.4 | Comparison of alternative technologies. Solely diesel is the most expensive. As the use and the variety of clean energy increase, the unit cost goes down substantially. It is more cost efficient to design solar panel area based on high demand and spill some of the renewable energy in low demand periods.

Figure 5.5 | Alternative Technologies - Compared to other hybrid systems, the system with pumped hydro storage is the most cost efficient design with much lower unit cost. Having two-level reservoir with a pump system reduces the intermittency effect of renewable sources and the system can utilize more clean energy. In particular, total hydro production increases with even smaller reservoir size compared to conventional system.

Figure 5.6 | Comparisons of solar production profile of one day in November for conventional **(a)** and pumped hydro systems **(b) -** Total production is scaled up by increased the solar panel area as the solar radiation curve is same. The role of solar energy in pumped system is twofold: internal solar and pumped solar. Solar energy used internally for the 4th and 5th time periods are the same between two systems as extra solar energy is spent for pumping, however shaded area in 3rd and 6th time periods in (**b)** represents extra solar internal.

Figure 5.7 | Pinder basin is added to the system to show that geographic aggregation of alternative sources can reduce the variability of renewable sources.

Figure 5.8 | Spectral distribution of stream flow and demand for Bhagirathi River and Delhi case study. **(a)** The spectrum for 51 years of 3-hour resolution data shows the yearly fluctuations. **(b)** The diurnal fluctuations observed in March is showed on the spectrum for 1 month data with 3-hour resolution. **(c)** Spectral distribution of demand is presented for 1 year data of 3-hour resolution and therefore seasonal fluctuations can not be observed in the figure.

Figure 5.9 | Spectral distribution of intermittent sources. **a,** The spectrum of solar data used in base case studies (Delhi) with 1-hour resolution. **b,** The spectrum of wind data used as an example and collected from NREL website.

Figure 5.10 | Energy generation distribution comparisons for solar and wind. **a,** Proportion of wasted, internally used and pumped energy as a percentage of demand. **b,** Proportion of hydro energy generated by natural stream flow or pumped water.

Acknowledgment

First of all, I owe my greatest gratitude to my advisor, Professor Vijay Modi, whose guidance and support in this work have been invaluable to me. I thank him for having been always patient, understanding and supportive during the challenges in my academic and personal life throughout the years. I always felt fortunate to be a part of his research group. I am grateful him for helping me and teaching me so much in sustainable energy systems.

I would also like to thank Prof. Tim Huh for giving me a chance at the beginning of my graduate study and that was the time when my journey started. His wisdom and encouragement at critical times made very important steps possible for me.

Besides my advisors, I would also like to thank the members of my PhD thesis committee -Prof. Upmanu Lall, Prof. Tuncel Yegulap and Prof. Paul Duby for their valuable time and advice.

I am very lucky to work with very talented people in the Sustainable Engineering Laboratory. Susan Kum with her GIS skills made things faster and productive for me and I wish the best for her. I am thankful to Roy Han who introduced Python to me and helped so much during the learning process. Alex Zvoleff, Aly Sanoh and Carlos Abad are great researchers and it was always great to work and collaborate with them. It has been always great pleasure to interact with the other former and current members of our group, Edwin Adkins, Chris Natali, Matt Zebiak, Shaky Sherpa, Daniel Soto, Matt Berg, Jiehua Chen, Bianca Howard, Mike Waite, Brett Gleitsmann, Brittany Wright, Pam Sanchez, Jonathan Carbajal, Salah Chafik, Jonathan Demierre and many others. This work would have never been possible without their help. I am truly honored to have had this opportunity to work so closely with such brilliant researchers. Their perpetual energy and enthusiasm in research always motivated me.

I greatly enjoyed the collaboration between many groups both inside and outside of the Columbia and our group. They were certainly critical for this work to be done. Within Columbia, I want to thank Prof. Upmanu Lall, Dr. Tara Troy and members of Columbia Water Center, Prof. Iyengar, Prof. Yegulap, Prof. Wong and the people in their group and other Earth and Environmental Engineering and Industrial Engineering and Operations Research professors for their time and help along the way. I would also like to thank Prof. Olav Fosso from Norwegian University of Science and Technology and Eilyan Bitar from Cornell University for the fruitful discussions.

Finally, I would like to thank my family for their unconditional love and support. My parents deserve more credit than I can ever express in words here. I can not imagine more open-minded and supportive parents. They were always very understanding and respective to all the decisions I made. The smile on their faces is the greatest accomplishment I can achieve. Anne, baba, thank you very much for everything, I love you! My beloved brother, Serhat, I felt always safe knowing that he is always there for me. Our expanded family has been enriched much further with his dear wife, Sena and his little daughter, Berin.

Last but not least, I wish to give my deepest gratitude to my husband and my son for their endurance and sacrifice over the years. My beloved husband, Serdar, encouraged me to do my best all the time, and was always there for me with all his support and love. I could not do any of these without his support. I am more complete with him in my life since the last 12 years. There is no doubt that the reason for my greatest moments during my PhD is my son, Emre. I am thankful for the happiness that came to my life with him. He is always the reason for my smile. I would like to thank him for being so mature and I apologize for the time I took from him.

To my angel, Ece,

To my son, Emre,

To my husband, Serdar.

Chapter 1

Introduction

The International Energy Agency estimates that the number of people who do not have access to electricity is nearly 1.3 billion and a billion more have only unreliable and intermittent supply. These numbers are expected to increase unless investments in providing modern energy services are expanded significantly. Moreover, current supply for electricity generation mostly relies on fossil fuels, which are finite and one of the greatest threats to the environment. Rising population growth rates, depleting fuel sources, environmental issues and economic developments have increased the importance of sustainable energy planning and applications of Operations Research techniques in energy industry. Mathematical optimization provides a formal framework that enables systematic and clear decision making in energy operations. This thesis through its methodologies and algorithms enable tools for energy generation, transmission and distribution system design and help policy makers to make cost assessments in energy infrastructure planning rapidly and accurately.

To this end, in Chapter 2, we first focus on power distribution systems planning for rural electrification using techniques from combinatorial optimization. As governments of the developing countries attempt to increase the electricity coverage, they are seeking an understanding of the various modalities of public and private sector contributions. These modalities require a detailed understanding of the cost structure of electrification to make rapid assessment of the progress in rural electrification. With the emergence of off-grid and distributed approaches, there is a need amongst infrastructure planners to evaluate the costs of networked or grid approaches vis a vis off-grid approaches The investment costs of off-grid approaches are easier to estimate, but the investment costs of networked approaches are more difficult to estimate, taking into account both the spatial distribution of demand and the optimal placement of infrastructure to meet that demand.

For distribution systems planning problem studied in Chapter 2, we digitized QuickBird satellite imagery using remote sensing techniques to identify household-level demand points in some villages in Sub-Saharan Africa. Given the detailed local-level information from remote-sensing data, we developed new methods to estimate the cost of local-level distribution systems for a least-cost network, and to compute additional information of interest to policymakers, such as the marginal cost of connecting additional households to a grid as a function of the penetration rate while taking into account populations structure. If it is possible to connect only some portion of the households to electricity due to limited funding, how to choose that portion is a very critical question and detailed mathematical analysis provided in this chapter can help answer this type of questions.

Moreover, in Chapter 2 we present another heuristic algorithm for rural electrification projects which designs multi-level distribution system layouts while minimizing overall cost of infrastructure costs; specifically the combined costs of transformers and the two-tiered network together. To our knowledge, this algorithm is the first heuristic algorithm that selects the locations and service areas of transformers without requiring candidate solutions and simultaneously builds two-level grid network in a green-field setting. It allows one to specify different costs for the higher throughput lines upstream of the transformer as compared to downstream of the transformer This algorithm can serve as a tool for network engineers and planners to make rapid assessments assisting them with (a) estimates of total cost of distribution, (b) layouts of initial designs and (c) breakdown of total costs into transformer cost and medium and low voltage line costs and giving them a good starting point for more detailed smart grid projects. Proposed methodologies can easily be adapted to other infrastructure problems such as designing communication networks and water supply network. This discussion forms and completes Chapter 2.

In Chapter 3 and 4, we focus on power generation and transmission using clean energy sources. Current supply for electricity generation mostly relies on fossil fuels. However, fossil fuels are finite and their combustion causes global warming and health hazards. To reduce the role of fossil fuels and ease the concerns on the electricity generation, energy models which involve clean and renewable energy sources are necessitated. However, renewable sources are generally are intermittent and heavily dependent on the spatial location (e.g. sun doesn't shine constantly at any given time in any given place). Intermittency causes limited control on power output because of variability and partially predictability of the renewable sources such as solar and wind and dependence on the spatial location causes a mismatched between potential of renewable energy generation and where the energy will be ultimately consumed. One of the ways to mitigate the intermittency of renewable sources is to design hybrid systems which operate as a combination of alternative resources. Energy storage, long distance transmission line and demand response programs are other critical components of electricity generation and transmission system to improve power grid reliability/efficiency and integrate intermittent renewable energy sources.

To this end, in Chapter 3 and 4, we discuss the problem of modeling hydro and solar energy production and allocation, including long-term investments and storage, capturing the stochastic nature of hourly supply and demand data. We imagine a country in the future where hydro and solar are the dominant sources and fossil fuels are only available in minimal form, in the shape of diesel generators as an example. In this country, we first identify candidate basins for hydro power stations and aggregated demand point locations (the cities or the states of the country). Within demand points solar energy production is possible. Then, we determine the possible transmission network between supply and demand points. We mathematically model this hybrid energy generation and allocation system, where time variability of energy sources and demand is balanced with the water stored in the reservoirs of hydropower stations. The aim of this model is to obtain the optimal size of infrastructure needed to meet the demand, combining three major components of power systems to reduce the intermittency: hybrid systems, long distance transmission lines and energy storage. In these chapters, we present results with several cases studies from India which help answer whether solar energy in addition to high hydro power potential in Himalaya Mountains would be enough to meet growing electricity demand if fossil fuels could be almost completely phased out from electricity generation.

In Chapter 3, a scenario based linear programming approach is described for modeling the hybrid solar and hydropower ststem with conventional hydro storage. In this system, incoming stream flows are stored in large reservoirs in dams and water release is deferred until it is needed. We first provide results for a single basin-single demand point case for one sceario to show the the basic results that the model can provide. Then, we provide multi basin-multi demand point cases to show how the unit cost of the systems is reduced when geographic aggregation of sources are possible with transmission lines. A sensitivity analyis of the cost parameters used in the model completes the Chapter 3.

Furthermore in chapter 4, to the same problem in chapter 3, we apply pumped hydro power stations in which water is pumped from lower reservoir to upper reservoir during periods of low demand to be released for generation when demand is high. We show that introduction of the pumped hydro storage increases the utilization of both hydro and solar sources and decreases the unit cost of the system significantly. Now, as the solar energy can be also stored, it helps reducing the variability of sources. In this chapter, we introduce a heuristic algorithm based on decomposition of scenarios for the stochastic linear programming model.

An innovative contribution of the work in chapter 4 and chapter 5 is the establishment of a new perspective to energy modeling by including fine-grained sources of uncertainty such as stream flow, solar radiations and demand in hourly level as well as spatial location of supply and demand in national level.

Chapter 2

Power Distribution Systems Design for Rural Electrification

2.1 Introduction

The International Energy Agency estimates that the number of people who do not have access to electricity is nearly 1.3 billion and more than eight out of ten people without modern energy access live in rural areas. Attempts to increase electricity coverage require detailed understanding of the cost structure of electrification in rural settings. This chapter through its methodologies and algorithms enable tools for energy planning and policy making to make cost assessments in energy infrastructure planning rapidly and accurately.

Rural energy planning has, thus far, focused primarily on the national to regional scale with aggregated supply and demand information. However, effective implementation of the new energy technologies requires a new planning approach that can consider information across spatial scales. In a technological landscape that is altered by the emergence of off-grid and distributed approaches, there is a need amongst infrastructure planners to evaluate the costs of networked or grid approaches vis a vis off-grid approaches. However, estimating the investment cost of networked approaches requires the information for both spatial distribution of demand and the optimal placement of the infrastructure to meet that demand. To this end, in this chapter we first present a new data set of local-level demand points developed from QuickBird satellite imagery and provide new tools to estimate the cost of green-field power distribution system rapidly and with high accuracy considering local-level spatial distribution data.

In the first part of this chapter, we address the question of how the population settlement patterns influence the cost of electrification. Understanding the impact of the spatial structure of the population on infrastructure costs is critical in rural energy planning. If it is possible to connect only some portion of the households to electricity due to limited funding, how to optimally choose that portion requires detailed mathematical analysis. Given the possibility of acquiring detailed local-level information from remote-sensing data, we develop a new method for estimating a more detailed, per-unit cost of infrastructure, while taking into account population structure. Since settlement patterns influence the cost of local distribution networks while not directly impacting the costs of transmission and generation, we focus on these local distribution networks.

Moreover, in this chapter, we provide a heuristic algorithm which designs multi-level distribution system layouts while minimizing overall cost of infrastructure costs; specifically the combined costs of transformers and the two-tiered network together. To our knowledge, this algorithm is the first heuristic algorithm that selects the locations and service areas of transformers without requiring candidate solution and simultaneously builds two-level grid network in a green-field setting. Proposed methodologies can easily be adapted to other infrastructure problems such as designing communication networks and water supply network.

The algorithm we propose does not require a set of candidate locations to be considered as transformer locations. The maximum service distance in a low voltage distribution network is also pre-specified and determined from engineering practice. Given these costs, the demand points, the location of the HV network, and the maximum distance of the demand point from the transformer, the algorithm automatically finds the locations and service areas transformers as well as the LV and MV network layout with the goal of minimizing the total costs. This algorithm can serve as a tool for network engineers and planners to make rapid assessments assisting them with a) estimates of total cost of distribution, b) layouts of initial designs and c) breakdown of total costs into transformer cost and medium and low voltage line costs and giving them a good starting point for more detailed smart grid projects.

2.1.1 Rural Electrification Background

Electricity access is one of the most important components of rural developments. It has been shown that better living conditions in developing countries cannot be achieved without investments in electricity [1]. In rural areas where renewable energy resources are widely available, small off-grid standalone systems appears to be an attractive alternative [2]. Moreover, decentralized technologies seems to be more suitable for rural and remote areas due to the fact that it helps avoid long distribution lines with low load densities, underutilized transformers and losses in distribution. It also has been discussed that whether decentralized alternatives which use locally available resources provide more reliable supply of energy [2, 3]. Thus, most of the earlier research aims to investigate primarily renewable energy alternatives and off-grid technologies [4-9]. It is worth noting that, although there has been a lot of attention to rural electrification projects, literature on the networked approaches is very limited.

Moreover, rural energy planning has, thus far, focused primarily on the national to regional scale. Experience has been gained in design and implementation of rural electrification projects at the national level, with reviews such as that by Bekker (2008) [10] in South Africa, and Haanyika (2008) [11] in Zambia providing a basis for future policy and systems design. However, much work remains to be done, particularly on the local level. Effective implementation of the new energy technologies requires a new planning approach that can consider information across spatial scales. Local-level spatial information helps planners optimally locate critical energy infrastructure. For example, energy storage is an important component of the power systems; especially when effectively utilized with renewable sources, storage allows improved efficiency and lower cost generations. Moreover, optimal location of storages and transformers can help reduce distribution losses and increases the efficiencies of the power systems. The development of design strategies that consider local-level data in conjunction with the regional scale picture is important for planning cost-effective rural energy networks. In this chapter, an attempt is made on estimating the cost of rural networks to facilitate the rural energy decision making based on purely cost comparison without considering other consequences of off-grid and grid approaches.

2.1.2 Literature Review

The so-called power distribution system problem, in general, has been studied extensively in the literature [12-21]. Techniques developed in prior efforts for this complex problem usually divide the problem into sub-problems at each level and then solve each sub-problem separately using various optimization techniques [12-16]. These studies differ from each other in how they represent the problem components as well as in the algorithms utilized. None of these studies address the problem of designing both LV and medium voltage (MV) networks in a single framework. However, dividing the problem into sub-problems and solving them separately reduces the probability of reaching an optimal solution and prevents us from seeing the effects of different cost parameters on the final network layout. The methods that have been proposed in the literature are based on either mathematical programming techniques such as Mixed Integer Programming, Branch and Bound Method [14, 15, 19] or heuristic algorithms such as Genetic Algorithm [13, 20, 21]. However, complexity of the models and the algorithms reduces their applicability to estimate the cost of networked approaches in rural electrification discussions when spatial distribution of a very large data set (demand points) is available. In addition, regardless of the solution methods, all studies mentioned here, except for [13], includes pre-assumption of candidate locations for transformers or feeders. These studies do not provide a method to update the candidate transformer locations during the search for an optimum solution.

Therefore, the final feeder network is strictly dependent upon the initial selection of candidate locations. In practice, however, determination of candidate locations is not always a simple task, and if the methodology has to scale to a larger number of demand points, clearly the transformer locations should be an outcome of the optimization process.

2.2 Input Data

The algorithms of this chapter have been tested on household level data from nine sites in Sub-Saharan Africa shown in Figure 2.1. The data were digitized from QuickBird satellite imagery of sites; details of the digitization method are discussed in Zvoleff and Kocaman et al. (2009) [22]. For most of the sites, a QuickBird image covers an area of $10x10km^2$ which covers a large representative area. Even though all structures do not necessarily correspond to households, we assume that each structure represents a demand point which needs to be electrified.

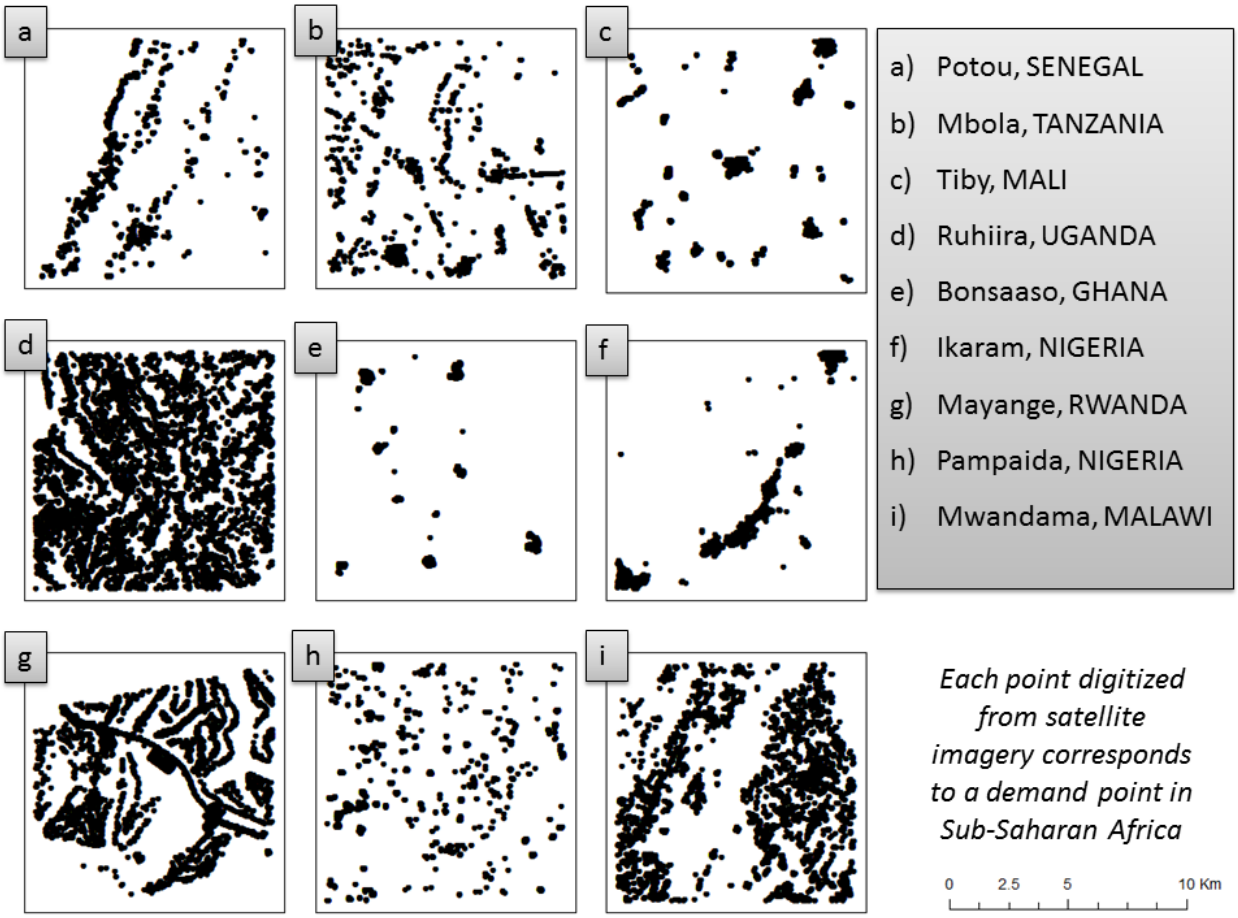


Figure 2.1 | Demand point locations for nine Sub-Saharan Africa sites digitized from satellite imagery

Although the data acquired thus far shows a broad range of spatial distributions of populations, there are common features that have been observed. Qualitatively, sites can be classified as sparse or dense, and nucleated or dispersed. A sparse site has a relatively low density as compared to a dense site. A nucleated site shows clustering of population around certain centroids, whereas a dispersed site is closer to a random distribution of points around the landscape.

Using these relative terms, a site can therefore be nucleated but sparse, or dense yet dispersed. While Ruhiira, Uganda is densely populated (compared to the other areas we discuss here) it is not nucleated, although topography and local road networks influence settlement locations to some extent. Tiby, Mali exemplifies the most nucleated areas, with a small number of very dense clusters of structures dotting the landscape. Potou, Senegal and Mbola, Tanzania fall in between these two extremes. Potou shows nucleation in the south and along the coast in the west, and is otherwise dispersed. Mbola features one large cluster in the south, while the remaining area is sparse.

2.3 Single Level Power Distribution System Design

Here, we discuss a methodology to estimate the cost of local-level distribution systems for a least-cost network. The main questions we want to answer are: i) How does rural population structure affect the cost of infrastructure investments? ii) If only pre-specified percentage of the households are connected, how does the per-unit cost of grid construction vary with this percentage (this percentage will be referred as the penetration rate)? We also provide additional information of interest to policy makers, such as the marginal cost of connecting additional households to a grid as a function of the penetration rate.

To see the effect of settlement patterns of households on the distribution system cost, we first made several simplifications. In this section, the distribution system consists of only Low Voltage (LV) lines we assume that total cost of the network is linearly related to the lengths of the network. Moreover, additional cost of transformers is assumed to simply be proportional to low-voltage network length. This is actually a realistic assumption as in in rural settings the local, low-voltage lines network is the dominant components of the infrastructure [23]. Thus, here we can define the cost in units of lengths instead of monetary units. To be able to compare different possible networks, we define the unit cost of a network as the mean inter-household distance (MID) which is simply calculated by dividing the total length of the network with the number of connections in that network.

2.3.1 Full Connectivity

In full connectivity case, we are interested in connecting households to the electricity with 100% penetration rate. The least cost network which spans all the households without creating a cycle is called minimum spanning tree (MST). An MST easily be calculated by using the well-studied Minimum Spanning Tree (MST) algorithms [24-26] which aim to find a tree (i.e. network containing no cycles) that spans all the points minimizing the total length of the network with the guarantee of the exact optimal solution. Although there are different MST algorithms in the literature; in this thesis, Prim's algorithm [24] is implemented as it has better running time performance for dense sites [27]. Prim's MST algorithm starts with choosing a starting point and adds the shortest segment of this point to the network. Until all nodes are spanned, the shortest segment emanating from the existing points on the network is added. The connections that would create cycles are avoided. We note that since the MST algorithm finds the optimal network, changing the starting point will not affect the result as all starting points will end up with the same network. A five node illustration of MST algorithm can be seen in Figure 2.2.

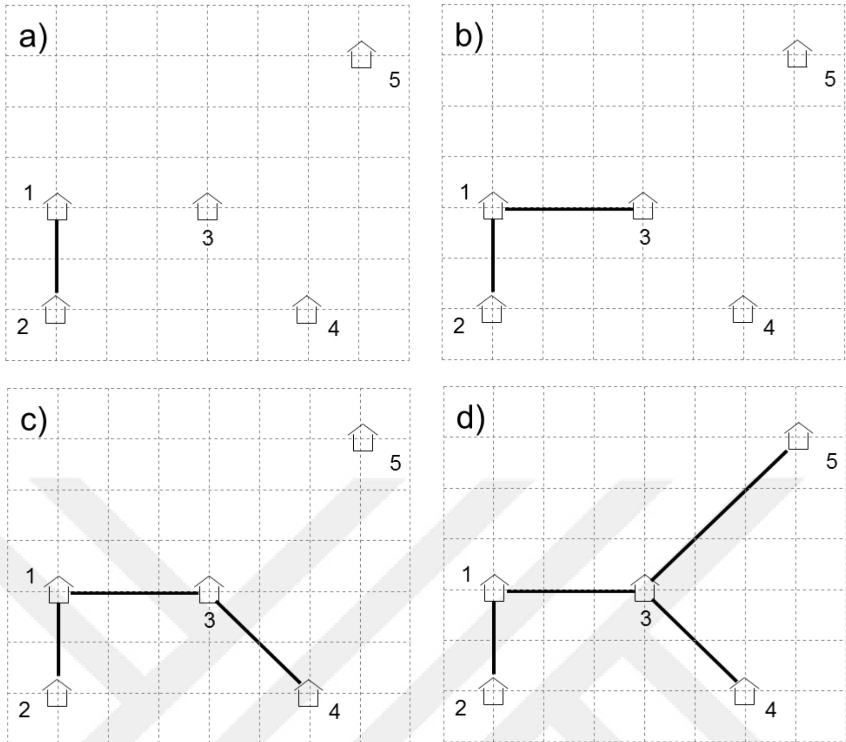


Figure 2.2 | Illustration of MST Algorithm

2.3.2 Partial Connectivity

Given limited funding, one of the most important concerns for infrastructure planners is the initial decision to extend the network into an unserved area. If it is possible to connect only some portion of the households to electricity, how to choose that portion is a very critical question. One common approach in this case is to expand the network to the strategic demand points where the total cost or per-household demand is minimized. This approach usually results with excluding sparsely inhabited areas from the network to avoid very long cable runs and underutilized transformers and exposes the question of how the population structures offer potential cost savings due to partial penetration or connectivity.

Partial connectivity problem that we study here is to construct a network such that the mean inter-household distance (MID) is minimized subject to the requirement that the network connects a pre-specified percentage of the households. In graph theory, an abstraction of this problem is known as the k-MST problem. The objective of the k-MST problem is to find the least cost network which spans at least k nodes on a graph. When penetration rate is 100% (k equals to the number of all points), optimal solution can easily be calculated with the greedy minimum spanning tree algorithms. However, k-MST problem is an NP-hard problem [28] and several attempts have been made to find the approximation algorithms for the problem [29-33].

For the problem of minimizing the network length subject to the penetration rate, Composite Prim's Algorithm (CPA), a simple and computationally efficient approach, is developed by Zvoleff and Kocaman et al. (2009) [22] based on the Prim's algorithm. For a fixed penetration rate, Prim's algorithm is modified as follows. For each initial demand point, the Prim's algorithm is run until the required penetration rate is achieved. By running this algorithm repeatedly, using a different initial demand point for each run, a series of different networks is calculated. Then for each penetration rate, the network which has minimum MID is selected.

As the CPA is based on Prim's algorithm, it is important to recognize that it necessarily builds subgraphs of the MST. Therefore, any network produced by the CPA can be expanded to full penetration with no penalty. The same is not necessarily true for expansion of a minimum cost network spanning a subset of the population to another network also providing less than full coverage. For example, the optimal network for 10% penetration may not be a subgraph of the optimal network for 20% penetration, as the two may start from different starting points. For example, a planner interested in constructing a network spanning only 10% of the grid and who is planning for later expansion up to 20% connectivity could minimize their expansion costs by estimating with the CPA the optimal network for 20% connectivity, and building initially building a network that is a sub-graph of this network. The CPA appears to handle these cases reasonably well, based on the observed empirical results; however we don't know the how the solution of the heuristic algorithm is compared to optimal solution. For this reason, we compare the results of the CPA algorithm with another approximation algorithm which has a proven performance guarantee.

An α-approximation algorithm for a minimization problem runs in polynomial time and guarantees that the cost of the solution is no more than α times of the optimal solution. The value α is called performance guarantee or approximation ratio of the algorithm. Here, we present the results of the CPA algorithm together with the results of a 2-approximation algorithm proposed for k-MST problem by Chudak et al. (2001) [33]. Their approach is based on the fact that the Lagrangian version of the *k*-MST problem is a well-studied Prize Collecting Steiner Tree problem, for which a similar guarantee has been developed using the so-called primal-dual algorithm [34]. The primal-dual algorithm searches for the best possible solution by keeping track of both the feasible solution and its shadow prices related to the constraints. While the algorithm by Chudak et al. (2001) [33] has a theoretically appealing property, its implementation is much more complex than that of the CPA. We have implemented both algorithms, and find that both have similar performances. As seen in Figure 2.3, the mean cost curves obtained from these two algorithms match closely.

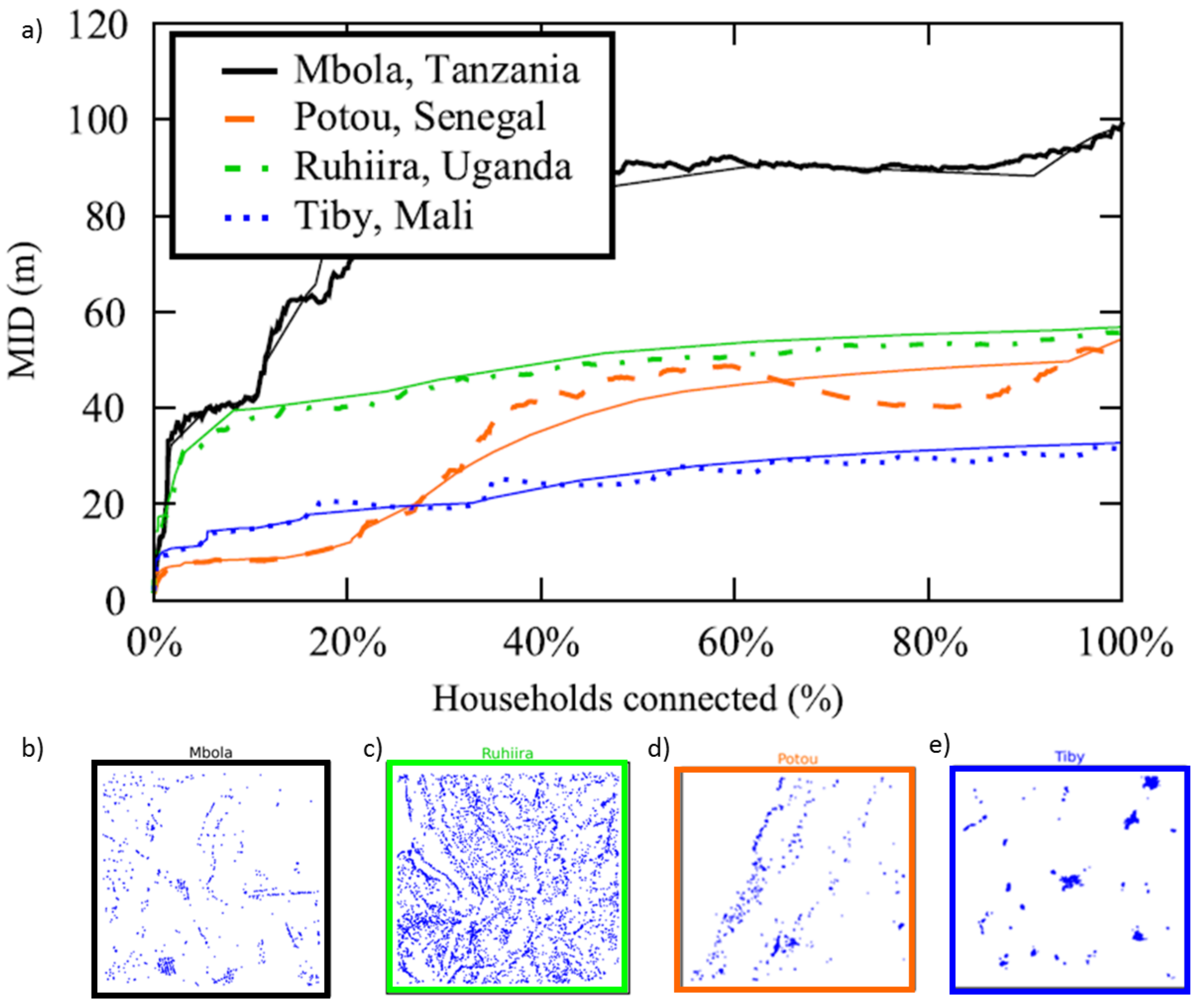


Figure 2.3 | a, Comparison of results of the composite Prim's algorithm (thick lines) with those obtained from the k-MST algorithm proposed by Chudak et al. (2001) (thin lines). **b-e,** Households distribution in Mbola, Ruhiira, Potou, Tiby.

To understand the effect of spatial distribution of households in the cost of electrification, Figure 2.3a should be examined together with Figures 2.3b-e. The area of Mbola, Tanzania considered here shows a sparse pattern of population, with little clustering of households except for two areas in the southwest. These dense areas allow interconnections within a portion of the population (up to about 15%) with a relatively low MID. Connecting the remaining population, however,is expensive. Ruhiira, Uganda is far more densely populated than the other sites considered here. Ruhiira clearly has the least clustering of the four distributions shown here. The lack of clustering leads to a situation where after 5% penetration, the MID is essentially the same.

Potou, Senegal is sparsely populated, with two nucleated areas; one in the south and another in the northwest along the coast. The MID rise slowly as the largest population center (in the south of the image) is connected to the grid. After 20% penetration, the MID begins to rise more quickly as the outskirts of the main cluster are connected to the grid. After 60% penetration, the network begins to reach nucleated areas along the coast. The population of Tiby, Mali is split into several nucleated clusters, with few outliers. The ease of connecting these dense, nucleated clusters allows connection of the entire population with an MID of only 32.7 m. However, the large separation between population centers in Tiby leads to "jumps" in the curve as connections are made between clusters.

2.4 Multi-level Power Distribution System Design

In the previous section, we assumed that each household can be connected to electricity with low voltage lines without considering a capacity constraint. We also assumed that transformer costs are linearly related to the length of the network. Here, we study a more realistic problem where we have three different cost parameters: 1) the cost per meter of LV line, CLV; 2) the cost per meter of MV line, CMV; and 3) the unit cost of a transformer, CT. The capacity constraint on the transformers, Dmax1, is defined as the maximum service distance and modeled as the radius of coverage of a transformer.

The algorithm we present in this section combines the transformer location problem and the Low Voltage (LV) and Medium Voltage (MV) network design problem into a single problem and solves them in a single optimization framework. We propose a heuristic algorithm to design a two-level radial power distribution system. The *first level* includes the determination of the numbers, locations and capacities of transformers that feed an LV distribution network. The

[1] This distance would vary over the network with local geography and topography but is assumed constant here.

transformers represent load points for an upstream MV network and the MV network is also determined as a part of the first level. The *second level* includes the determination of the layout of the low voltage network between the transformers and the specified ultimate demand points. Note that the high-voltage (HV) network (for that matter source points) further upstream of the MV network are assumed to be known[2]. One could have further generalized the problem to include the determination of the HV networks as well, making it a three level problem, but here we consider the HV network as pre-specified for simplicity.

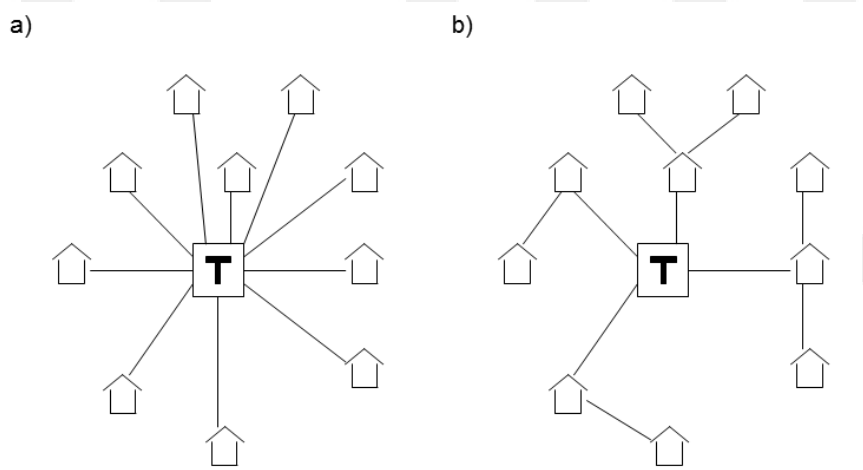


Figure 2.4 | Examples network configurations. **a,** Point to Point network (star configuration). **b,** Multi-point network.

Locations of ultimate consumers are called "demand points" in the rest of the chapter. Each demand point is assumed to have the same load and the load is assumed not to change over time making the problem "static". Distribution system is designed to be radial, to have one path between demand points and transformers, due to the fact that it is the most widely used form of

[2] The medium voltage network connects these transformers to electricity sources further upstream where they could either be sub-stations of a high voltage transmission network or power generating stations. Note that the costs associated with what we call here for convenience High-voltage networks or transformations from High-voltage to Medium-Voltage are not considered here.

distribution design and it is the cheapest and the simplest alternative compared to loop and networked designs [35]. In radial design, since there is only one path between demand points and transformers, power flow is certain and the system can be operated easily. The major drawback to radial feeder design is reliability. Any equipment failure will interrupt service to all customers downstream from it. However, low statistical rate of failure of equipment on the low voltage level makes the adaptation of radial systems easier [14].

Within the service areas of the transformers, the low voltage network is permitted to be multi-point, in that, in order to minimize costs the wire to a demand point further in distance can first go through one or more intermediate demand points. This architecture is called a "multi-point" LV network here (See Figure 2.4b). Maximum distance capacity of an LV line is then defined as another design parameter and called L_{max} (i.e. the maximum LV line used to connect a demand point to the transformer directly or through other demand points should be less than L_{max}[3]). L_{max} value should be used to limit the maximum total load on LV line and should be greater than or equal to D_{max} so that each demand point within the service area of a transformer gets connected.

Given the cost parameters and subject to the constraints described above, the desired outputs of the algorithm are:

- number and locations of the transformers;
- medium voltage (MV) network that connects a source point to the transformers; and
- low voltage (LV) network between the demand points and transformers.

---

[3] *Lmax* can also be considered as a constraint on distribution losses in LV level as the losses and wire lengths are linearly related.

Our objective function is the minimization of total system cost, which includes cost of transformers, cost of low voltage and medium voltage networks. Schematic illustration of our problem formulation can be seen in Figure 2.5.

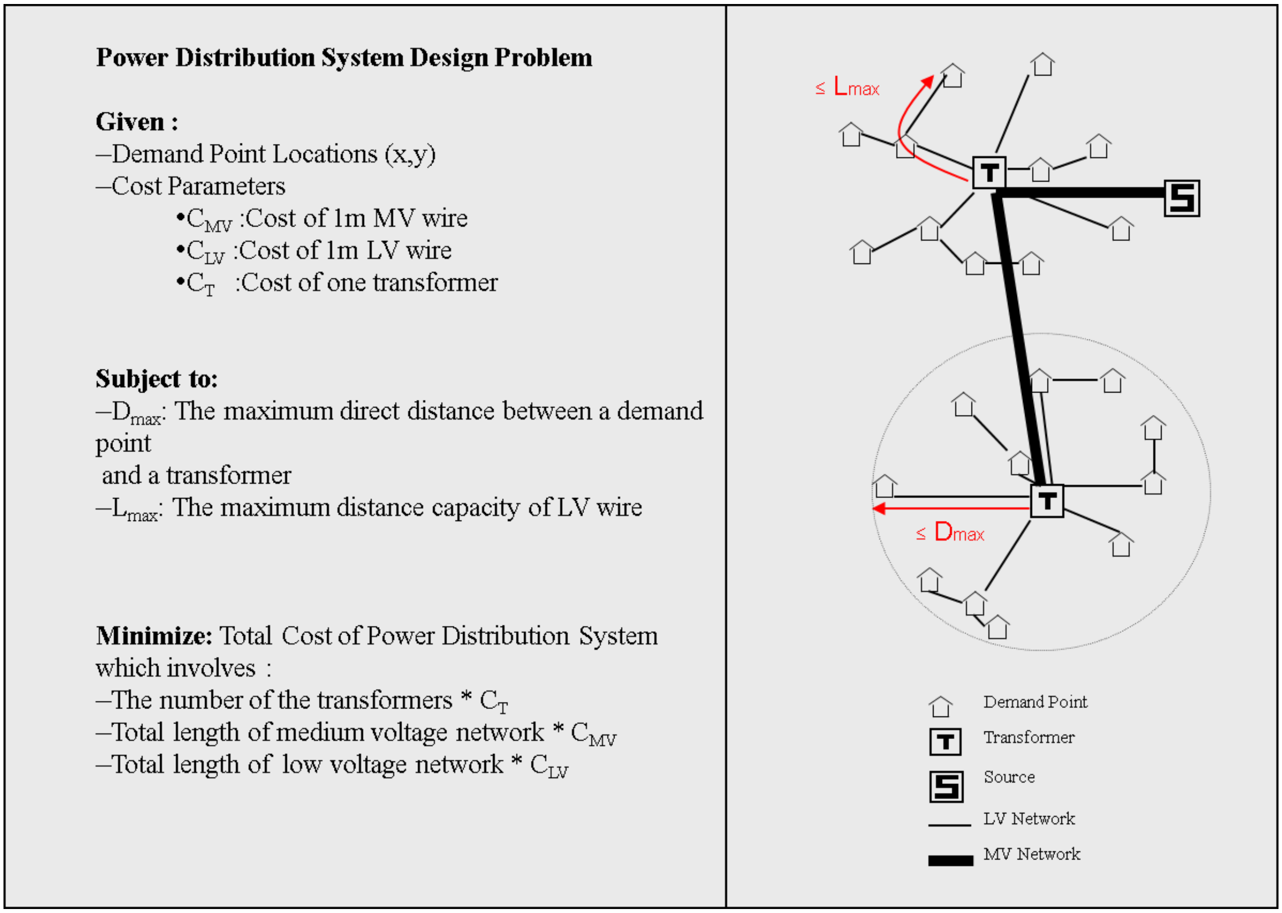


Figure 2.5 | Illustration of Power Distribution System.

2.4.1 Methodology

Given the difficulty of the problem, a heuristic algorithm is developed to place transformers and locate the networks. The algorithm relies on a “greedy” approach that starts with a stage that each demand point has one transformer (i.e. for n demand points, there would be n transformers) and iteratively decreases the number of transformers. Initially, transformers are connected to a prescribed single source point and to each other with a least cost medium voltage network; there is no LV network at this stage. The cost of this design is computed and provided as an initial upper bound to the cost of the network design since it is a feasible solution to the problem. It contains the maximum possible number of transformers and maximum length of MV line (which per unit length is more expensive than LV line). Then, one begins the process of eliminating transformers one at a time while removing some of the MV lines and adding new LV lines in the process to find a least cost network. The algorithm consists of the repeated applications of the following iterations:

- Search for the closest pair of transformers which can be replaced by a single transformer located at the centroid (center of mass) of the demand points without violating, D_{max} constraint.
- Build the MV network between the updated set of transformers and the source point. (See Section 1.4.1.2 for details)
- Build the LV network between the demand points that are no longer served directly by transformers using LV line, ensuring constraint $L_{max.}$ (See Section 1.4.1.3)
- Compute the new overall cost.

The heuristic algorithm continues this iterative process until the number of transformers cannot be reduced any further without violating the D_{max} constraint. (Note that the solution with the least number of transformers is not necessarily the least cost since the design with the least number of transformers may have been obtained by adding more LV line length, and this trade-off may not be favorable to the total cost.).

At this stage, all the computed costs during the transformer elimination process are compared and the least cost network design is selected. With one transformer replacing a "pair", and process repeated, one can think of the sets of demand points being served by one transformer as a “cluster”. With this perspective the algorithm is an agglomerative algorithm, using a bottom-up approach to iteratively agglomerate (merges) the closest pair of points. (See Section 1.4.1.1 for details)

The algorithm is further analyzed in its three components; i) selecting the transformers to be removed, ii) creating a MV network among the transformers and iii) connecting the demand points and the transformers with an LV network. A flow chart of the algorithm can be seen in Figure 2.6.

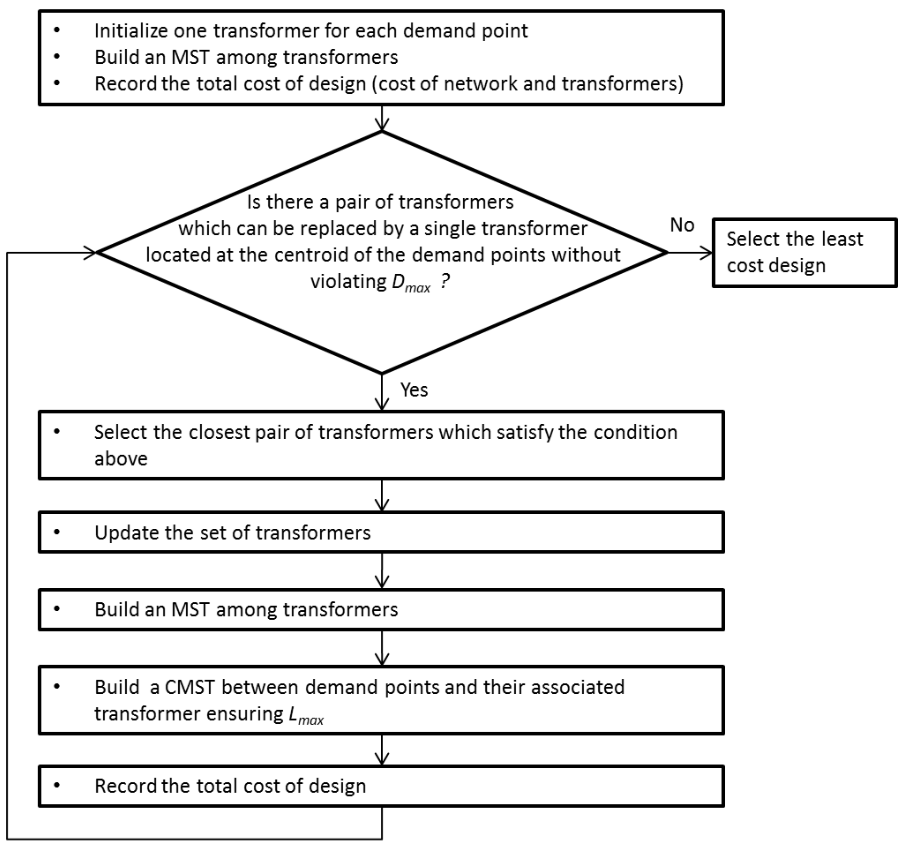


Figure 2. 6 | Flow Chart of the Heuristic Algorithm

2.4.1.1 Locating the transformers

A capacitated agglomerative hierarchical clustering method is adapted to find the locations of transformers since it does not require pre-selected candidate locations for transformers. Note again that this is different than most of the work in the literature and is a must for our target demand points, households in rural Africa with little or none existing infrastructure.

An agglomerative hierarchical clustering method starts with as many clusters as the number points to be clustered. At each step, the clusters are merged according to a rule and eventually only one cluster remains where all points are connected. In contrast, a *capacitated* agglomerative hierarchical clustering method has no assumption on the final number of clusters [36] and finds the minimum possible number of clusters that can be achieved with the given constraints. In clustering methods, many rules can be used depending on the problem definition. In our problem, to be able to incorporate the D_{max} conveniently, the closest pair based on the Euclidean distance between transformers has higher priority to be merged. The applications of clustering methods on similar problems using Euclidean distance can also be seen in [36-40]. Here, as opposed to stopping the process when the best possible agglomeration violates the capacity constraint (infeasible) [37], we choose the next best feasible agglomeration (if there exists one) as proposed by [39].

An illustrative example of our agglomerative clustering approach is presented in Figure 2.7. In this example, we have five demand points and D_{max} is 2. Figure 2.7a represents the initial configuration and Figure 2.7b-d show how the closest pairs of transformers which do not violate the capacity condition are connected one by one. No further change is possible in Figure 2.7 d since the agglomeration of the final two clusters would violate the D_{max} constraint.

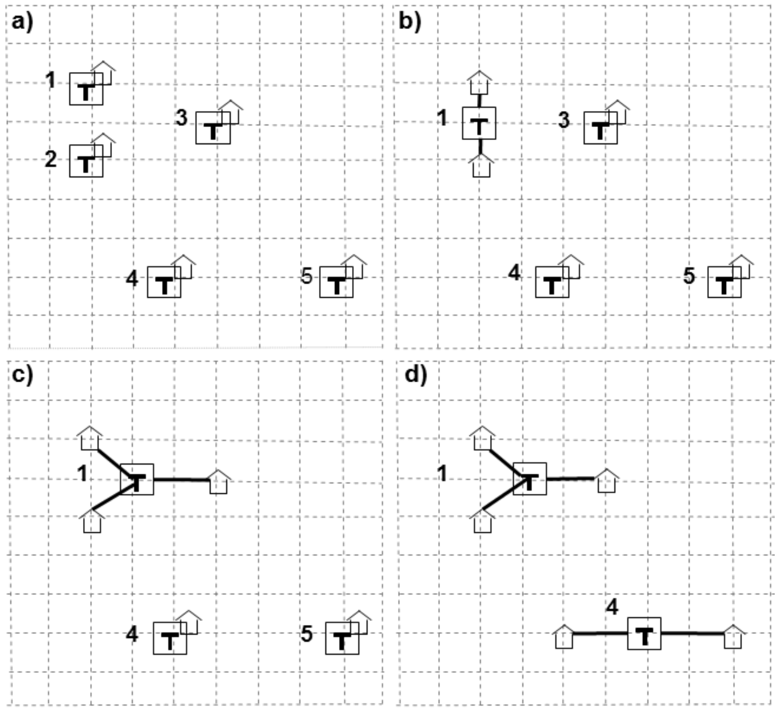


Figure 2.7 | Agglomerative Clustering Example. **a,** Initial configuration where each demand point has one transformer. **b,** The closest pair (1 and 2) gets connected. **c,** Next closest pair is 1 and 3. **d,** 4 and 5 are connected and no further change is possible without violating D_{max} constraint.

2.4.1.2 Medium Voltage Line Layout

At any iteration, when transformer locations are known, the problem is to find the least cost layout that connects the transformers and the given source point. This can easily be solved by using the well-studied Minimum Spanning Tree (MST) algorithms discussed in Section 2.3.1, which aim to find a tree (i.e. network) that spans all the points minimizing the total length of the network[4] with the guarantee of the exact optimal solution [26]. For the problem at hand, the points represent the transformers and the source point.

2.4.1.3 Low Voltage Line Layout

---

[5] Since there is no constraint on the maximum length of MV line, the least cost solution is simply the least total length.

As "clusters" emerge during each iteration cycle, an LV network needs to be laid out within each cluster. The total length of the connections is minimized while ensuring that the length of LV line between a transformer and the demand points are always less than a given L_{max} value.

Constructing a least cost multi-point network (cost efficient compared to star configuration, see Figure 2.4) is similar to the minimum spanning tree (MST) problem, but with an additional L_{max} constraint. This extra condition converts the problem into a capacitated minimum spanning tree (CMST) problem. CMST is a minimal cost spanning tree which has a designated root (transformers) and a capacity constraint which ensures that the length of a sub-tree incident on the root does not exceed a certain distance (L_{max}). This problem is well studied for its applications to centralized communication design. Unlike MST, CMST cannot be efficiently solved in polynomial time. However, Essau-Williams [41] and Sharma [42] developed heuristic algorithms to solve the CMST problem and Chandy and Russell [43] showed that these heuristics find near optimal solutions within 10 percent (often 5 percent) of the optimal solution.

Essau-William's heuristic algorithm is implemented in order to get the least total length LV layout within each cluster. A CMST algorithm starts with connecting all demand points to the transformers using the star configuration. The procedure to go from star configuration to a multi-point configuration is simply the successive iterations of calculating the trade-off values from removing the direct connections between the demand points and the transformer ,as well as, adding indirect connections through their neighbors. At each iteration, the trade-off value is computed for every demand point and the largest trade-off value (i.e. the greatest improvement) that does not lead to a violation of the L_{max} constraint is used to update the network. The algorithm terminates when there is no further improvement possible.

An example of CMST algorithm with L_{max} value of 5 is presented in Figure 2.8. It starts with the star configuration in Figure 2.8a, and constructs the multi-point connections in Figure 2.8b-2.8d until there is no further change possible with the given L_{max}. Notice that although the trade-off value of point 1 is greater, its direct connection to the transformer is not removed due to the constraint.

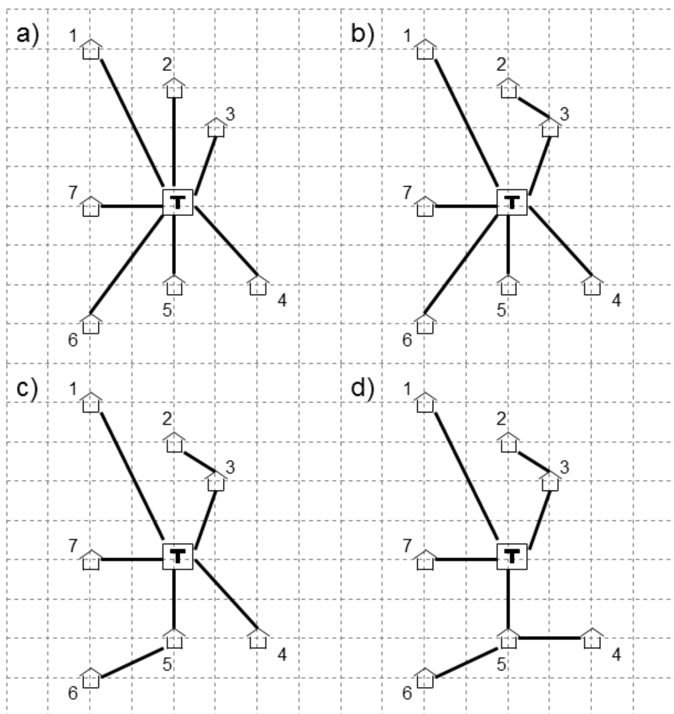


Figure 2.8 | A seven-point example of Essau-William's CMST algorithm. **a,** Initial star configuration. **b,** Maximum trade-off value is for Point 1 (2.23) however it violates Lmax. Point with the next best trade-off value (Point 2) is selected. **c**, Next best trade-off value is for Point 6. **d,** Final configuration is reached after Point 4 is connected.

2.4.2 Results on the sub-Saharan Africa sites

The algorithm is tested on household level data from nine sites in Sub-Saharan Africa shown in Figure 2.1. In rural electrification programs, the clusters of demand points and loads are small, so we use representative costs corresponding to the typical small 25kVA transformer [35, 36]. Cost parameters and other constraints consistent with rural electrification practice [37-40] are assumed to be: D_{max} = 500 meters, L_{max} = 600 meters, C_{LV} = \$10/meter, C_{MV} = \$25/meter, C_T = \$5000.

In Table 1, statistics and the outcome of our algorithm are shown for all sites. These results can help network engineers and planners estimate the number of transformers and the LV, MV line lengths easily for a particular location. In addition, by using our algorithm, they can also quantify their empirical observations. For example, from Figure 2.1 b-e, Mbola seems more dispersed than Bonsaaso (both have around 1000 households). Therefore the dynamics of their networks are expected to be different by observation. The algorithm indeed outputs 90 transformers for Mbola and 18 for Bonsaaso, and, the cost per household is more than 2.5 times for Mbola than for Bonsaaso. Similarly, Tiby (Figure 2.1 c) seems highly nucleated and this leads to shorter (cheaper) connections. Hence, with \$691 Tiby has the lowest per households cost among all sites.

Table 2.1 | Algorithm results for nine sub-Saharan Africa sites

	Area of the Image (km^2)	Number of Demand Points	Number of Transformers	Total Length		Average Length per Household		Cost Distribution			Total Cost (USD)	Average Cost per Household
				MV (m)	LV (m)	MV (m)	LV (m)	Transformers	MV	LV		
Potou, SENEGAL	95.5	1797	71	57,644	63,742	32.08	35.47	14.60%	59.20%	26.20%	2,433,515	1,354
Mbola, TANZANIA	100	1175	90	70,762	81,320	60.22	69.21	14.80%	58.30%	26.80%	3,032,239	2,581
Tiby, MALI	100	2496	32	42,026	51,361	16.84	20.58	9.30%	60.90%	29.80%	1,724,264	691
Ruhiira, UGANDA	100	6434	212	124,293	351,299	19.32	54.60	13.80%	40.50%	45.70%	7,680,302	1,194
Bonsaaso, GHANA	100	993	18	27,775	22,334	27.97	22.49	8.90%	68.90%	22.20%	1,007,723	1,015
Ikaram, NIGERIA	100	1484	33	33,572	46,417	22.62	31.28	11.20%	57.20%	31.60%	1,468,479	990
Mayange, RWANDA	100	3909	114	68,746	185,987	17.59	47.58	13.70%	41.40%	44.80%	4,148,506	1,061
Pampaida, NIGERIA	100	1570	88	73,576	68,863	46.86	43.86	14.80%	62%	23.20%	2,968,042	1,890
Mwandama, MALAWI	100	4230	152	97,763	213,116	23.11	50.38	14.20%	45.80%	39.90%	5,335,237	1,261

From the test results, it is also possible to reach some generalized conclusions. For all sites, overall transformer costs are between 8%-15% of the total cost. In addition, for sites that has 1000 to 2500 household for 100 km^2, the bulk (~60%) of the total cost is composed of MV voltage lines. In the dense sites (Ruhiira, Mayange and Mwandama), there is a greater number of households per transformers and the total cost of LV line is comparable or even higher than the total cost of medium voltage lines despite the fact that the MV line is more expensive than LV line.

Furthermore, as a result of the rural electrification programs, some of the sites such as Mwandama, Pampaida and Mbola have already partial existing grid. When we compare our medium voltage network with the existing medium voltage line in these sites, we observe a highly good match. In Figure 2.9, the overlap between the existing grid and the proposed grid is shown. This indicates that the planners may benefit from our algorithm in estimating the network structure and related costs also for these sites where there is an existing partial network.

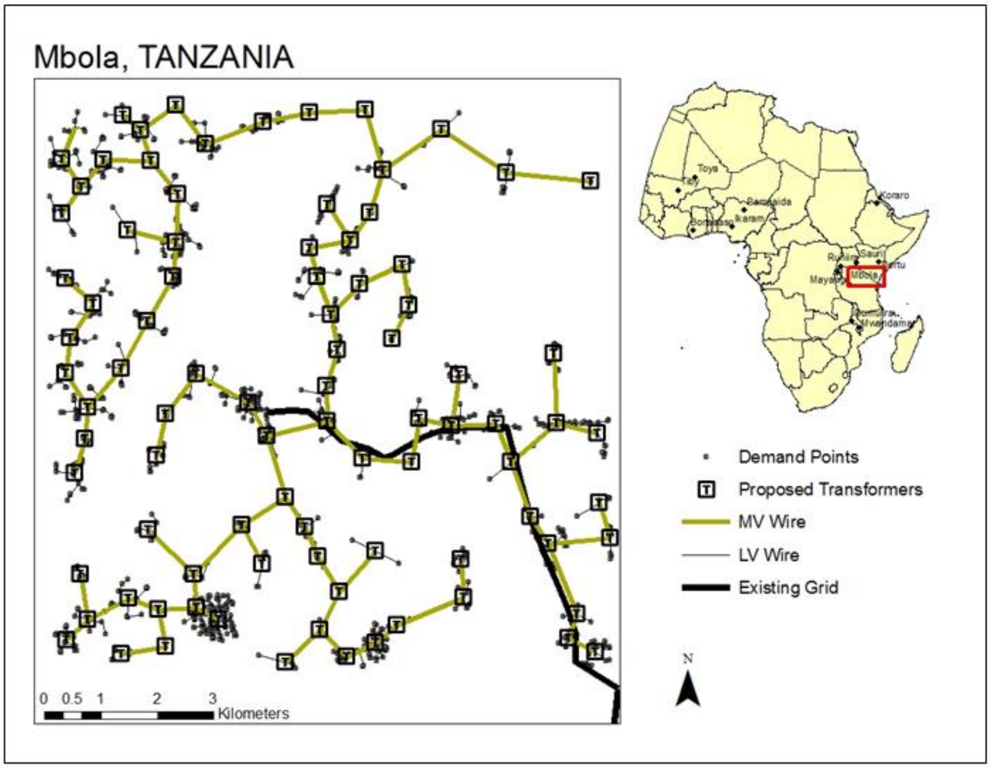


Figure 2.9 | Match between proposed network and existing grid. Proposed transformers, LV and MV networks compared to partial existing grid. Algorithm outputs 90 transformers for 1175 demand points.

2.5 Discussion on Multi-level Power Distribution Systems Results

In this section, we first assess the sensitivity of the two-level power distribution networks in terms of the cost parameters for LV, MV lines and transformers (Section 1.5.1). This analysis can be significantly useful for policy planners to estimate the total cost fluctuations due to individual cost parameters. Next, considering that the complexity and the noise in the real data from Sub-Saharan Africa may complicate the understanding of our results, we also test our algorithm on simplified (simulated) datasets (Section 1.5.2) and see that we get consistent results for the artificial data [49]. Then, we compare our algorithm with a sequential approach (Section 1.5.3) to see the relative performance and finally, we discuss the limitations and possible extensions of the algorithm for more detailed planning of power distribution systems and other parts of infrastructure problems such as siting of health or educational facilities (Section 1.5.4).

2.5.1 Network Sensitivity Analysis in Terms of Cost Parameters

In our base case results, the cost parameters; C_{LV}, C_{MV}, C_T; are chosen as \$10, \$25 and \$5000 respectively. We perform a sensitivity analysis to understand whether there is a drastic change in the final network due to slight movements in these cost parameters. For the sake of generality, we present the general behaviors of the algorithm on uniformly distributed randomly generated points (1000 demand points on 10X10km^2). At the end, we also present several runs of the algorithm on the data from Sub-Saharan Africa and verify the generalizations in real datasets.

2.5.1.1 Analysis with MV and LV Line Costs

First, we define a ratio, p, between cost parameters of MV and LV lines (i.e. $p=C_{MV}/C_{LV}$). When the transformer cost parameter, $C_{T,}$ is set to zero, the differences in final networks help us understand the sensitivity of final network design to the ratio between cost of MV and LV lines. Initially every demand point has one transformer; therefore maximum amount of MV and zero LV lengths are used in the system. Figure 2.10 shows how the total length of MV and LV lines change as the algorithm reduces the number of transformers from number of demand points to the minimum number of transformers subject to D_{max} constraint.

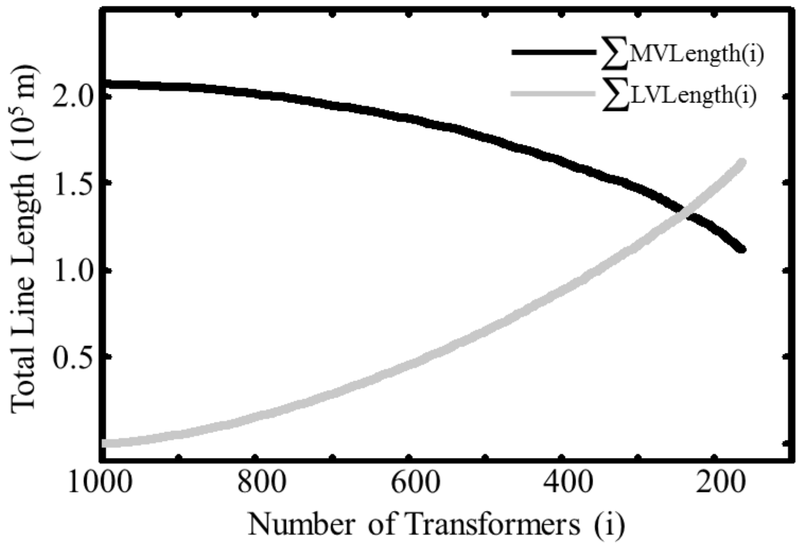


Figure 2.10 | Change in total length of MV and LV lines. As the algorithm decreases the number of transformers for 1000 uniformly distributed randomly generated demand points data within a 10X10km^2 area, total MV line length decreases, while total LV line length increases

In Figure 2.11 a, we show the change in number of transformers as the *p* ratio is increased from 1 to infinity. We interestingly observe that there is a critical value (p^*) such that for all values less than p^* one transformer for each demand point (maximum number of transformer case) is the minimum cost design. For other values, greater than or equal to p^*, the solution includes almost the minimum possible number of transformers without violating the maximum distance (D_{max}) constraint between demand points and transformers. Here, the critical p value (p^*) is observed around 1.70.

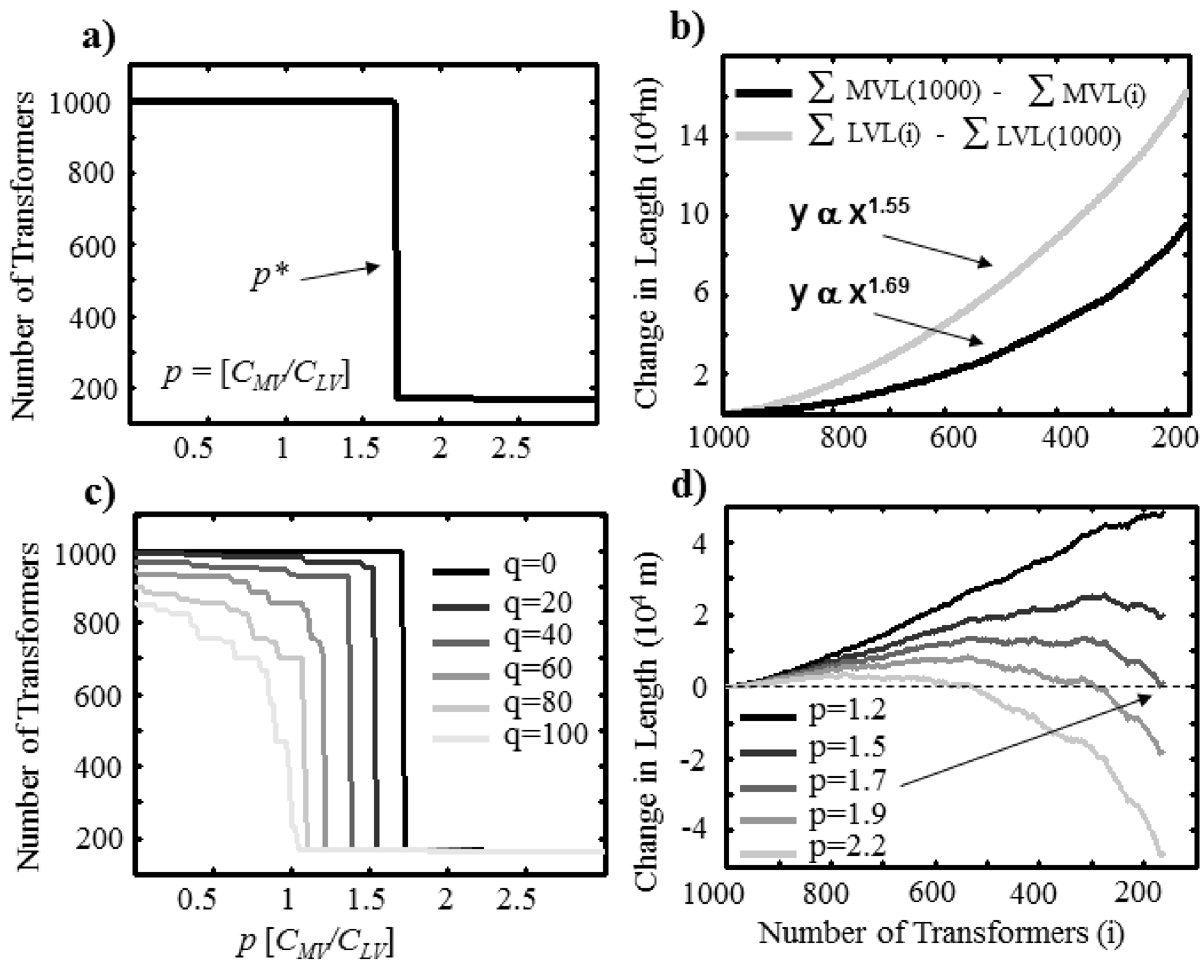


Figure 2.11 | Network Sensitivity Analysis. a, The sudden change in the number of transformers as the ratio between C_{MV} and C_{LV} cost parameters increases. **b,** Change in total MV and LV lengths from the 1000 transformer case. **c,** The change in the number of transformers for different *q* ratios as *p* ratio increases.(See Section 5.1.2 for q). **d**, The difference between the curves in b with MV weighted with different *p* ratios.

To understand this sudden change further, we need to think about how the algorithm works in each step. The total cost is given by the following equation;

$$TotalCost_i = C_T \times i + C_{MV} \times \sum MVLength_i + C_{LV} \times \sum LVLength_i$$

And when we assume there is no transformer cost, it becomes

$$TotalCost_i = C_{MV} \times \sum MVLength_i + C_{LV} \times \sum LVLength_i$$

The algorithm compares the total cost values in determining the number of transformers. Therefore, the only way to decrease the number of transformers is if

$TotalCost_i \leq TotalCost_j$ where $i < j$. This means;

$$C_{MV} \times \sum MVLength_i + C_{LV} \times \sum LVLength_i \leq C_{MV} \times \sum MVLength_j + C_{LV} \times \sum LVLength_j$$

and

$$p\left(\sum MVLength_j - \sum MVLength_i\right) \geq \left(\sum LVLength_i - \sum LVLength_j\right)$$

Here, for small p values, the change in the transformer number is not profitable because it also makes LV longer and cost of the LV is not cheap enough to make above equation hold (see the black (darkest) line in Figure 2.11d). Then algorithm results in one transformer for each demand point. However, after a critical p value (p=1.70 in Figure 2.11a) cost saving becomes possible by decreasing number of transformers within the allowed configurations. Once the algorithm favors less number of transformers, for all higher p values it goes all the way down to the minimum number of transformers subject to the D_{max} constraint because decrease in the MV length is faster than the increase in LV length; the fewer the number of transformers the less total cost (see Figure 2.11 b).

We perform the same analysis for some of the Sub-Saharan African sites and again obtain similar results; most notably, a sudden drop in the number of transformers. Table 2.2 summarizes the results. We observe that p^* differs based on the number of demand points and their spatial distributions. For example, for clustered sites (Bonsaaso and Ikaram) p^* value is smaller than the rest of the sites and p^* is highest for Mbola which is known as its dispersed settlement pattern. Since p^* is actually the ratio between the changes in LV and MV as the algorithm proceeds, it is expected to have smaller ratios for clustered sites where the decrease in total MV line is much faster than the increase in LV line for small number of transformers. Moreover, when D_{max} is set to 750 meters, algorithm finds less transformers than the 500 meters case and the sudden drop point happens at smaller *p** values as expected.

Table 2.2 | Critical *p** ratio for some of sub-Saharan sites

	Dmax=500 m				Dmax=750 m			
	p*	Drop		Ratio	p*	Drop		Ratio
		From	To			From	To	
Bonsaaso, GHANA	1.20	993	18	0.018	1.18	993	17	0.02
Ikaram, NIGERIA	1.31	1484	35	0.024	1.27	1484	30	0.02
Pampaida, NIGERIA	1.64	1570	288	0.183	1.63	1570	67	0.04
Mbola, TANZANIA	1.74	1169	93	0.080	1.64	1174	61	0.05
random1000	1.70	1000	169	0.169	1.54	1000	91	0.09

2.5.1.2 Analysis with LV Line and Transformer Costs

To be able to see the effect of transformer cost on the final network design, another ratio between C_T and C_{LV} is defined here as $q=C_T/\,(C_{LV}*D_{\max})$ and the cost of MV line is set to zero (since the cost contributions from C_T and C_{MV} are in the same way). Unlike the p ratio, D_{max} is introduced in the *q* ratio to make it dimensionless.

The total cost is given by

$$TotalCost_j = C_T \times j + C_{LV} \times \sum LVLength_j$$

To be able to decrease the number of transformers

$$TotalCost_i \leq TotalCost_j$$

$$C_T \times i + C_{LV} \times \sum LVLength_i \leq C_T \times j + C_{LV} \times \sum LVLength_j$$

$$\left(\sum LVLength_i - \sum LVLength_j\right) \leq qD_{\max}(j-i)$$

This time there is no sudden change and the number of transformers decreases gradually as *q* increases (Figure 2.12a). The S-shaped behavior is due to the fact that closer nodes are connected first so the cost from LV length is initially smaller and the total cost is more sensitive to the change in transformer cost.

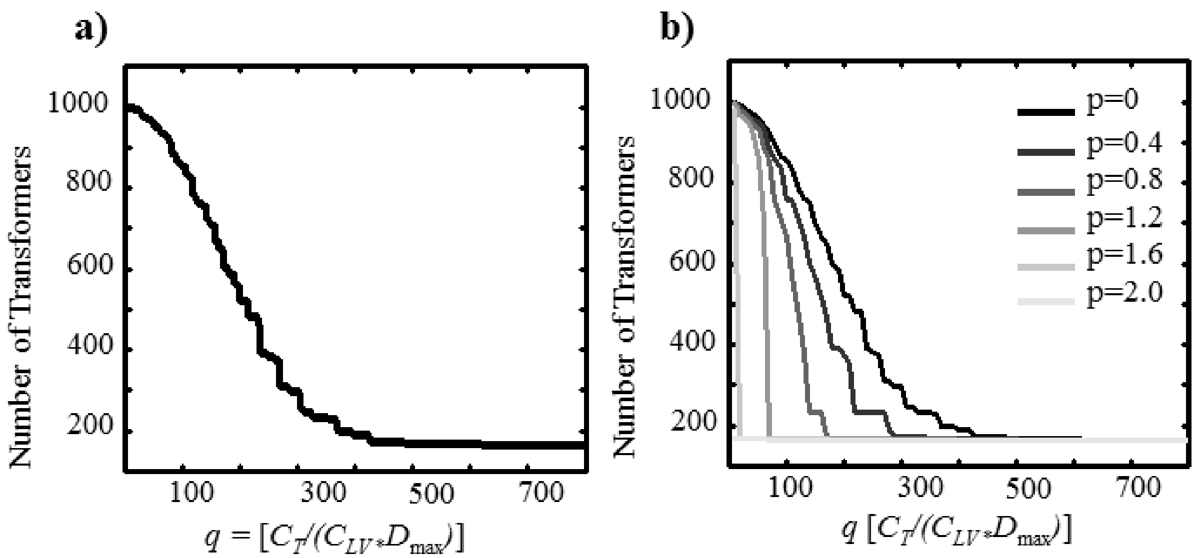


Figure 2.12 | Network sensitivity analysis. **a**, The change in the number of transformers as the C_T/C_{LV} ratio increases. **b**, The change in the number of transformers for different *p* ratios as *q* ratio increases.

2.5.1.3 Analysis with all cost parameters

In Figure 2.11c and Figure 2.12b, we looked at more realistic cases where all the cost parameters are present. Figure 2.11c shows *p* ratio analysis for different *q* values. For high *q* values, effect of S-shaped behavior of *q* dominates the sudden drop effect of *p* and we observe higher critical *p* values for small *q* values. In Figure 2.12b, *q* ratio analysis with different *p* values is presented and for high *p* ratios the sudden drop effect of *p* ratio dominates the smooth decrease effect of *q* ratio.

In conclusion, depending on the price changes the design with minimum cost may change drastically. For the values (p = 2.5, q = 1) that we use for calculating the grid in the nine Sub-Saharan Africa sites, network generated by our algorithm is not sensitive. However, policy makers can use our algorithm as a tool to understand whether or not there is a possible design change with future prices.

2.5.2 Testing of Algorithm on Simulated Data

Our results for African sites are specific to each site because all sites have different number of demand points and spatial distribution patterns. The noise and the variety in the data from Sub-Saharan Africa sites can make it difficult to draw conclusions from the results of the algorithm. This is why we test our algorithm on simulated data which provide more intuitive sense of what results would be. Same cost parameters and constraints are used for the base case run of simulated data. (C_{LV}=10, C_{MV}=25, C_T=5000, D_{max}=500, L_{max}=600)

2.5.2.1 Multivariate Normally Distributed Randomly Generated Data

We present six 10 X 10km^2 artificial sites with 1000 demand points that are randomly generated using multivariate normal distribution (generalization of Gaussian distribution in two dimensions). To filter out the noise, we stretch out the points in two dimensions by increasing standard deviations of the multivariate normal distribution (from 250 to 1500). The generated sites are shown in Figure 2.13a-c, e-g, and the results are presented in Figure 2.13d, h. As expected, the number of transformers goes up with higher standard deviation (Figure 2.13d) and as the number of transformers increases, the total MV line used increases, while the total LV decreases consistently (Figure 2.13h).

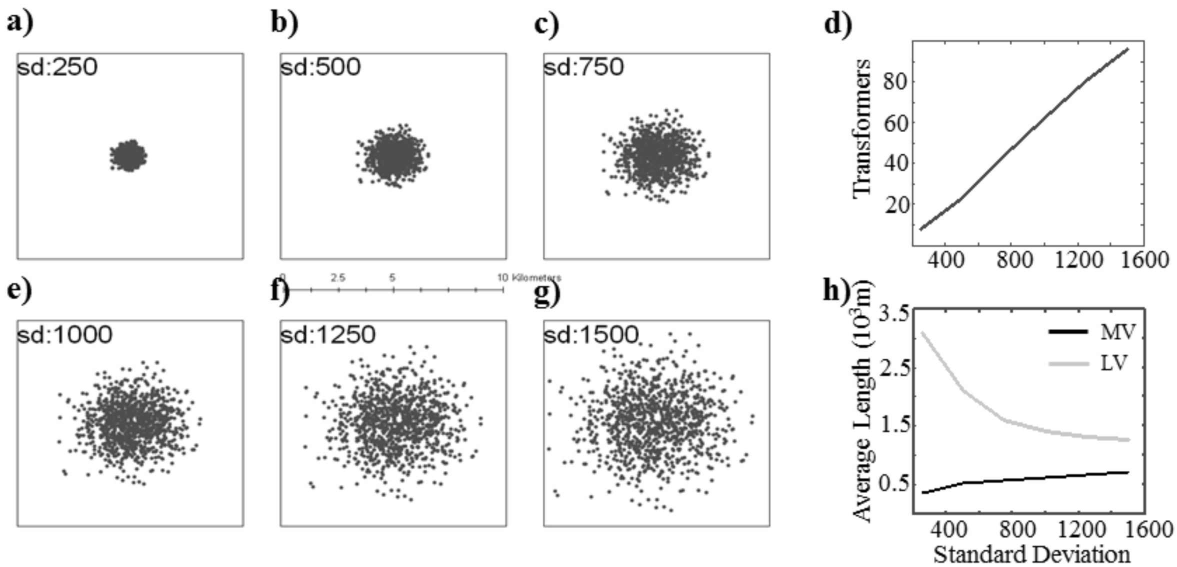


Figure 2.13 | Multivariate normally distributed random data and results. **a-c, e-g**, Multivariate normally distributed 10X10 km^2 sites with 1000 demand points. **d**, The behavior of algorithm as the standard deviation increases. **h**, Average LV line per transformer declines as the average MV line and the number of transformer increases.

2.5.2.2 Uniformly Distributed Randomly Generated Data

Next, we present 6 different sized areas (4X4, 6X6, 8X8, 10X10, 12X12, 14X14 km^2) with again 1000 demand points that are generated randomly using uniform distribution (i.e. the likelihood of generating a demand point on any point of the site is same). Since these sites have different areas but the same number of points the mean distance between demand points is the highest for 14X14 km^2 site and decreases as the site area gets smaller. The generated sites are presented in Figure 2.14 and the results are summarized in Figure 2.15.

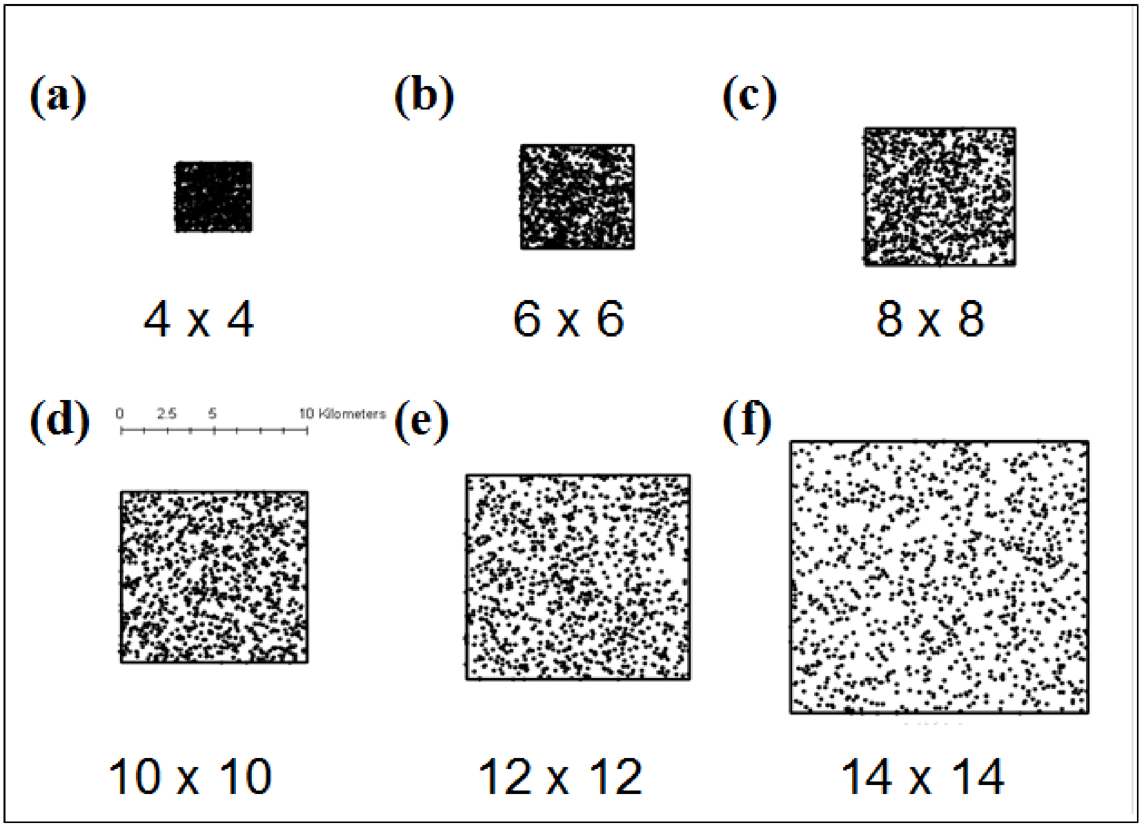


Figure 2.14 | Uniformly distributed randomly generated data on different sized areas.

To facilitate the quantitative comparison between sites, we refer to "average nearest distance" defined by Clark and Evans [44]. We calculate the average nearest distance using a Geographic Information System (GIS) tool for each different sized site and they are shown in the x-axis of Figure 2.15a-b (60 for the densest site (4X4 km^2), around 260 for the largest site (14X14 km^2)). Due to the lower number of demand points within the radius of D_{max}, number of transformers increases almost linearly (Figure 2.15a) as the average nearest distance increases (same total number of demand points, same D_{max}) and when we keep the ratio between the average nearest distance and the D_{max} same, we obtain the same number of transformers for each site (Figure 2.15b).

In addition, we also test our algorithm on the same sites setting the D_{max} to various numbers between 50 and 1000 and present the results in Figure 2.15c. As expected, the number of transformers increases as the D_{max} decreases and converges to the number of demand points. This is in general more sensitive in the dispersed areas making the slope steeper (slope in the Figure 2.15a) for lower D_{max} values.

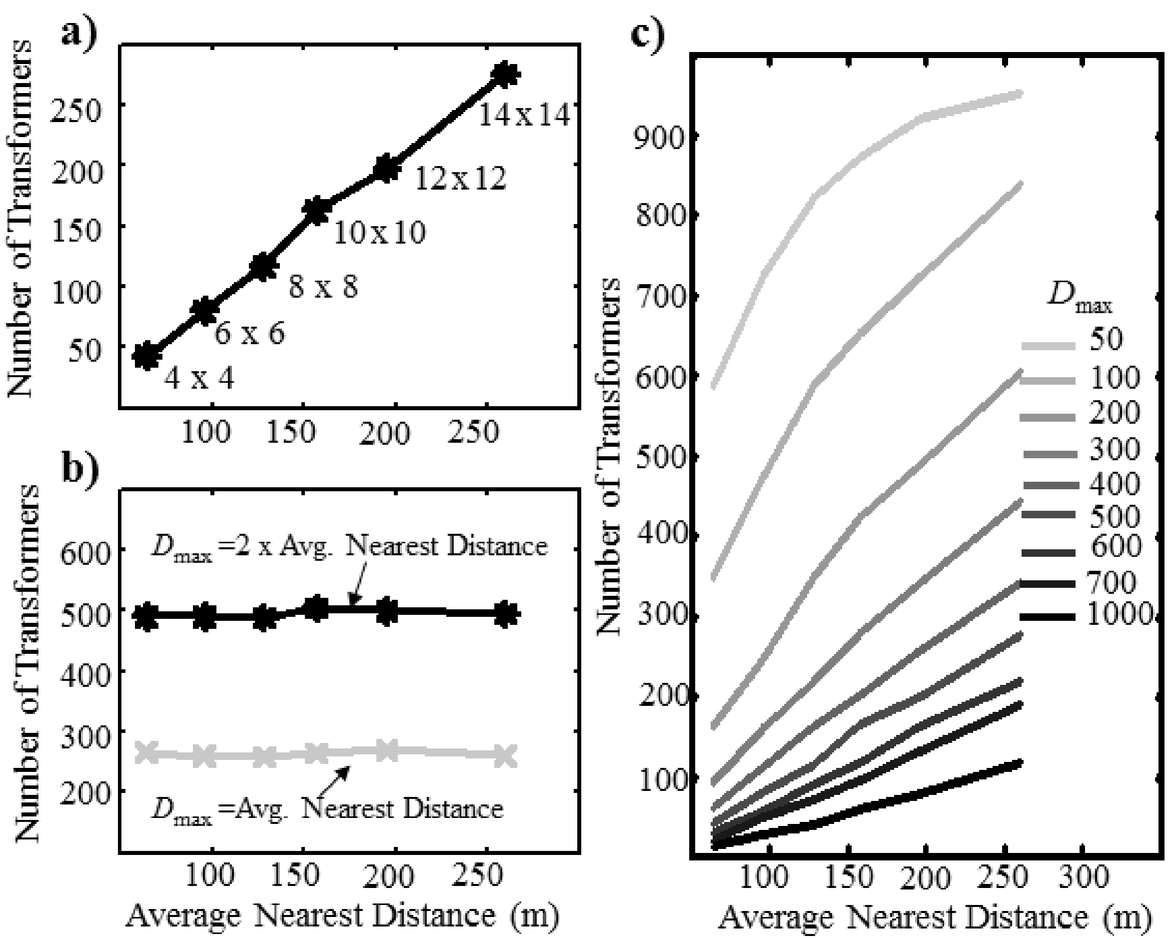


Figure 2.15 | Results for uniformly distributed random data. **a,** Number of transformers outputted by algorithm for the sites which have same number of demand points but different average nearest distances. **b,** The number of demand points when the ratio between D_{max} and Average Nearest Distance is kept constant for each site. **c,** Number of transformers versus average nearest distance for different values of D_{max} constraint.

2.5.3. Comparison of Our Algorithm with a Sequential Approach

As it was discussed in Section 2.1.2, to our knowledge, none of the previous studies in rural electrification and power engineering literature exactly matches with the objective of our paper. Complexity of the existing models [14, 15, 19] limits their suitability for the large data sets of demand points (up to 6500 households per site in our case). However; we can still compare our results with a relatively simpler sequential approach. In this approach, problem is divided into three sub-problems: transformer location problem, MV network design problem and LV network design problem. Then, each sub-problem is solved sequentially.

In the sequential approach, a greedy approach proposed for set covering problem[5] by Chavatal [45] is implemented to solve transformer location problem. Chavatal proves that the cost returned by the heuristic algorithm is at most *H(d)* times of the cost of an optimal solution where $H(d) = \sum_{i=1}^{d} 1/i$ and d is the size of the largest set found by the algorithm. Once the locations of the transformer and their service areas are known, the MV and LV networks are found using MST and CMST algorithms, respectively. Then, total cost of transformers and cost of networks are calculated.

Final number of transformers and total cost results for both sequential approach and our algorithm are presented in Table 2.3. It is shown that our algorithm tends to perform better than the sequential approach in terms of the total cost providing 4.5% improvement across all sites. We note that both approaches discussed here are based on polynomial time algorithms and provide reasonably good solutions to an NP-hard problem.

[5]Set covering problem aims to find the minimum number of sets subject to the constraint that each demand point should be covered (served) by a facility within a certain coverage criterion.

Table 2.3 | Comparison of our algorithm with a sequential approach

	Number of Demand Points	Number of Transformers		Total Cost ($)		
		Our Algorithm	Sequential Approach	Our Algorithm	Sequential Approach	Difference
Potou, SENEGAL	1797	71	70	2,433,520	2,534,654	-4.2%
Mbola, TANZANIA	1175	90	96	3,032,250	3,180,655	-4.9%
Tiby, MALI	2496	32	31	1,724,260	1,716,742	0.4%
Ruhiira, UGANDA	6434	212	199	7,680,315	8,323,259	-8.4%
Bonsaaso, GHANA	993	18	16	1,007,715	996,358	1.1%
Ikaram, NIGERIA	1484	33	34	1,468,470	1,546,644	-5.3%
Mayange, RWANDA	3909	114	107	4,148,520	4,327,479	-4.3%
Pampaida, NIGERIA	1570	88	98	2,968,030	3,143,463	-5.9%
Mwandama, MALAWI	4230	152	154	5,335,235	5,832,993	-9.3%

2.5.4. Limitations of Our Algorithm and Possible Extensions

The model described in this paper intended for obtaining quick estimates for the network structure and associated costs as a part of feasibility analysis, rather than being used for detailed implementation. Below we list a set of simplifications that we adopted:

- Our present model does not take some concepts into consideration such as power flow, power loss, voltage regulations, and transformer sizes for simplicity purposes.
- We use constant cost parameters throughout the entire system for transformers and wires. Thus, we limit our model to have single type of transformer (25kVA) and single LV and MV technology (three-phase).

- We assume that our demand points are distributed with the similar loads and spatial characteristics and do not model bulk MV and non-homogenous loads.
- Upfront capital costs dominate the operations and maintenance (O&M) costs and are treated as “overnight” costs (i.e. it is assumed that the entire system investment is made at once).
- We do not include in our model network control devices such as voltage regulators, switches etc.

Our intent in this paper is to keep the model as simple as possible with the assumptions above while focusing on designing two-level network such that the overall distribution system is optimized in one framework as suggested by [46]. A multi-level network design problem which includes multi-point network configuration in one level has enough complexity; however our model can still be extended to include some of the important concepts mentioned above.

For example, based on the number of demand points served by a transformer, transformer size can also be determined by the model and different cost parameters can be used to calculate the total transformer costs. For example, instead of using 25kVA for $5000 each, we can prefer assigning cheaper 16kVA transformers to the clusters which serve small number of points in sparse areas. Thus, we could decrease the total cost and avoid underutilized transformers.

Furthermore, it is also possible to put a capacity constraint (to incorporate the power losses and voltage regulations) in MV network even though this would make the problem even more complicated. Using the source as root, we can use CMST algorithm to design the MV level instead of using MST. For the LV level, L_{max} can be determined from the power loss and voltage drop constraints. Based on the number of transformers and the total length of the network, number of network control devices can be estimated and the cost of these devices can also be included in the objective function (total cost) of the system.

Another modification in the algorithm that could be done is that instead of using, multi-point network configuration in the LV level, star configuration can also be employed as it might be preferable for some situations. In this case, at each step of the algorithm CMST algorithm will be skipped and total cost of LV network will be found after calculating the direct distance between transformers and their associated demand points.

It is also possible to introduce time in our model with an assumption on the lifetime of a grid. For depreciation purposes, the lifetime of a grid is considered between 20 and 30 years but in reality it is usually more with a proper maintenance [35]. It is also acceptable to use a number between 1/8 and 1/30 of the capital cost for an estimate on the O&M costs in annual basis [35]. Thus, given annual costs and life time the grid, O&M costs can be discounted to the present value and be included in our objective function.

Furthermore, another potential application of our algorithm would be in facility location problem where given a set of household locations, planners are interested in finding how many schools or health facilities they need [44-48]. The unique situations of rural areas, in particular the sites in Sub-Saharan Africa as explained in the introduction; prevent many of the existing algorithms from being applicable as they usually require a set of candidate facilities as an input and there is no way to refine the candidate locations on the two dimensional coordinate system of the ground in these sites. By removing the cost of the network from the objective function, our algorithm can easily be modified for this purpose. This will simplify our algorithm to an agglomerative clustering algorithm that minimizes the total cost of opening facilities subject to the D_{max} constraint, which specifies the maximum distance between each household and the facilities (discussed in Section 4.1 in detail).

2.6 Conclusion

With the local-level household data developed from satellite imagery, we look at how the settlement patterns of households (demand points) affect the cost of power distribution systems. Rural energy planning has mainly focused primarily on the national to regional scale. This study is one of the first examples which focus on the local level analysis and help energy planners make the rapid assessment of distribution systems planning and partial electrification in rural settings. In the first part of this chapter (single level electrification), we show that some population settlement patterns offer the potential for savings (on a per-unit basis) in upfront investments through an initial roll-out that covers part of the population; later expansion of these partially spanning networks can be undertaken at little additional cost in the long-term.

In the second part (multi-level), a new heuristic algorithm for the design of two-level power distribution systems has been introduced. It has been presented that the algorithm finds the number and locations of MV/LV transformers without giving any candidate locations and finds a multi-point low voltage network between demand points and transformers. The proposed algorithm ignores transmission losses, load flow considerations and local topography. Hence it should be viewed as a quick tool which simplifies a complex problem and provides good starting point for decision makers and practitioners. However; our algorithm is flexible such that it can be simplified to other infrastructure problems (for example; facility location problem) or it can be extended to include more distribution system components such as transformer sizes.

Bibliography

[1] Modi, V., McDade, S., Lallement, D., & Saghir, J. (2005). Energy services for the Millennium Development Goals. *Energy services for the Millennium Development Goals*.

[2] WEC and FAO. (1999).The challenge of rural energy poverty in developing countries. London: World Energy Council.

[3] Rural Energy and Development: improving energy supplies for two billion people. World Bank, 1996.

[4] Silva Herran, D., & Nakata, T. (2012). Design of decentralized energy systems for rural electrification in developing countries considering regional disparity.*Applied Energy*, *91*(1), 130-145.

[5] Bhandari, R., & Stadler, I. (2011). Electrification using solar photovoltaic systems in Nepal. *Applied Energy*, *88*(2), 458-465.

[6] Saheb-Koussa, D., Haddadi, M., & Belhamel, M. (2009). Economic and technical study of a hybrid system (wind–photovoltaic–diesel) for rural electrification in Algeria. *Applied Energy*, *86*(7), 1024-1030. [7] Urban, F., Benders, R. M., & Moll, H. C. (2009). Energy for rural India. *Applied Energy*, *86*, S47-S57.

[8] Herran, D. S., & Nakata, T. (2008). Renewable technologies for rural electrification in Colombia: a multiple objective approach. *International Journal of Energy Sector Management*, *2*(1), 139-154.

[9] Kanniappan, P., & Ramachandran, T. (2000). Goal programming model for sustainable electricity production from biomass. *International journal of energy research*, *24*(1), 1-18.

[10] Bekker, B., Eberhard, A., Gaunt, T., & Marquard, A. (2008). South Africa's rapid electrification programme: Policy, institutional, planning, financing and technical innovations. Energy Policy, 36(8), 3125-3137.

[11] Haanyika, C. M. (2006). Rural electrification policy and institutional linkages.Energy Policy, 34(17), 2977-2993.

[12] Hsu, Y., & Jwo-Hwu, Y. (1996). Planning of distribution substations, feeders and sectionalizing switches using heuristic algorithms. *International Journal of Electrical Power & Energy Systems*, *18*(5), 315-322.

[13] Hongwei, D., Yixin, Y., Chunhua, H., Chengshan, W., Shaoyun, G., Jian, X., ... & Rui, X. (1993, October). Optimal planning of distribution substation locations and sizes-model and algorithm. In *TENCON'93. Proceedings. Computer, Communication, Control and Power Engineering. 1993 IEEE Region 10 Conference on* (pp. 351-354). IEEE.

[14] Hamam, Y. M., & Hindi, K. S. (1987). Optimized design of low-voltage distribution networks a comprehensive algorithm. *Systems, Man and Cybernetics, IEEE Transactions on*, *17*(3), 502-507.

[15] Aoki, K., Nara, K., Satoh, T., Kitagawa, M., & Yamanaka, K. (1990). New approximate optimization method for distribution system planning. *Power Systems, IEEE Transactions on*, *5*(1), 126-132.

[16] Ramezani, M., Falaghi, H., Moghaddam, M. P., & Haghifam, M. R. (2006). Genetic based algorithm for optimal placement of distribution transformers. In*Power Engineering Society General Meeting, 2006. IEEE* (pp. 5-pp). IEEE.

[17] Khator, S. K., & Leung, L. C. (1997). Power distribution planning: a review of models and issues. *Power Systems, IEEE Transactions on*, *12*(3), 1151-1159.

[18] Gonen, T., Mahmoud, A. A., & Calborn, H. W. (1983). Bibliography of Distribution System Planning. *Power Engineering Review, IEEE*, (6), 53-53.

[19] Gonen, T., & Ramirez-Rosado, I. J. (1987). Optimal multi-stage planning of power distribution systems. *Power Delivery, IEEE Transactions on*, *2*(2), 512-519.

[20] TURKAY, B., & ARTAC, T. (2005). Optimal Distribution Network Design Using Genetic Algorithms. *Electric Power Components and Systems*, *33*(5), 513-524.

[21] Miranda, V., Ranito, J. V., & Proenca, L. M. (1994). Genetic algorithms in optimal multistage distribution network planning. *Power Systems, IEEE Transactions on*, *9*(4), 1927-1933.

[22] Zvoleff, A., Kocaman, A. S., Huh, W. T., & Modi, V. (2009). The impact of geography on energy infrastructure costs. *Energy Policy*, *37*(10), 4066-4078.

[23] ESMAP (Energy Sector Management Assistance Program, World Bank). (2007).Technical and Economic Assessment of off-grid, Minigrid and Grid Electrification Technologies. Technical Paper ;121/07. December. World Bank, Washington, D.C.

[24] Prim, R. C. (1957). Shortest connection networks and some generalizations.*Bell system technical journal*, *36*(6), 1389-1401.

[25] Kruskal, J. B. (1956). On the shortest spanning subtree of a graph and the traveling salesman problem. *Proceedings of the American Mathematical society*, *7*(1), 48-50.

[26] Boruvka, O. (1926). Prıspevek k rešenı otázky ekonomické stavby elektrovodnıch sıtı (contribution to the solution of a problem of economical construction of electrical networks). *Elektronický obzor*, *15*, 153-154.

[27] Skiena, S. (1998). *The Algorithm Design Manual: Text* (Vol. 1). Springer.

[28] Fischetti, M., Hamacher, H. W., Jørnsten, K., & Maffioli, F. (1994). Weighted k-cardinality trees: Complexity and polyhedral structure. *Networks*, *24*(1), 11-21.

[29] Blum, A., Ravi, R., & Vempala, S. (1996, July). A constant-factor approximation algorithm for the k MST problem. In *Proceedings of the twenty-eighth annual ACM symposium on Theory of computing* (pp. 442-448). ACM.

[30] Garg, N. (1996, October). A 3-approximation for the minimum tree spanning k vertices. In *Foundations of Computer Science, 1996. Proceedings., 37th Annual Symposium on* (pp. 302-309). IEEE.

[31] Arya, S., & Ramesh, H. (1998). A 2.5-factor approximation algorithm for the k-MST problem. *Information Processing Letters*, *65*(3), 117-118.

[32] Arora, S., & Karakostas, G. (2000, February). A 2+ ε approximation algorithm for the k-MST problem. In *Proceedings of the eleventh annual ACM-SIAM symposium on Discrete algorithms* (pp. 754-759). Society for Industrial and Applied Mathematics.

[33] Chudak, F. A., Roughgarden, T., & Williamson, D. P. (2001). Approximate k-MSTs and k-Steiner trees via the primal-dual method and Lagrangean relaxation. In *Integer Programming and Combinatorial Optimization* (pp. 60-70). Springer Berlin Heidelberg.

[34] Goemans, M. X., & Williamson, D. P. (1995). A general approximation technique for constrained forest problems. *SIAM Journal on Computing*, *24*(2), 296-317.

[35] Willis, H. L. (Ed.). (2000). *Distributed power generation: planning and evaluation*. CRC Press.

[36] Johnson, R. A., & Wichern, D. W. (2002). *Applied multivariate statistical analysis* (Vol. 5, p. 767). Upper Saddle River, NJ: Prentice hall.

[37] Liao, K., & Guo, D. (2008). A Clustering-Based Approach to the Capacitated Facility Location Problem1. *Transactions in GIS*, *12*(3), 323-339.

[38] Min, H. (1996). Consolidation terminal location-allocation and consolidated routing problems. *Journal of Business Logistics*.

[39] Barreto, S., Ferreira, C., Paixao, J., & Santos, B. S. (2007). Using clustering analysis in a capacitated location-routing problem. *European Journal of Operational Research*, *179*(3), 968-977.

[40] Mittenthal, J., & Gray, B. (2009). The impact of stopping rules on hierarchical capacitated clustering in location routing problems. *Academy of Information and Management Sciences Journal*, *12*(1), 13.

[41] Essau, L. R., Williams, K. C. (1966). A method for approximating the optimal network. *IBM Systems Journal*, 5(3), 142-147.

[42] Sharma, R. L., & El-Bardai, M. T. (1970, June). Suboptimal communications network synthesis. In *Proceedings of the IEEE International Conference on Communications* (Vol. 19, pp. 11-16).

[43] Chandy, K.M., Russell R.A.(1972). The design of multi-point linkages in a teleprocessing tree network. *IEEE Transactions on Computers* (Vol. 21, pp. 1062-1066).

[44] Clark, P. J., & Evans, F. C. (1954). Distance to nearest neighbor as a measure of spatial relationships in populations. *Ecology*, *35*(4), 445-453.

[45] Chvatal, V. (1979). A greedy heuristic for the set-covering problem.*Mathematics of operations research*, *4*(3), 233-235.

[46] Carter-Brown, C. G., & Gaunt, C. T. (2002, October). Model for voltage drop apportionment in MV/LV rural electricity distribution. In *Africon Conference in Africa, 2002. IEEE AFRICON. 6th* (Vol. 2, pp. 929-932). IEEE.

[44] Mahdian, M., Ye, Y., & Zhang, J. (2002). Improved approximation algorithms for metric facility location problems. In *Approximation algorithms for combinatorial optimization* (pp. 229-242). Springer Berlin Heidelberg.

[45] Chudak, F. A., & Williamson, D. P. (1999). Improved approximation algorithms for capacitated facility location problems. In *Integer programming and combinatorial optimization* (pp. 99-113). Springer Berlin Heidelberg.

[46] Chudak, F. A., & Williamson, D. P. (1999). Improved approximation algorithms for capacitated facility location problems. In *Integer programming and combinatorial optimization* (pp. 99-113). Springer Berlin Heidelberg.

[47] Barcelo, J., Fernandez, E., & Jörnsten, K. O. (1991). Computational results from a new Lagrangean relaxation algorithm for the capacitated plant location problem. *European Journal of Operational Research*, *53*(1), 38-45.

[48] Melkote, S., & Daskin, M. S. (2001). Capacitated facility location/network design problems. European Journal of Operational Research, 129(3), 481-495.

[49] Kocaman, A. S., Huh, W. T., & Modi, V. (2012). Initial layout of power distribution systems for rural electrification: A heuristic algorithm for multilevel network design. Applied Energy, 96, 302-315.

Chapter 3

Modelling of a Hybrid Energy System with Conventional Hydro Storage

3.1 Introduction

The importance of sustainable energy planning has increased substantially with rising population growth rates, environmental issues and economic developments. International Energy Agency (IEA) estimated that primary sources of electricity in 2009 consisted of 40.6% coal, 21.4% natural gas and 5.1% petroleum summing up to a 67.1% share for fossil fuels in primary electricity consumption in the world [1]. However, fossil fuels are finite and their combustion results in greenhouse gas emissions, which contribute to global warming and health hazards. Therefore, energy models that involve clean and renewable energy sources are necessitated to ease the concerns on the electricity generation that meets the projected demand.

Transition to alternative renewable energy sources is inevitable. However, renewable sources are generally are variable and heavily dependent on the spatial location (e.g. sunshine while more predictable, is limited to daytime hours, and the total annual insolation is also spatially varying. Annual wind energy potential is even more spatially heterogeneous.). Thus if a future energy system wants to predominantly rely on these sources, it must utilize a mix of variable and dispatch-able resources that are interconnected, thus requiring investments in transmission, or utilize back-up dispatch-able resources (likely to be fossil fuels or hydro in the near term), or utilize some form of storage (e.g. pumped hydro or compressed air energy storage), or allow some of the energy generated to be curtailed or use intelligent demand side management. For cost-effectiveness of the overall system, the approach is likely to be “all of the above” [2-4]. Here, for the sake of demonstration we imagine a long-term scenario that primarily relies on solar and hydro as the renewable resources and assume that fossil fuels will be expensive and hence judiciously used.

We imagine the demand profiles of a specific country, and model the long-term investments and storage while capturing the volatility of hourly supply and demand. In this country, we first identify candidate basins for hydro power stations and aggregated demand point locations (the cities or the states of the country). We assume that the first use of solar will be for immediate in time as well as locally. Then, we determine the possible transmission network between supply and demand points. We mathematically model this hybrid energy generation and allocation system, where time variability of energy sources and demand is balanced with the water stored in the reservoirs of hydropower stations.

Hydropower plants are known as reliable parts of power systems due to their inherent instantaneous starting, stopping and load variation ability [5] and can be designed with or without storage. A plant without storage is called *run-of-the-river* system and produces electricity by diverting river flow through turbines that spin generators without materially altering the normal course of the river. Hydropower plants with storage can be either a conventional plant where incoming stream flows are stored in large reservoirs in dams and water release can be varied or deferred as per need or a pumped hydro plant where water can also be moved between a lower and a upper reservoir for later use. In this chapter, we present a model for a hybrid system which includes conventional hydropower stations and decentralized solar power stations working together to meet the variable demand. To increase the reliability of the system, diesel generators (as a proxy for expensive fossil resources) are used as a backup source where we ideally use only when there is no hydro and solar power production possible due to the intermittency of renewable sources. The objective is to find the least-cost design for the power stations and transmission lines, which also minimizes the diesel usage. Then, in the next chapter, we present a pumped hydro model where excess electricity generated from solar energy can be transmitted to hydropower stations and used to pump water to the upper reservoir to store energy in the form of potential energy.

The main motivation of the model is to determine the optimal capacities of infrastructure needed to match projected demand and supply in the most cost effective way. We answer questions such as; -where to locate hydro and solar power stations, -where, when and how much solar and hydro energy should be produced during each time interval, -what should be the size of reservoirs so that the demand and supply would be matched, –from/to where and how much energy should be transmitted? In this section, we also provide results for alternative hybrid system designs such as solar-diesel, run of the river-diesel, solar-conventional hydro-diesel and show the quantitative importance of storage while using renewable sources.

An innovative contribution of this work is the establishment of a new perspective to energy modeling by including fine-grained sources of variability such as stream flow, solar radiation in hourly level as well as spatial location of supply and demand in the national/regional level. The model is formulated as a stochastic linear program. Stochastic nature of inputs such as stream flow is addressed by determining scenarios from the time series. The model will be presented with a case study of India and helps answer whether solar energy in addition to high hydro power potential in the Himalayas would provide a back-bone for a low carbon economy in the face of growing electricity demand if fossil fuels could be almost completely phased out from electricity generation.

3.2 Background

3.2.1 Literature Review

This study is relevant to some well-studied problems in the literature such as planning of hybrid energy systems, long-term energy investment planning problems, and electricity generation expansion problem.

The idea of a hybrid system is to obtain the most cost efficient system using alternative sources. In order to obtain electricity from a hybrid system reliably and at an economical price, it must be designed optimally in terms of operation and component selection. Many different hybrid systems which has been proposed in literature, involves renewable resources such as solar photovoltaic, wind and hydro with or without existence of storage alternatives such as pumped hydro or batteries [2-4]. Mathematical modeling and optimization of hybrid systems is not an easy task as they usually involve many components and decision variables. Especially in the existence of storage, the fact that all time units in the planning horizon are linked to each other complicates the solution of the model. Therefore, hybrid systems have generally been proposed more for localized and decentralized systems without including transmission part of the power systems to reduce the complexity of the models. However, there is a need for feasibility studies in the literature which help understand contribution of the renewable sources in national energy system planning.

In the macro level, several national level energy planning models have been proposed [6-11]. These models provide policy makers with extensive details on energy generations and consumption technologies and how to meet some of the long-term goals related to government policies such as phasing out fossil fuels or decreasing greenhouse gas emissions. Previously proposed studies (except for [6]) include time component in their model with increments from 1 to 5 years and use the average values for energy sources and demand. However, it has been shown that models that utilize intermittent sources such as solar and wind tend to understate their value when averages are used [12]. These sources look more valuable when production periods are set to as short as a few hours. In addition, all these models use aggregated supply and demand without explicitly representing spatial locations. Therefore it is not possible to answer specific investment questions such as where to locate solar power station or how to expand the transmission lines.

3.2.2 Background for India Case Study

India, with 1.27 billion people, is the second most populous country in the world as of 2013. According to World Energy Outlook, 2012 version from IEA, India is expected to overtake China soon after 2025 becoming the most populous country and its population will exceed 1.5 billion in 2035 [13]. In India, nearly 25 percent of the population lacks basic access to electricity and electrified areas suffer from electricity blackouts [14]. Moreover, India heavily depends on coal for meeting its current energy demand (with a share of 42% in 2009). It is currently the third-largest generator of coal-fired power after China and United States and estimated to overtake United States to become the second-largest by the end of 2025 [14]. Therefore, increasing rate of energy consumption, heavy dependence on petroleum fuels and volatility of world oil market increase the importance of clean energy sources in order to be able to balance the need for electricity and address the environmental concerns for sustainable development in India. Furthermore, there is also additional factors such as global pressure, voluntary targets for greenhouse gas emission reduction and intensification of rural electrification program that promote the use of renewable sources in the country [15].

India, with a vast land area, is very rich in terms of renewable energy sources like solar, hydro, wind and biomass [16]. Given the Himalayan ranges in the north having numerous rivers and streams with perennial flows; one of the biggest renewable potential in India is the hydropower. Stream flow in the Himalayan Rivers is generated from rainfall and concentrated snow. The flows are observed throughout the years and steep geographical slopes make all the streams in Himalayans high potential sites for hydropower generation [17]. As the potential is quite significant, both large-scale and small-scale hydropower stations can be considered for installation. Our goal in this chapter is to evaluate both alternatives. In national level, large power plants seem to be more cost effective as they include "water storage (reservoirs)" instead of substantially expensive "energy storage" (batteries) and small-scale plants (e.g. run-off-the river stations) is environmentally friendly and sufficiently meets the need in smaller demand areas.

Moreover, India lies in the sunny belt of the world and a very promising place for solar energy generation. The average intensity of solar radiation received in India is 200 MW/km^2 with 250–300 sunny days in one year [18]. With the increasing urbanization rate, cooling demand also increases in India and the same trend in solar radiation and cooling demand makes solar energy utterly suitable source to meet the peak demand caused by the air conditioners. In our model, we consider the energy generated from solar power as a part of the nation-wide energy planning optimization processes and find the optimized solar panel area for each demand point using the hourly solar radiation data.

In addition to hydro potential and solar energy, wind power and biomass are the other clean energy resource currently being exploited in India [19]. Wind power is widely distributed energy resource and has its own unique advantages especially for the developing world such as being able be installed quickly in areas where electricity is urgently needed. Total wind energy potential in India has been assessed as 45 GW in 2008 [20] and as of September 2013 19.8 GW of these capacity has been installed [21]. In our model, we include neither wind energy nor biomass in our problem; however, once the data is available, it is fairly easy to extend the model to include more renewable sources as can be seen below.

3.3 Problem Statement

Dispatch-able power is the power that can be called on when needed, in contrast to base load power, which is essentially always on. Dispatch-able generation is a premium power source since it is controllable and not wasted. For example; when air conditioners are off and refrigerators are not running, there is no need for power inflow and in this case, having controllable supply power saves energy. In a carbon-constrained world, we want to use as much renewable sources as possible; however, because of the intermittency of renewable sources, they are mostly non-dispatch-able. Hybrid power systems, energy storage, long distance transmission and demand response programs can help reduce intermittency of the renewables and allow the grid accommodate more variation on both supply and demand. In this problem, our goal is to see how combining multiple renewable sources which have different variability, storage and transmission can help reduce the intermittency and variability of sources and increase the reliability of the power systems.

Our model is designed to help infrastructure planners make long-term investment decisions based on the results for electric dispatch, energy resource allocation and storage over one-year horizon. The objective of the model is to minimize the sum of the investment costs and expected penalty cost for the demand which can not be met by renewable sources. Given demand points and candidate basin locations for hydropower stations, we are interested in optimal sizing of power stations (hydropower and solar) and transmission lines (between basins and demand points) while minimizing the infrastructure cost of solar and hydropower stations and cost of the back-up energy source (diesel generators) that we use in case there is no stream flow or solar radiation available. In this system, water stored in the reservoirs can mitigate volatility of supply and demand. Water release from the reservoirs can be controlled and deferred until it is needed. Thus, reservoirs facilitate energy transfer from low use periods to peak use periods, allowing the system operate based on demand load while maintaining high system reliability.

Storage is the key enabling technology for intermittent energy; however, it complicates the design of optimization problems by coupling all the time periods together. While working with sources that are not constantly available such as solar, the time increment that we use in the optimization model becomes quite important. As it will be discussed in Section 3.3.2.1, stream flow data in Himalayas, solar radiations as well as demand show both seasonal and diurnal variability. To accurately capture the diurnal variability, it is necessary to model energy supply and demand in hourly time increments. Moreover, because of the seasonal variability of the sources, it is also crucially important to use at least one year as time horizon. An approach that avoids capturing every time increment over a year by simply sampling different time periods (e.g. different time of the years and time of the days) fails to accurately model the storage. Moreover, modeling reservoir systems is principally more complicated than modeling other storage types such as a battery which stores energy during the day and releases at night. Here, we may put water in reservoir storage in September so it can be used in dry seasons, couple of months in the future.

The nature of hydropower generation, storage and the stochastic aspect of the key variables like stream flow, solar radiation and demand make optimization problems quite difficult to be dynamically solved as high number of units is involved. Here, following the typical strategy in stochastic programming, we present a scenario based static model with multiple time periods and time periods are coupled by storage. Possible random situations are represented by scenarios with the associated probabilities. By scenario approach, a set of prototype 1-year series with 3 hourly time increments are selected from the time series as a particular realization of the uncertain data. A drawback of a scenario-based approach is the fact that scenarios are generated in advance, and this limits their ability to capture the interaction between decisions and exogenous events. We assume that effect of this drawback can be minimized during the real-time operations of the power systems. For example, in case of very rainy season which is not foreseen and captured by scenarios, the water in the reservoirs can be controlled to be prepared for the season.

3.3.1 Hybrid System Components

The hybrid model described in this chapter has three sub-systems; hydropower stations, solar power stations and transmission network between hydro and solar power stations. Design of individual power systems is not in the scope of this problem so several assumptions are made to reduce the complexity of the model.

3.3.1.1 Hydropower Systems

Here, we identify several basins as candidate locations for hydropower stations and optimization model is solved for the size of the reservoirs and generators to determine the type and capacity of the power systems. If the generator size is zero for a candidate basin, there is no need for a hydropower station at that candidate basin. Moreover, when generator size is positive whereas the reservoir size is zero, then it means a run-of-the-river system is the optimal choice of the basin. It should be noted that these decisions are based only on supply and demand data. Other site-specific concerns such as effects of dams on fishing, recreational activities and tourism or environmental constraints are not taken into account.

In a more detailed hydropower plant design problem, diameter of the pipe from the reservoir to turbines would be another constraint to consider. Here, we assume that pipeline sizes are proportional to reservoir size and cost can be included in reservoir cost. Moreover, empirical information shows that there is no operational cost based on the output level of the hydropower station [22]. Water represents the only variable that could be in the form of opportunity cost and is not taken into account in our model. For our case studies, we use $1/m^3 as unit cost of reservoir capacity (i.e. constant incremental cost of installing reservoir capacity) which is within the ranges given in [38] and [39]. For the unit cost of powerhouse (generator, turbine/pump, transformers), we use $500/kW as this value is the lower bound of the range given in [40] for the capital cost of the medium to large hydropower stations (assuming that the lowest cost system has no storage). The cost parameters can clearly vary from site to site and this can easily be incorporated in the model assigning site specific unit cost for each investment variable.

In this model, although we are interested in finding the size of multiple reservoirs, we assume that each hydropower plant operates independently and each plant is assigned to one reservoir. Due to frictions in the tunnel, turbines and generators, 12-14% of the potential energy of the water can be lost while generating electricity [22]. Therefore, we use 88% efficiency for all plants here. Losses due to evaporation from the reservoirs are ignored.

The potential for power production at a reservoir site mainly depends on the flow rate of water that can pass through generation turbines and the potential head available. Potential head usually depends on the topology and the constructed wall of the dam because based on the design of the dams; the water level stored in the reservoirs can have an important influence on the energy potential of water. For example; in Norwegian statistics, the vertical height of a waterfall is measured from the intake to the turbines [22]. Given the steep slopes of the Himalayas, it would be reasonable to use the same statistic here. Thus, we use a constant head for each reservoir during the operations and do not consider the reduced electricity conversion efficiency, which is caused by the fact that height of water falls is reduced as the reservoir is drawn. Moreover, usually reservoirs have dead storage capacities for providing a firm head for hydropower production taking care of the sedimentation requirements. Since our optimization model solves for only the active storage, a prescribed fraction of the total reservoir cost can simply be added to objective function with a minor change.

3.3.1.2 Solar Power Systems

Sun is with no doubt is the largest energy source present and the amount of energy it provides to the Earth (1.8×10^{11} MW) in one hour is more than the total energy consumed in an entire year. However, generated power from the solar sources is less than a percent of the world's power consumption [23]. Many research groups have been studying various ways to increase the energy generation efficiency and storage capability [24, 25]. The two main device types that are utilized for this purpose are photovoltaic (solar cells to generate electricity directly via the photoelectric effect) and concentrated solar power (capturing solar thermal energy for use in power producing heat processes). In both types, there are techniques developed to enhance the efficiency such as designing the materials and the systems that are used to absorb sunlight or using sun-trackers to compensate the Earth's motions keeping the best orientation relative to the sun [23-25].

Here, we use a simplistic approach and set the efficiency to 12% [25] and assume that solar power systems cost is linearly dependent on the size of the solar panels. Upfront capital costs dominate the operations and maintenance (O&M) costs and are treated as "overnight" costs (i.e. it is assumed that the entire system investment is made at once). The estimated range for overnight capital cost of a solar photovoltaic system is between \$2500/kW and \$9500/kW [40]. Assuming 200 MW/km^2 average solar radiation and 12% efficiency, a solar power station with 1 km^2 solar panel could generate 24 MW. Then, the cost of a 24 MW system with 1km^2 solar panel is estimated to be between \$60 and \$228 million. In our model, we use \$150 million as the estimate for the representative solar power system unit area (1 km^2 solar panel). Therefore, we use approximately \$6200/kW as our overnight capital cost for a solar photovoltaic system.

No battery storage introduced in the solar power stations here because the solar power system will only function in the day light and the generated power will be sold to the grid.

3.3.1.3 Transmission Network

Transmission cost in a power network usually depends on the capacity, distance from generation sites to demand points and related power losses in the lines. In our model, we use a process that allows us to have the transmission cost dependent on both the distance and the capacity of the lines. The details of the cost analysis along with the data are provided in APPENDIX A and here we will provide a short description of the process and some fundamental details.

We first calculate how the unit cost of unit power transmission varies based on the distance (further details given in the next paragraph). Next, we calculate the spherical distance between two points and multiply it with the corresponding cost multiplier to estimate the unit investment cost of the transmission line between the two points per unit power [26]. Then, we use this unit cost in the objective function of our optimization model to minimize the total investment cost. The data used in the model is compiled by Selcuk Korel, MS’2010, Columbia University, using the sources [41-43]. The way we calculate the cost per unit power per distance is as follows:

We initially determine what type of transmission lines can be used to transmit a certain amount of power by a certain length of distance using surge impedance loading factor. Then, the investment costs of alternative transmission lines are calculated. For example, in order to transmit 300 MW by 250 miles, one can choose to use 500 kV DC Bipole, 230 kV AC Double, 345 kV AC Single Circuit or 500 kV AC Single Circuit transmission lines where investment costs are estimated to be $442 million, $275.8 million, $308 million and $537.5 million respectively. We then repeat this procedure and calculate the cost of 1 MW power transmission by 1 km for 27 alternatives including various distances and for various power levels. After we have the unit cost values for all the alternatives, we then fit a curve to estimate the incremental cost per unit power per distance. In our analysis, we have determined a linear relationship between the amount of power and the cost. The slope of this relationship decreases with the distance transmitted. As a result, we have computed three cost parameters for three distance ranges (0-500 km, 500-1000 km and >1000 km) as $1.1 million, $0.8 million and $0.6 million per giga watt per kilometer, respectively.

In our model, there is a loss parameter that is proportional to the distance. Underground cable transporting and the cost for the stations have been also assumed to be proportional to the distance of the connection and included in the unit cost calculations.

Possible network flow directions from sources to demand points are prescribed with dedicated lines and designed as a point-to-point topology. Here we neither model the grid itself nor consider real power flow equations / phase angle differences and assume that power flows over lines can be independently assigned. This level of detail is only required for operational models and for certain types of regional planning models that aims to identify the bottlenecks in the grid. Our model does not have an explicit representation of the grid, but utilizes point-to-point distances to compute transportation costs and transmission losses with individual links for power station-to-state transmission. This representation of power flows, which captures point-to-point movements without explicitly modeling the grid, is a common approximation made in policy studies [6]. A more detailed discussion on transmission systems and how they can be linearly modeled can be found in [27].

3.3.2 Input Data

Aside from the physical features of the hybrid system, the model needs the following input data to determine the scenarios and calculate the estimated operation in the planning horizon; stream flow data for each candidate hydropower locations, solar radiation data and demand profile for each demand point.

3.3.2.1 Stream Flow Data

Stream flow in the Himalayan Rivers is generated from rainfall and the melting from accumulated snow-pack. The flows are observed through the all years and steep slopes make all the streams in Himalayans potential sites for hydropower generation [17].

Forecasting the inflows and capturing the structure of the processes is of vital importance to hydropower models. This issue is discussed in detail in [28]. For our model we identified several basins from [29] in Himalaya Mountains which are either proposed or under construction areas for hydropower generation. Then, 3-hourly stream flow data for the years between 1951 and 2004 for each candidate basins is obtained from the Variable Infiltration Capacity (VIC) land surface model which is a large scale hydrological model. This model can be implemented at grid cells from 1/8° to 2° latitude by longitude and with temporal resolutions from hourly to daily. For this study model is ran at 1° at 3 hourly resolutions. The details of the VIC model can be found in [30, 31]. General statistics about basins and details of the data are given in Table 3.1.

Table 3.1 | General Statistics for Basins

No	River	Project	Lat (°E)	Long (°N)	Min	Period 1951-2004 (Stream Flow m3/s) Max	Average	Standard Deviation	Coefficient of Variation
1	Bhagirathi	Tehri	30.38	78.48	2	36550	217	592	2.73
2	Pinder	Devsari	30.41	79.37	0	13734	165	230	1.39
3	Chenab	Pakal Dul	33.46	75.81	169	48334	1765	1957	1.11
4	Marusudar	Bursar	33.29	75.76	44	12749	466	516	1.11
5	Lohit	Demwe	28.03	96.45	77	18915	599	616	1.03
6	Dibang	Dibang	28.34	95.78	46	21362	284	305	1.08
7	Barak	Tipaimukh	24.23	93.02	0	46528	2086	1968	0.94
8	Siang	Siang	28.17	95.23	878	1863500	13879	23303	1.68

In India, the normal onset of Monsoon is expected to be observed around June and its withdrawal completes by around October every year. Therefore, as it can be observed from the stream flow time series data of Bhagirathi River for the years 2003 and 2004 in Figure 3.1a, stream flow data shows significant seasonal variation. There is also a substantial contribution from snowmelt runoff to the annual stream flows of the Himalayan Rivers [17]. The water yield from a high Himalayan basin is roughly twice as high as that from an equivalent basin located in the peninsular part of India. A higher water yield from the Himalayan basins is mainly due to the large inputs from the snowmelt and glaciers [32]. Most of the snow melts occur in the summer period correlated with the sun light and cause diurnal variation in the stream flow. An example to show diurnal variability is presented in Figure 3.1b.

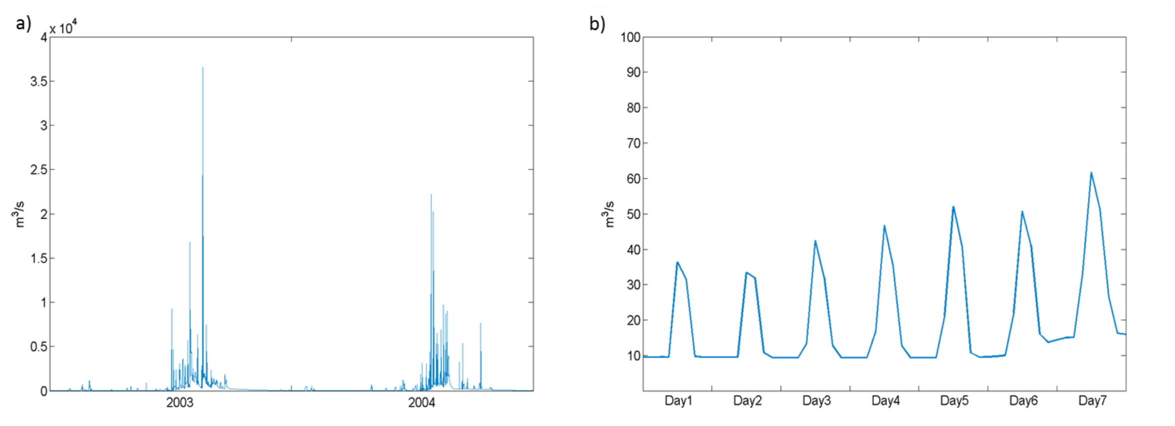


Figure 3.1 | a, Stream Flow Data of Bhagirathi River for 2003 and 2004. Monsoon is expected to be observed around June and its withdrawal completes by around October every year. **b,** Stream Flow Data of Bhagirathi River for a week in March. Most of the snow melts occur in the summer period correlated with the sun light and cause diurnal variation in the stream flow.

3.3.2.2 Demand Data

Aggregated electricity demand data in 3-hourly resolution is collected for Delhi and 7 other states which are located in northern part of India. Total monthly power availabilities and requirements for each state for the year 2012 are provided on Central Electricity Authority, Power Ministry of India (CEA) website [33]. On the website of Load Dispatch Center of Delhi, daily load profiles for the days where peak demand occurred in each month are provided. A few other states such as Chhattisgarh, Assam and Punjab also publish on their website daily, weekly or monthly reports for daily load profiles for a number of days or for the times when the minimum and maximum demands are observed. We have used the collected data to accurately estimate the 3-hourly demand load profile of each state for one year. If there is missing data for some days or hours within a day, interpolation/extrapolation methods are performed to project the data. When there is only limited number of days that we can use as a representative of all days in a month, we generated data from a normal distribution with mean equal to observed demand of given days and standard deviation equal to 5 percent of the observed data.

For the states which we do not have access to daily load profiles such as Uttar Pradesh, Bihar, Jharkhand and West Bengal, we used Chhattisgarh as a reference state. Daily load profiles of Chhattisgarh are rescaled by the ratio between total monthly demands collected from CEA's website. The location of the basins and demand points are presented in Figure 3.4 and the list of the states with the estimated annual demand for the year 2012 is provided in Table 3.2. Population data provided in the table is based on 2011 Population Census [34].

Table 3.2 | List of states used as aggregated demand points

	Population (Million) (2011 Census)	Estimated Annual Demand in 2012 (GWh)	Load Factor
Delhi	16.8	30,013	0.71
Punjab	27.7	47,534	0.59
Uttaranchal	10.1	12,786	0.82
Himachal Pradesh	6.9	7,744	0.72
Uttar Pradesh	199.6	87,916	0.79
Bihar	103.8	13,774	0.76
West Bengal	91.3	40,777	0.79
Jharkhand	33	5,663	0.78
Assam	31.2	5,162	0.63
Chhattisgarh	25.5	17,718	0.81

The daily demand profiles of Delhi for each month in 2012 are presented in Figure 3.2. The highest demand is observed in summer months and the demand is lowest in winter months. The fact that the highest demand occurs in summer and daily peak demand is observed in the afternoon can be explained by the increasing cooling demand in summer. In winter, it is possible to observe two peaks in the daily load profile, one in the morning and one in the evening and this can be explained by the lightning demand. In Figure 3.3 we present the monthly total demand of the ten states listed in Table 3.2. The daily distribution of the demand used in the case studies of this thesis can be obtained from Appendix B.

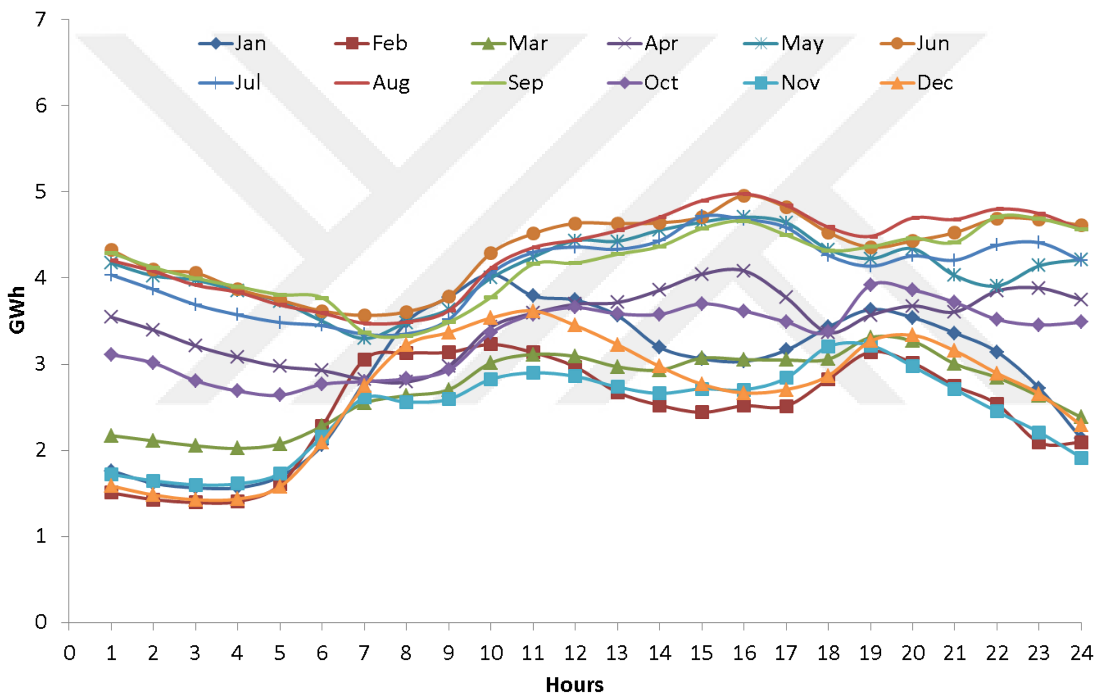


Figure 3.2 | Daily demand profile of Delhi in 2012

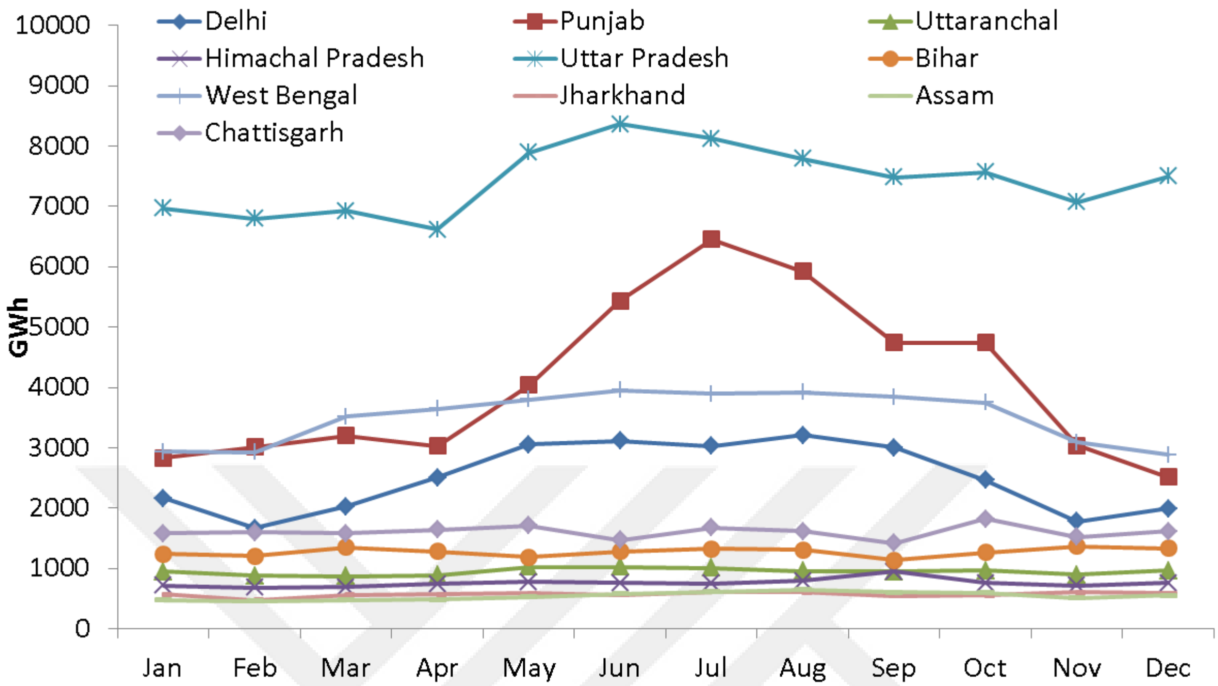


Figure 3.3 | Monthly total demand of ten states in India in 2012

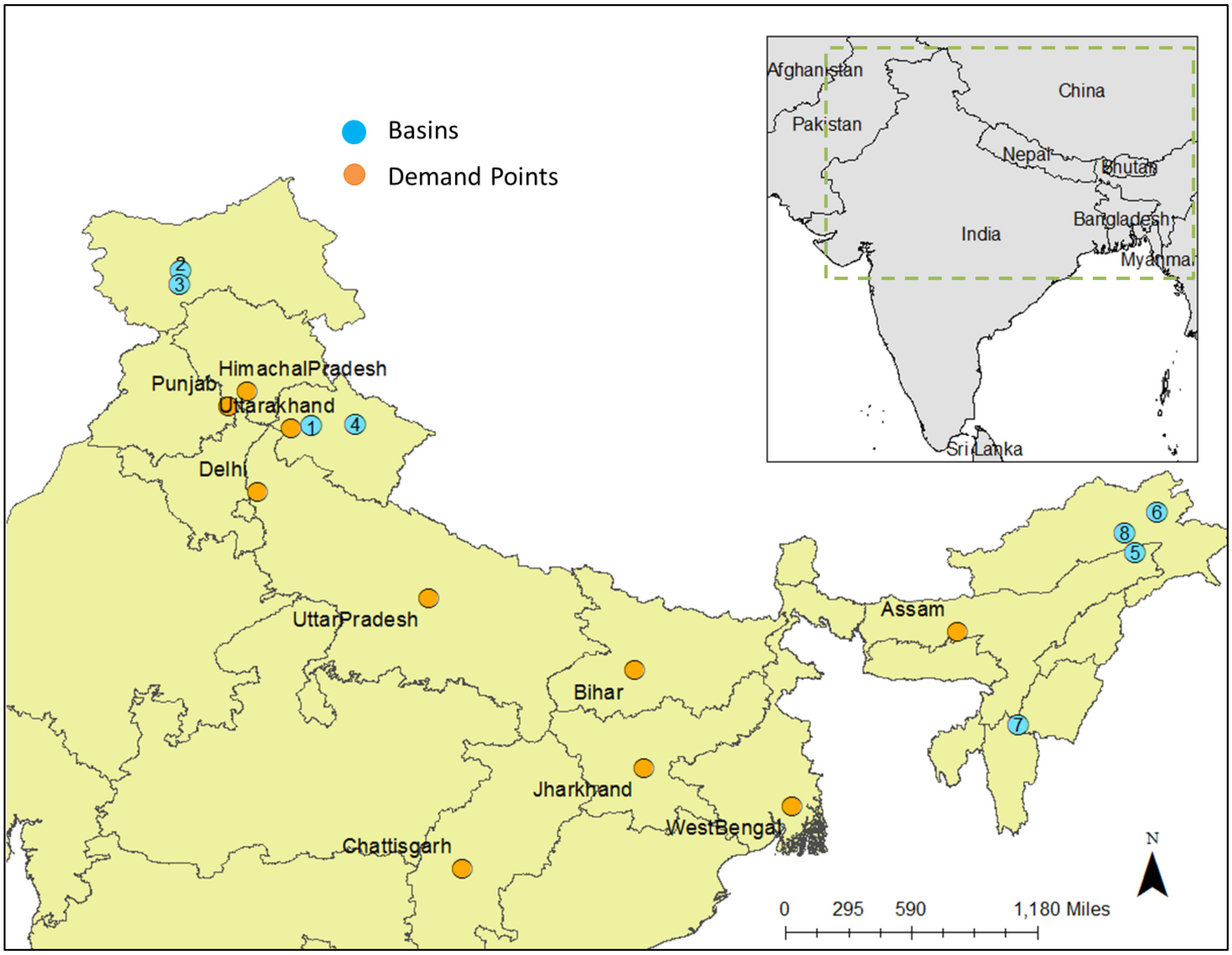


Figure 3.4 | Basins and demand points determined in India for analysis. Data is collected from CEA (Central Electricity Authority, Power Ministry of India) and other official websites to accurately estimate the 3-hourly demand load profile of each state for one year. If there is missing data for some days or hours within a day, interpolation/extrapolation methods are performed for projection.

3.3.2.3 Solar Radiation Data

Site and time specific, high resolution solar radiation data was developed using weather satellite by the U.S. National Renewable Energy Laboratory (NREL) in cooperation with India's Ministry of New and Renewable Energy. Global and direct irradiance at hourly intervals on the 10-km grid for all of India for the years 2001-2008 is available on NREL's website. Solar radiation data for all demand points used in the model is presented with some statistics in Table 3.3.

Table 3.3 | Solar Radiation Data

			Period 2001-2008 (Global Horizontal Irradiance : W /m2)				
	Lat (°E)	Long (°N)	Min	Max	Average	Standard Deviation	Coefficient of Variation
Delhi	29.02	77.38	0	1004	213	295	1.38
Punjab	30.79	76.78	0	980	208	287	1.38
Uttaranchal	30.33	78.06	0	998	209	289	1.38
HimachalPradesh	31.10	77.17	0	1013	213	292	1.37
UttarPradesh	26.85	80.91	0	1002	221	320	1.45
Bihar	25.37	85.13	0	1006	221	321	1.45
WestBengal	22.57	88.37	0	999	212	313	1.47
Jharkhand	23.35	85.33	0	1009	224	224	1.00
Assam	26.14	91.77	0	974	193	273	1.41
Chattisgarh	21.27	81.60	0	1003	232	312	1.34

3.4 Problem Formulation

As discussed above, we propose a stochastic linear programming approach to formulate and solve the described problem where uncertainties in the input data will be facilitated in the form of scenario realizations. Tables 3.4, 3.5 and 3.6 summarize the indices, parameters and variables used in the model.

Table 3.4 | Indices for parameters and decision variables

i:	hydropower generation point 1,…,I, with a total of I locations
j:	demand (solar power generation) point 1,…,J, with a total of J points
t:	time period 1,…,T, with a total of T periods
ω:	scenarios 1,…, Ω, with a total of Ω scenarios

Table 3.5 | Parameters of the model

n:	length of time periods

d:	dimensionless annualization parameter used to express the investment cost on a yearly basis
l_{ij}:	percentage of power loss while transmitting electricity from hydropower generation point i to demand point j.
g:	standard acceleration due to gravity (~9.8 m/s^2)
h_i:	height of the reservoir in hydropower generation point i
α:	efficiency of hydropower station
γ:	efficiency of solar panels
C_{Si}:	unit cost of reservoir capacity in hydropower generation point i
C_{PGi}:	unit cost of generator capacity in hydropower generation point i
C_{Mj}:	unit cost of solar array in demand point j
C_{Tij}:	unit cost of transmission line capacity between hydropower generation point i and demand point j
μ_j:	unit cost of generating electricity using diesel generator (i.e. penalty for mismatched demand in demand point j)
p_ω:	weight of scenario ω, where $\sum_{\omega=1}^{\Omega} p_\omega = 1$and $p_\omega \geq 0$

Table 3.6 | Variables of the model

Exogenous Variables:	
$W_i^{\omega t}$:	water runoff to hydropower generation point i in period t in scenario ω
$N_j^{\omega t}$:	solar radiation in point j in period t in in scenario ω
$D_j^{\omega t}$:	demand in point j at time t in in scenario ω

State/Decision Variables:

$S_i^{\omega t}$: water stored in the reservoir in hydropower generation point i at the end of period t in scenario ω

$Z_j^{\omega t}$: mismatched demand in demand point j in period t in scenario ω

$T_{ij}^{\omega t}$: electricity sent from hydropower generation point i to demand point j in period t in scenario ω

$L_i^{\omega t}$: water spilled from the reservoir in hydropower generation point i in period t in scenario ω

$R_i^{\omega t}$: water released from the reservoir in hydropower generation point i in period t in scenario ω

$Smax_i$: active reservoir capacity in hydropower generation point i

M_j: size of solar panels at demand point j

$PGmax_i$: generator size in hydropower generation point i

$Tmax_{ij}$: maximum energy transmitted from hydropower generation point i to demand point j

3.4.1 Objective Function

The objective of the model is to minimize the sum of the investment costs and expected penalty cost for the mismatched demand. Unit costs of investments are assumed to be the constant incremental cost of installing capacities and indexed by the location so that different costs parameters can be used for different locations. Objective function has five components:

i) Cost of Reservoirs:

$$O_1 = \sum_i C_{Si} * Smax_i$$

ii) Cost of Hydropower Generators:

$$O_2 = \sum_i C_{PGi} * PGmax_i$$

iii) Cost of Solar Power Stations:

$$O_3 = \sum_j C_{Mj} * M_j$$

iv) Cost of Transmission Lines:

$$O_4 = \sum_i \sum_j C_{Tij} * Tmax_{ij}$$

v) Expected Cost of Mismatched Demand:

$$O_5 = \sum_{jt\omega} p_\omega * Z_j^{\omega t} * \mu_j$$

Objective function can be stated as:

$$\min\ (O_1+O_2+O_3+O_4)*d+O_5$$

3.4.2 Constraints

The equality and inequality constraints of the problem are stated below:

(1) $S_i^{\omega t} \leq Smax_i \qquad \forall\, i, t, \omega$

(2) $S_i^{\omega t} = S_i^{\omega(t-1)} + W_i^{\omega t} - R_i^{\omega t} - L_i^{\omega t} \qquad \forall\, i, t\colon t > 1, \omega$

(3) $S_i^{\omega 1} = Smax_i + W_i^{\omega 1} - R_i^{\omega 1} - L_i^{\omega 1} \qquad \forall\, i, \omega$

(4) $S_i^{\omega T} = Smax_i \qquad \forall\, i, \omega$

(5) $f_{Gi}(R_i^{\omega t}) \leq PGmax_i * n \qquad \forall\, i, t, \omega$

(6) $\sum_j T_{ij}^{\omega t} = f_{Gi}(R_i^{\omega t}) \qquad \forall\, i, t, \omega$

(7) $T_{ij}^{\omega t} \leq Tmax_{ij} * n \qquad \forall\, i, j, t, \omega$

(8) $D_j^{\omega t} \leq Z_j^{\omega t} + f_{Sj}(M_j) + \sum_i T_{ij}^{\omega t} * (1 - \mathfrak{l}_{ij}) \qquad \forall\, j, t, \omega$

(9) $\quad S_i^{\omega t}, Smax_i, PGmax_i, R_i^{\omega t}, L_i^{\omega t}, M_j, T_{ij}^{\omega t}, Tmax_{ij}, Z_j^{\omega t} \ \geq 0 \quad \forall\, i, j, t, \omega$

The constraint in (1) ensures that water stored in the reservoir is limited by the size of the reservoir at each time period of every scenario. Constraints in (2-4) represent the mass balance equations in reservoirs. Constraint in (2) couples the reservoir levels between subsequent time periods. In (3) and (4), beginning and ending balance of reservoirs are set. Here, we assume that operations begin and end with full reservoirs at each scenario. In the model, scenarios start in September, which is almost end of Monsoon season in India and end at the end of August next year. Thus, it is quite reasonable to assume that reservoirs are full at this time of the year. Constraint in (5) ensures that generated energy defined by the function $f_{Gi}(R_i^{\omega t})$ is limited by the generator capacity at each time period of every scenario and $f_{Gi}(R_i^{\omega t}) = R_i^{\omega t} * g * hi * \alpha$. Constraint in (6) ensures that at any period in any scenario, total energy transmitted to the demand points from a hydropower generation is equal to generated energy in that hydropower generation point. Constraint in (7) ensures that transmitted energy is limited by the transmission line capacity. Constraint (8) ensures that demand $D_j^{\omega t}$ is met by sum of the energy transmitted from hydropower generation points, energy generated in solar power stations and energy generated using diesel generators within demand point j in time period t in scenario ω. Energy generated in solar power stations is defined by the function $f_{Sj}(M_j)$ where $f_{Sj}(M_j) = N_j^{\omega t} * M_j * \gamma$. For operational purposes, it is important to keep as much water as possible in the reservoirs although full kept amount will never be used. Here, since for one scenario the policy is "anticipatory" of what's happening in the future, the system spills the water that would not be needed. Although it would also have been optimal to keep water as much as possible in the reservoir, the solver can choose the solution with less water. For this reason, in the objective function we add another term: $\sum_{it\omega} S_i^{\omega t} * \mathcal{E}$ where $\mathcal{E}$ is a very small amount. This term tilts the balance so that the solver will choose the option with more water in the reservoir. When we report the final cost, we omit this term.

3.5 Results

Our linear program is implemented and solved in IBM ILOG CPLEX Optimization Studio (CPLEX) [35]. We present multiple case studies for India to emphasize the different aspects of our results. We first present results for *single basin-single demand point* case study to illustrate the basic information that our model can provide. Parameters that are used in analysis of the data are shown in Table 3.7. We also assess the sensitivity of the system in terms of cost parameters. Then, we introduce more basins and demand points to the system in order to observe how the network (transmission lines) and storage options can aid the integration of intermittent renewables as geographic aggregation smoothes the variability of the stream flows and solar radiations. Although our model is able to include different scenarios to take the uncertainty of the stream flows and solar radiations into account, in this chapter, we present results for a deterministic case where we assume how the exogenous variables will unfold within a year. Uncertainty of the data will be considered in the next chapter.

For the cases studied here, we use $1/m^3 as unit cost of reservoir capacity (i.e. constant incremental cost of installing reservoir capacity) which is within the ranges given in [38] and [39]. For the unit cost of powerhouse (generator, turbine/pump, transformers), we use $500/kW as this value is the lower bound of the range given in [40] for the capital cost of the medium to large hydropower stations (assuming that the lowest cost system has no storage). According to estimates given in [40], the cost of a 24 MW system with 1km^2 solar panel is estimated to be between $60 and $228 million. In our model, we use $150 million as the estimate for the representative solar power system unit area (1 km^2 solar panel). These cost parameters can clearly vary from site to site and this can easily be incorporated in the model assigning site specific unit cost for each investment variable. Efficiency of solar and hydropower systems are used as 12% [25] and 88% [22] respectively. Planning horizon is assumed to be 25 years and discount rate or risk-free interest rate is used as 5% [44]. Using discount rate and planning horizon, a dimensionless annualization parameter is calculated as 0.07 using the following formula to express the investment cost on a yearly basis. Let the planning horizon be n and interest rate be i,

$$\text{Annualization Parameter} = i / (1 - (1+i)^{-n})$$

Table 3.7 | Parameters used in the model

Unit Cost of Reservoir Capacity C_{Si}:	\$1/m3
Unit Cost of Generator Capacity C_{PGi}:	\$500/kW
Unit Cost of Solar Array C_{Mj}:	\$150/m2
Unit Cost of Diesel μj:	\$0.15/kWh
Efficiency of Hydropower System α:	88%
Efficiency of Solar Panels γ:	12%
Discount Rate:	5%
Planning Horizon:	25 years

3.5.1 Case Study I: Single Basin- Single Demand Point (Bhagirathi River and Delhi) Case

Bhagirathi is a Himalayan river in the state of Uttarakhand. On Bhagirathi, one of the tallest dam in Asia (260 m head), Tehri Dam, and Koteshwar Dam have been built as a part of the Tehri Dam Project [36]. According to VIC land surface model, average discharge of the river between 1954 and 2002 is estimated to be around 215 m^3/s. A sample one year data with 13.3

km^3 annual inflow is taken from the time series data for this analysis. Stream flow data for the sample year has been presented in Figure 3.5a. Demand data is projected with the method described in Section 3.3.2.2. Annual total demand is estimated to be around 30000 GWh for 2011-2012 and three-hourly demand load curve presented in Figure 3.5c. If all of the 13.3 km^3 inflow could be used to turn the turbine to generate electricity, it could generate about 3200 GWh electricity (assuming 100 m head and 88% efficiency). Since demand is an order of magnitude higher compared to hydro potential, in this section we present results for both normal demand and low demand profiles where low demand profile being equal to one fourth of normal demand. Solar radiation data of Delhi for 2002 is used for this analysis and data is presented in Figure 3.5b.

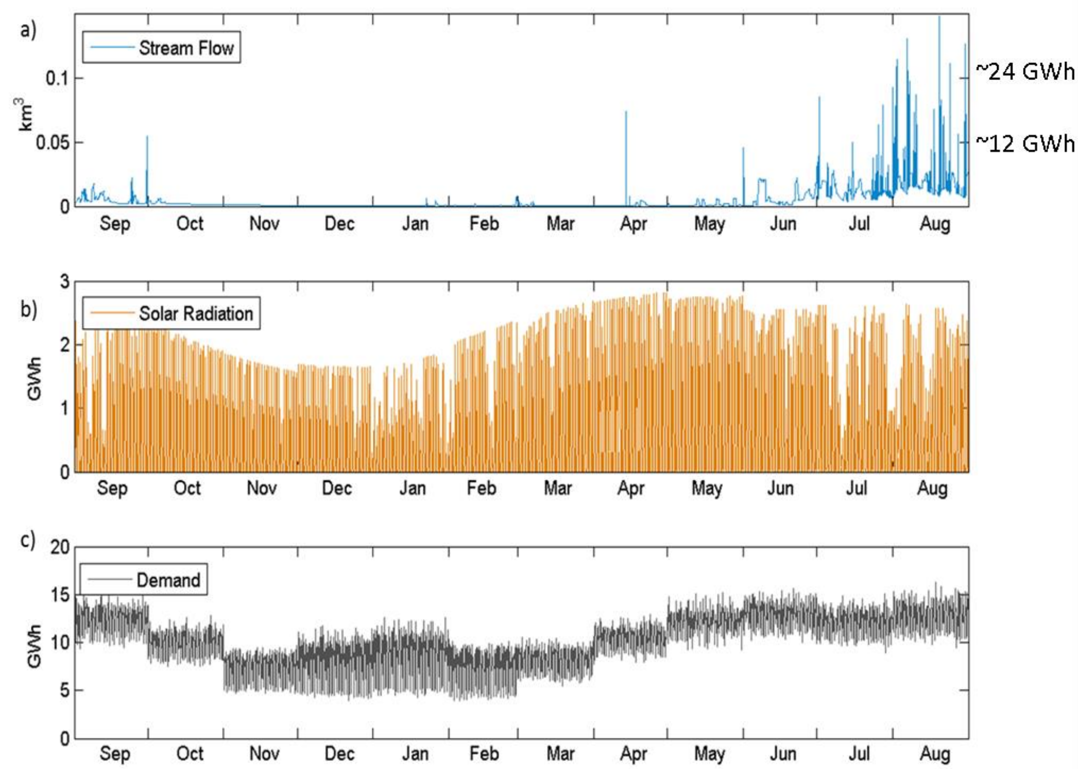


Figure 3.5 | a, 3-hourly stream flow data of Bhagirathi River for one year (Sep 1970-Aug 1971). **b,** Solar energy for every three hour per kilo meter square for year in Delhi. **c,** Demand load curve for one year in Delhi.

Table 3.8 compares the results for the actual and low demand profiles. A smaller reservoir and a bigger generator are needed in actual demand case where the demand is significantly higher (~10 times) than the energy potential of the river. The reason for a smaller reservoir in the higher demand case can be explained by the fact that most of the inflow goes directly to the turbine to meet as much demand as possible at that time period instead of being stored in the reservoir for future time periods. This can also be observed from the residence time of the reservoirs. Residence time is a widely used term in hydrology to express the average time a water molecule spends in that reservoir. Relying on the conservation of mass principle, residence time of the reservoir can be estimated by dividing the volume of the reservoir by the rate by which reservoir either gets filled or depleted [36]. Therefore, conceptually, this term can be also expressed as the time that takes reservoir to empty from full if no water were to enter. In this chapter, we estimate our residence time (RT) in days by the following formula:

$$RT_i^{\omega} = SUmax_i * 365 / \sum_t R_i^{\omega t}$$

where $SUmax_i$ is the size of the reservoir i and $\sum_t R_i^{\omega t}$ represents the total amount of water released from the reservoir i in one year (our time horizon) under scenario ω. Using the formula above, residence times of the reservoirs are estimated to be 5.25 and 11.75 days for actual and low demand cases respectively.

Table 3.8 | Summary results for single basin-single demand point case

	Actual Demand	**Low Demand (Actual Demand/4)**
Investments		
Reservoir Size (km^3) (1km^3 ~240 GWh)	0.18 (~44 GWh)	0.29 (~70 GWh)
Solar Panel Size (km^2) (1km^2~0.024GW)	63.67 (~1,5 GW)	12.5 (~0.3 GW)
Generator Size (GW)	1.86	1.06
Transmission Line Size (GW)	1.86	1.06

Production	GWh	% Demand	GWh	% Demand
Solar	11279	37.58	2533	33.76
Hydro	3045	10.15	2184	29.11
Diesel	15689	52.27	2786	37.13
Total	30013	100	7503	100

Demand		
Average Demand (GW)	3.43	0.86
Peak Demand (GW)	5.40	1.35
Load Factor	0.63	0.63

Diesel		
Average Diesel Usage (GW)	1.79	0.31
Peak Diesel Usage (GW)	5.14	1.29

Solar		
Efficiency	79%	90%
Average Solar Production (GW)	1.63	0.32
Peak Solar Production (GW)	7.18	1.41
Capacity Factor	0.23	0.23

Hydro		
Average Hydro Production (GW)	0.35	0.25
Peak Hydro Production (GW)	1.86	1.06
Capacity Factor	0.19	0.24

Transmission		
Average Transmission (GW)	0.35	0.25
Peak Transmission (GW)	1.86	1.06
Capacity Factor	0.19	0.24

Unit Cost		
Unit Cost of Hydro ($/kWh)	0.03	0.04
Unit Cost of Solar ($/kWh)	0.06	0.05
Unit Cost of Diesel ($/kWh)	0.15	0.15
Unit Cost of the System ($/kWh)	0.10	0.08

Distribution of each "fuel" (supply) type that has been used to meet the demand is also presented in Table 3.8 with annual aggregated figures. The change in this distribution over a year is also presented for low demand profile in Figure 3.6 below. As can be seen from the Figure 3.5 b-c, with having lowest values in winter and peak values in summer, solar radiation and demand have similar trends. This causes the percentage wise average solar energy contribution to be quite constant throughout the year. However, we still observe that solar contribution in Monsoon season is highly fluctuating due to the cloudy days of this rainy season (The variability of the solar radiation in Monsoon season can also be observed in Figure 3.5b). Contribution of hydro goes up to 100% in Monsoon season when there is excessive inflow to the reservoir and there is almost no need for diesel's contribution. In the other seasons, hydro energy and diesel work as complementary to each other.

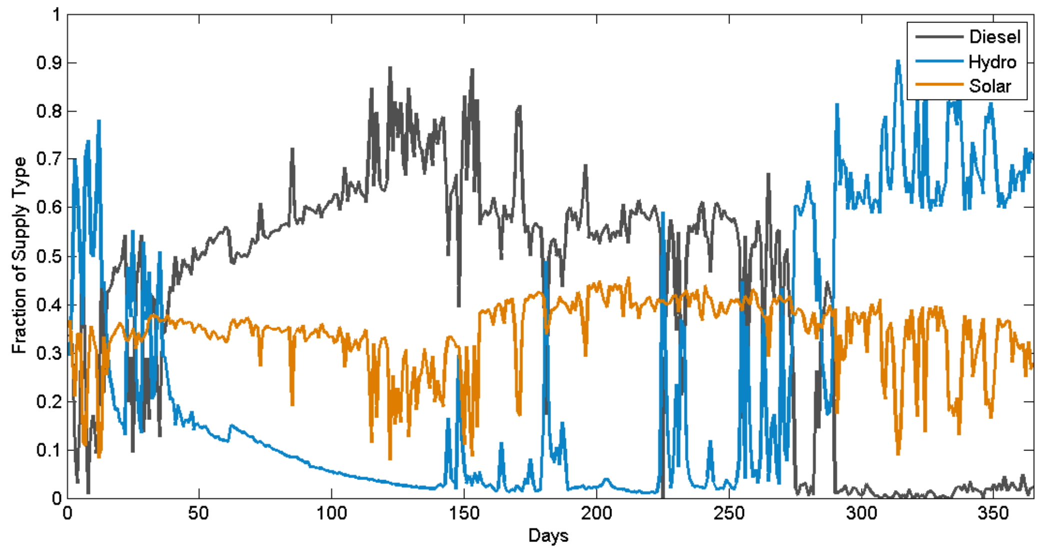


Figure 3.6 | Contribution of each “fuel” (supply) type that has been used to meet the demand through the year for each day. Solar energy contribution is quite constant throughout the year with some fluctuations in the Monsoon. Hydro and diesel work as complementary to each other.

3.5.1.1 Comparison of Alternative Technologies

When 100% of the demand is met by only one component of the system (e.g. diesel), unit cost of the system will also be the unit cost of that component (for diesel it is $0.15/kWh). Here, we start with no renewable source in the system (100% diesel) and gradually modify the system so that it includes more advanced use of renewable source and finally obtain our model that we have discussed above. In particular we look at diesel and solar, diesel and run-of-the-river hydro, diesel and conventional hydro and finally diesel, solar and conventional hydro combinations respectively. For this analysis, without formulating a new problem for each alternative, we use our model described in Section 3.4 assigning extreme cost parameters for the system components that we want to exclude. For example, when we want to find the optimal system for diesel and run of the river alternative, we make sure that we assigned high enough costs for reservoir and solar panel that solution does not include these components. Then, model finds the least cost solution for the components with regular cost parameters. Results for this process are presented in Figure 3.7. The detailed results for components of alternative systems are presented in Table 3.9. Table 3.9 also shows how much diesel capacity is displaced as alternative technologies are included into the system.

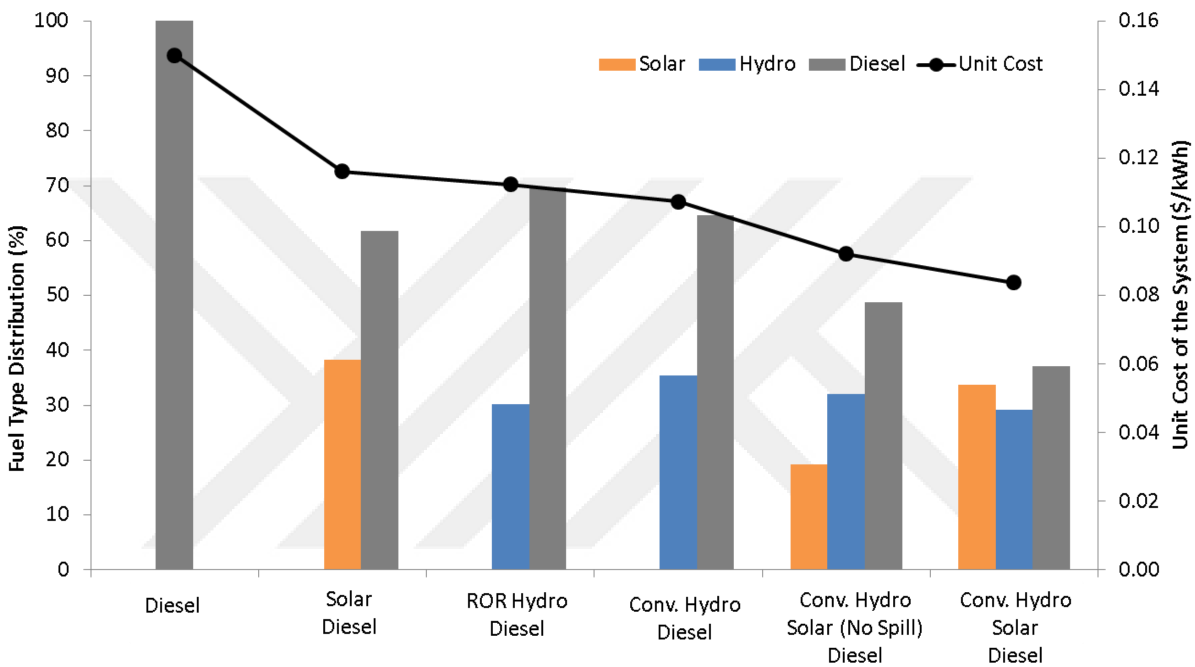


Figure 3.7 | Comparison of alternative technologies. Solely diesel is the most expensive. As the use and the variety of clean energy increase, the unit cost goes down substantially. It is more cost efficient to design solar panel area based on high demand and spill some of the renewable energy in low demand periods.

Table 3.9 | Summary results for alternative technologies

	DIESEL+ SOLAR	DIESEL+ ROR	DIESEL+ CONV HYDRO	DIESEL+ SOLAR (no Spill)+ CONV. HYDRO	DIESEL+ SOLAR+ CONV. HYDRO
Investments					
Reservoir Size (km^3) (1km^3 ~240 GWh)	NA	NA	0.24 (~57 GWh)	0.23 (~56 GWh)	0.29 (~70 GWh)
Solar Panel Size (km^2) (1km^2~0.024GW)	16.35 (~ 0.39 GW)	NA	NA	6.35 (~0.15 GW)	12.58 (~0.30 GW)
Generator Size (GW)	0	1.10	1.14	1.06	1.06
Transmission Line Size (GW)	0	1.10	1.14	1.06	1.06
Production (GWh)					
Solar	2861	0	0	1422	2548
Hydro	0	2275	2661	2406	2178
Diesel	4647	5233	4847	3679	2782
Total	7508	7508	7508	7508	7508
Demand					
Average Demand (GW)	0.86				
Peak Demand (GW)	1.35				
Load Factor	0.63				

Diesel					
Average Diesel Usage (GW)	0.53	0.60	0.55	0.42	0.32
Peak Diesel Usage (GW)	1.31	1.26	1.27	1.22	1.29

Solar					
Efficiency	78%			100%	90%
Average Solar Production (GW)	0.42	NA	NA	0.16	0.32
Peak Solar Production (GW)	1.84			0.72	1.42
Capacity Factor	0.23			0.23	0.23

Hydro					
Average Hydro Production (GW)		0.26	0.30	0.27	0.25
Peak Hydro Production (GW)	NA	1.10	1.14	1.06	1.06
Capacity Factor		0.24	0.27	0.26	0.23

Transmission					
Average Transmission (GW)		0.26	0.30	0.27	0.25
Peak Transmission (GW)	NA	1.10	1.14	1.06	1.06
Capacity Factor		0.24	0.27	0.26	0.23

Unit Cost					
Unit Cost of Hydro ($/kWh)	NA	0.02	0.02	0.02	0.03
Unit Cost of Solar ($/kWh)	0.06	NA	0.00	0.05	0.05
Unit Cost of Diesel ($/kWh)	0.15	0.15	0.15	0.15	0.15
Unit Cost of the System ($/kWh)	0.116	0.112	0.107	0.092	0.084

As can be seen in Figure 3.7, as the use of the renewable increase and gets more advanced, unit cost of the system decreases and the most cost efficient alternative is the one that we have used in our model. Figure 3.7 also shows that diesel-solar system is more expensive than the diesel-run-of-the-river system. When a reservoir is added to system (diesel-conventional hydro), flexibility of deferring water release decreases the system cost even more compared to diesel-run-of-the-river system.

In Figure 3.7, there is an additional system (Conv. Hydro - Solar (No spill) - Diesel) that we have included in order to make a point clearer and we needed to slightly modify our formulation for this additional system in the following way. In our regular formulation, we allow excessive solar energy to be spilled which means that some of the solar energy generated may not be used to fulfill the demand and solar system works with a lower efficiency than 100%. In order to see if allowing spill is a profitable decision, in constraint (8) of our formulation, we can replace the inequality with equality and force the system not to spill renewable energy. It is quite interesting to see that when we allow solar to be spilled, unit cost of the system is reduced. When we have a detailed comparison between the two systems, we see that solar panel area of the system in which spill is allowed is almost twice of the system in which it is not allowed. This is due to the fact that constraint (8) is satisfied for each time period and in case of equality, solar panel area is mainly determined based on the time periods when demand is low. However, it is more cost efficient to increase the solar panel area considering the time periods when there is high demand and spill some of the solar energy in periods when low demand occurs. Figure 3.8 shows solar energy production and demand for the systems when some renewable energy is spilled and when it is not spilled. In Figure 3.8a the orange area (solar) above the grey area (demand) is being spilled.

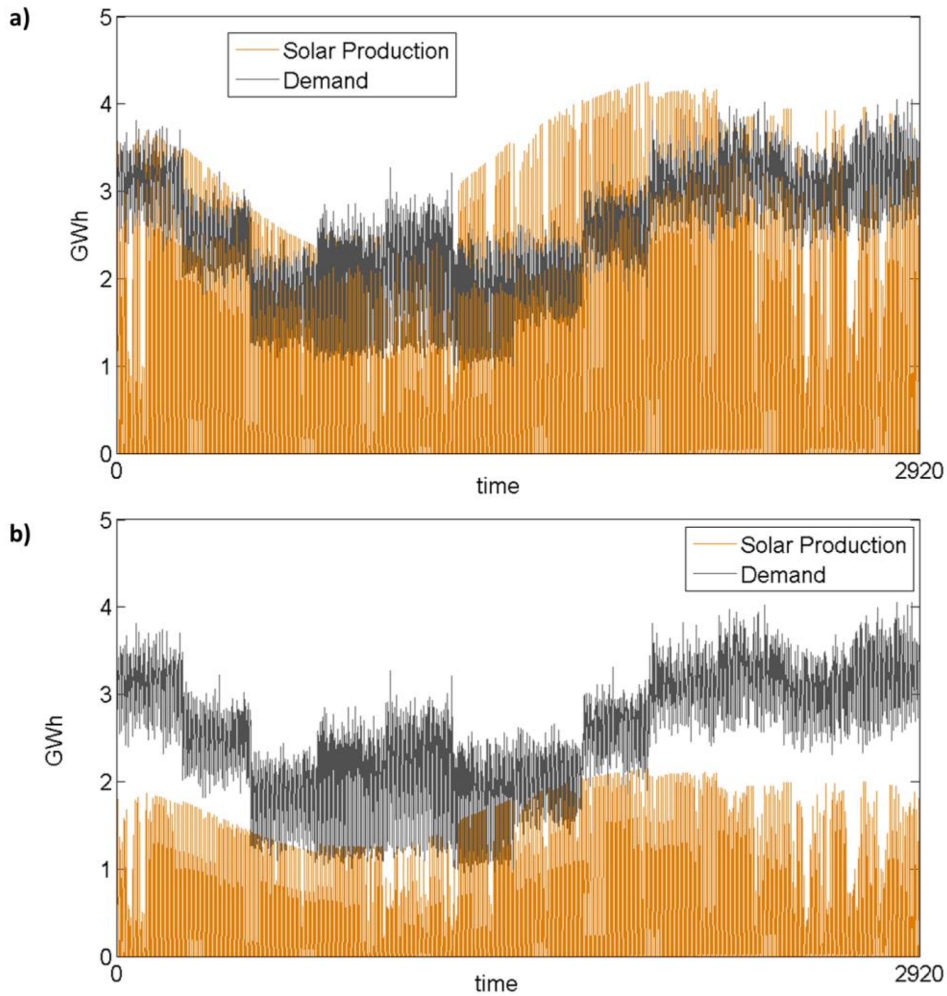
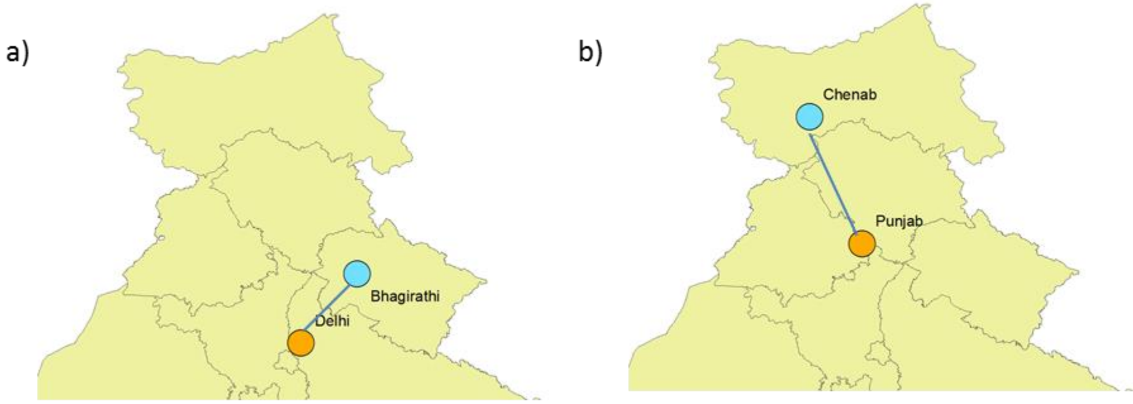


Figure 3.8 | Solar Production and demand for **a)** when spill is allowed, **b)** when spill is not allowed.

3.5.2 Case Study II: Multi Demand Points, Multi Basins System

In this section, we show the effect of combining multiple basins and demand points together. As individual systems are connected with transmission lines and work as a single system, the variability and intermittency of renewable sources is expected to be smoothed out. To see this effect, we first run our model with Chenab (another basin with 52 km^3 annual stream flow) and Punjab (demand point with approximately 12000 GWh (quarter of actual demand) annual demand. Then, we combine this case with our previous Bhagirathi (13.3 km^3 annual stream flow) and Delhi (7500 GWh annual demand) case study and run a two-basin and two-demand points case. Figure 3.9 summarizes the comparison between single-basin–single-demand point cases and Figure 3.10 shows the results for the two basins-two demand points' case.

Results in Figure 3.10 evidently show that with much smaller reservoirs (78% and 58% decrease in reservoir size respectively); combined system generates 2% more hydro energy. We note here that smaller reservoir may not necessarily imply a better solution since an additional inflow could drive larger reservoirs when higher hydro storage capacity is needed or smaller reservoirs when a system like run-of-the-river is sufficient. However, there are also other factors that we do not consider here but make smaller reservoirs favorable such as environmental effects of large dams. In addition, solar energy generation increases almost 6% and the improvements in hydro and solar energy generation provide a 10% reduction in diesel usage. It should also be noted that combined system is more cost efficient as total cost of single-basin-single-demand point cases is more than the total cost of combined system. In our case study, we need to pay 13% more transmission cost to build extra transmission line, yet we still save 5% from the overall total cost.



a)

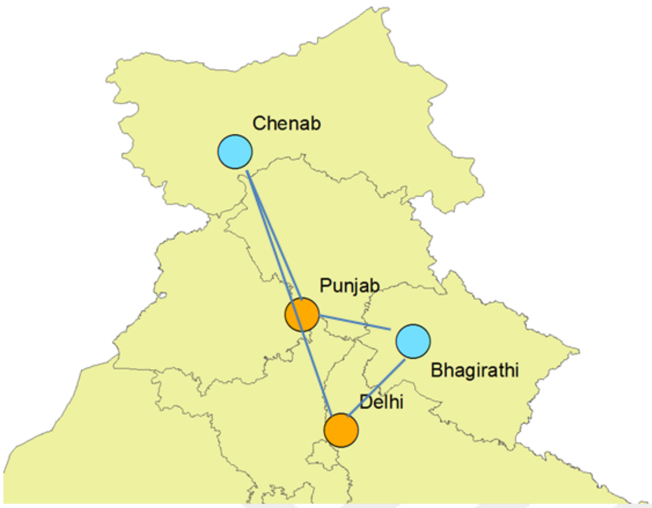
ReservoirSize (km^3) (1km^3~240GWh)		0.29 (~70 GWh)
SolarPanelSize(km^2) (1km^2~0.024 GW)		12.58 (~0.30 GW)
Generator Size (GW)		1.06
TransmissionLine (GW)		1.06
Unit Costs ($/kWh)		
Hydro		0.03
Transmission		0.01
Solar		0.05
Diesel		0.15
System		0.08
Energy Generation		
Solar	34%	2533
Hydro	29%	2186
Diesel	37%	2784
Total	100%	7503

b)

ReservoirSize (km^3) (1km^3~240GWh)		0.10 (~24 GWh)
SolarPanelSize(km^2) (1km^2~0.024 GW)		14.80 (~0.36GW)
Generator Size (GW)		2.11
TransmissionLine (GW)		2.11
Unit Costs ($/kWh)		
Hydro		0.01
Transmission		0.01
Solar		0.05
Diesel		0.15
System		0.05
Energy Generation		
Solar	26%	3054
Hydro	57%	6722
Diesel	18%	2107
Total	100%	11883

Figure 3.9 | Summarized results for *one-basin, one-demand point* case studies. Bhagirathi has 13.3 km^3 annual stream flow and Delhi has 7500 GWh annual demand. Chenab has 52 km^3 annual stream flow and Punjab has a 12000 GWh (quarter of actual demand) computed annual demand.

a)



b)

	Bhagirathi	Chenab
ReservoirSize (km^3)	0.065 (~16 GWh)	0.0424 (~10 GWh)
GeneratorSize (GW)	0.758	2.458
TransmissionLine to (GW)		
Delhi	0.416	0.733
Punjab	0.472	1.725

SolarPanelSize (km^2)	
Delhi	11.78 (~ 0.28 GW)
Punjab	17.97 (~ 0.42 GW)

Unit Costs ($/kWh)	
Hydro	0.013
Transmission	0.008
Solar	0.054
Diesel	0.150
System	0.060

c)

	Delhi		Punjab	
Energy Generation	GWh	% Demand	GWh	% Demand
Solar	2420	32.2%	3497	29.4%
Hydro	3105	41.4%	5984	50.4%
Diesel	1979	26.4%	2402	20.2%
Total	7503	100.0%	11883	100.0%

Figure 3.10 | Summarized results for *two basins-two demand points* case**.** Combined system generates 2% more hydro energy with much smaller reservoirs (78% and 58% decrease in reservoir size respectively). Solar energy generation increases almost 6% and the improvements in hydro and solar energy generation provide a 10% reduction in diesel usage. There is 13% more transmission cost to build extra transmission line but overall cost is 5% lower.

3.5.3 Case Study III: Multi Demand Points, Multi Basins System II

After we show the effect of combining multiple basins and demand points together in the previous section, in this case study, we analyze a much larger problem and provide results that include all the basins and demand points of India that we could collect the data for. The motivation behind this study is to demonstrate how an actual hybrid network could look like in the future country we are interested in where fossil fuels are only used in the minimal form.

The model is run with 8 basins and 10 demand points with their actual demand that are shown in Figure 3.2. Here, it is not possible to have comparison figures relative to the individual systems as we present in 3.5.2 due to the fact that basins and demand points are not one to one paired. However we can consider the utilization of transmission lines in a minimum cost network as the proof for higher hydro and solar contribution with smoother variations and decreased diesel contribution. As discussed in Section 3.3.1, unit cost of transmission lines between demand points and basins are calculated based on the spherical distance between points.

Table 3.10 summarizes the results and shows the proposed sizes for reservoirs and generators for each basin. With having 456 km^3 annual inflows Siang River has by far the highest potential and provides electricity to almost all the states in the case study. One should keep in mind that in this study we include lower/upper bounds for neither reservoir sizes nor generator capacities. Other environmental and geographic constraints which are specific to basins are also not in the scope of this thesis. These results are based on the stream flow potentials of the proposed basins.

Table 3.10 | Size of the hydropower stations proposed for basins

Rivers	Annual Inflow (km^3/ ~GWh)	Reservoir Size (km^3/ ~GWh)	Generator Size (GW)	Generation (%Demand)
Bhagirathi	13.51 /~3242	0.28/~67	1.22	1.0%
Pinder	6.29 /~1510	0.16/~38	0.71	0.5%
Chenab	45.57 /~10937	0.49/~118	4.27	3.8%
Marusudar	12.02 /~2885	0.23/~55	1.19	1.0%
Lohit	19.17 /~4600	0.07/~17	1.21	1.5%
Dibang	9.17 /~2200	0.02/~5	0.40	0.6%
Barak	57.71/~13850	0.08/~19	4.08	4.6%
Siang	456.47/~109552	0.81/~194	23.07	30.3%

As a summary, unit cost of the system is 6.9 cents/kWh (unit cost of diesel production is 15 cents/kWh and see Table 3.7 for other parameters). Hydropower is the cheapest source with having 1.2 cents/kWh. After including 1.7 cents average unit cost for transmission, adjusted cost for hydro power is about 3 cents/kWh. Unit cost of solar energy is estimated to be 5.4 cents.

In Table 3.11, solar panel areas that are proposed for installation in each demand point are presented. It can be seen that solar panel area is almost linearly related to the demand. The role of hydropower in meeting the demand changes between 36%- 59%. Total annual inflow of basins is about 620 km^3 and if all of the 620 km^3 inflow could be used to turn the turbine to generate electricity, it would generate about 168000 GWh electricity (assuming 100 m head and 88% efficiency). Given that sum of demand across all demand points is about 270000 GWh, the system utilizes hydropower fairly well and hydro portion in meeting the total demand is quite close to the theoretical (probably impractical) limit (62%).

Table 3.11 | Solar Panel Areas and Energy Generation by Type

Demand Points	**Demand (GWh)**	**Solar Panel Area (km^2/~ GW)**	**Energy Generation by Type (% Demand)**		
			Solar	Hydro	Diesel
Delhi	30013	49.7/~1.2	34%	36%	31%
Punjab	47534	75.1/~1.8	30%	46%	24%
Uttaranchal	11357	17.2/~0.4	25%	59%	16%
Himalach Pradesh	7744	13.6/~0.3	33%	46%	20%
Uttar Pradesh	87916	132.4/~3.2	29%	40%	31%
Bihar	13774	20.2/~0.5	29%	42%	29%
West Bengal	40777	55.0/~1.3	27%	45%	27%
Jharkhand	5663	8.7/~0.2	31%	41%	27%
Assam	5162	7.2/~0.2	25%	49%	26%
Chhattisgarh	17718	24.8/~0.6	30%	51%	19%

The major transmission lines (≥1GW) between hydropower stations and demand points can be seen in Table 3.12. The first four basins listed in the tables are located in the North part of India and the others are located in the north-east region. An important result to note here is that high stream flow potential of the basins in the north-east region was quite useful in fulfilling the demand of northern states and high capacity long transmission lines are preferred instead of local diesel generators within demand points.

Table 3.12 | Transmission line capacities between basins and demand points

Rivers/ States	Delhi	Punjab	Uttarak-hand	Himalach Pradesh	Uttar Pradesh	Bihar	West Bengal	Jharak-hand	Assam	Chattis-garh
Bhagirathi	0.2	1.0	0.7	0.0	0.0	0.0	0.0	0.0	0.0	0.1
Pinder	0.2	0.1	0.7	0.0	0.4	0.0	0.0	0.0	0.0	0.1
Chenab	1.6	1.8	0.8	0.4	0.0	0.0	0.0	0.0	0.0	0.0
Marusudar	0.0	1.2	0.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Lohit	0.0	0.0	0.0	0.0	0.0	0.0	1.2	0.0	0.0	0.0
Dibang	0.0	0.1	0.0	0.1	0.0	0.0	0.0	0.1	0.0	0.0
Barak	0.0	0.0	1.6	0.1	0.5	0.0	1.8	0.0	0.3	0.4
Siang	2.3	3.8	0.6	0.5	9.5	1.6	2.3	0.6	0.6	1.6

Table 3.13 | Capacity factors of the transmission lines between basins and demand points

Rivers/ States	Delhi	Punjab	Uttarak-hand	Himalach Pradesh	Uttar Pradesh	Bihar	West Bengal	Jharak-hand	Assam	Chattis-garh
Bhagirathi	0.1	0.2	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.3
Pinder	0.1	0.1	0.1	0.0	0.1	0.0	0.0	0.0	0.0	0.2
Chenab	0.2	0.3	0.2	0.4	0.0	0.0	0.0	0.0	0.0	0.0
Marusudar	0.0	0.2	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Lohit	0.0	0.0	0.0	0.0	0.0	0.0	0.4	0.0	0.0	0.0
Dibang	0.0	0.5	0.0	0.4	0.0	0.4	0.0	0.5	0.0	0.5
Barak	0.2	0.0	0.2	0.2	0.4	0.2	0.3	0.2	0.3	0.4
Siang	0.3	0.4	0.2	0.4	0.4	0.4	0.4	0.4	0.3	0.5

3.6 Discussion

3.6.1 Sensitivity Analysis

In this section, several sensitivity analyses have been performed to inspect how sensitive our model to the cost parameters as these analyses can provide meaningful insight about differences in the optimal solution of the problem in response to small changes in the cost parameters. We first show the sensitivity range of each investment variable for the calculated optimal solution, and then we vary the unit cost of diesel, solar panel area and reservoir size and resolve the linear program.

In Table 3.14, each investment variable, its reduced cost and the range over which its objective function coefficient can vary without forcing a change in the optimal basis is displayed. One needs to do this with some +/- range in the total investment. We see that we have very small ranges for each variable in which our current optimal solution remains the same. This means that with the default (realistic) values, there is a dynamic balance between the cost parameters and any foreseen significant change in the cost parameters should be seriously considered by the infrastructure planners before they make a decision as it may alter the optimal solution.

Table 3.14 | Sensitivity ranges for the optimal solution obtained with cost parameters used in the model

	Unit Cost	Down	Up	Sensitivity Range
Reservoir ($/m3)	1	-2.02%	0.01%	0.979-1.000
Generator ($/kW)	500	-0.07%	1.19%	499.6-505.9
Solar Panel ($/m2)	150	-0.16%	0.19%	149.7-150.2
Transmission Line ($/kW)	254	-0.03%	2.24%	253.9-259.6

Our objective function in this problem minimizes annualized investment cost of power systems and transmission lines and total cost of generating electricity using diesel generators. In the case studies in this chapter, we use $0.25/kWh, $\$1/m^3$, and $150 as the unit cost of diesel generation, reservoir size and solar panel area respectively. Diesel component of the network is ideally the expensive alternative and can be considered as a penalty that is paid for each unit of electricity that could not be generated using renewable sources. Therefore, unit cost of diesel is also marginal cost and our model compares that with the marginal cost of renewable sources. For example, if increasing the size of the reservoir to generate an additional kWh of hydro energy (without causing any size increase in generator or transmission lines) is cheaper than the unit cost of diesel, the model increases the size of the reservoir. Table 3.15 - 3.17 summarize the analyses performed for the unit costs.

Table 3.15 | Sensitivity Analysis for Diesel Cost

	μj=0.05	μj=0.1	μj=0.15	μj=0.2	μj=0.25	μj=0.3
ReservoirSize (km^3)	0.06	0.12	0.29	0.31	0.31	5.38
GeneratorSize (GW)	0.89	0.99	1.06	1.08	1.09	1.10
SolarPanelSize (km^2)	2.08	11.09	12.50	16.21	18.78	21.78
TransmissionLine (GW)	0.89	0.99	1.06	1.08	1.09	1.10
% Demand						
Solar	6%	31%	34%	38%	40%	42%
Hydro	31%	28%	29%	28%	28%	43%
Diesel	63%	40%	37%	34%	32%	15%
ResidenceTime of Reservoir (days)	2.3	4.7	11.7	12.7	13.3	145.4
Efficiency of Solar System	100%	95%	90%	78%	72%	65%

Table 3.16 | Sensitivity Analysis with the Unit Cost of Reservoirs

	$C_S = \$\,0.5\ /m^3$	$C_S = \$\,1\ /m^3$	$C_S = \$\,1.5\ /m^3$	$C_S = \$\,2\ /m^3$
ReservoirSize (km^3)	4.70	0.29	0.11	0.08
GeneratorSize (GW)	1.05	1.06	1.05	1.05
SolarPanelSize (km^2)	12.60	12.50	12.55	12.60
TransmissionLine (GW)	1.05	1.06	1.05	1.05
% Demand				
Solar	34%	34%	34%	34%
Hydro	43%	29%	28%	28%
Diesel	23%	37%	38%	39%
ResidenceTime of Reservoir (days)	127.1	11.7	4.7	3.3
Efficiency of Solar	90%	90%	90%	90%

Table 3.17 | Sensitivity Analysis for Unit Cost of Solar Panel

	$C_M = \$50\ /m^2$	$C_M = \$100\ /m^2$	$C_M = \$150\ /m^2$	$C_M = \$200\ /m^2$	$C_M = \$250\ /m^2$
ReservoirSize (km^3)	0.34	0.31	0.29	0.29	0.28
GeneratorSize (GW)	1.04	1.05	1.06	1.06	1.05
SolarPanelSize (km^2)	27.07	17.50	12.50	11.45	10.79
TransmissionLine (GW)	1.04	1.05	1.06	1.06	1.05
% Demand					
Solar	45%	39%	34%	32%	31%
Hydro	26%	28%	29%	30%	30%
Diesel	29%	33%	37%	38%	39%
ResidenceTime of Reservoir (days)	15.3	13.1	11.7	11.3	10.9
Efficiency of Solar	56%	75%	90%	94%	96%

In Tables 3.15-3.17, we observe some interesting sudden changes in the distribution in meeting the annual demand. For example, in Table 3.15 when diesel cost is increased to \$0.30/kWh, there is a significant increase in the size of the reservoir and proportion of hydro in meeting the annual demand. The reason behind this difference is as follows: the unit cost of reservoir capacity (i.e. constant incremental cost of additional capacity) that we used for our analysis is $\$1/m^3$. If an additional 1 m^3 of a reservoir is used only once (in one time period only) during our planning horizon (1 year), we could generate 0.24 kWh of electricity (with 100 m head and 88% efficiency) with that $1m^3$ water. In addition, annualized cost of an additional $1m^3$ reservoir is \$0.07 (with 25 year life time and 5% discount rate). Therefore, marginal cost of hydro can be roughly calculated as \$0.29/kWh. This is the main reason for the jump in Table 3.15 as the preference in terms of the units cost flips once the unit cost for the diesel generation becomes higher than \$0.29/kWh. Similar analyses can also be compiled to explain the increase in solar proportion when diesel cost is increased from \$0.05/kWh to \$0.10/kWh in Table 3.15 and the decrease in the reservoir size from 4.70 km^3 to 0.29 km^3 in Table 3.16.

Other two analyses with the unit cost of reservoir and solar panel area are presented in Table 3.16 and Table 3.17.

3.7 Conclusion

In summary, we have presented a model which is designed to help infrastructure planners make long-term investment decisions based on the results for electric dispatch, energy resource allocation and storage over one-year horizon. The objective of the model is to minimize the sum of the investment costs and expected penalty cost for the mismatched demand. We have shown that as we increase the use and complexity of the renewable sources, the unit cost for the overall system decreases and by combining multiple individual systems, we can have a further reduction in system cost with a possibility of smaller network components.

Bibliography

[1] Key World Energy Statistics (2011), International Energy Agency.

[2] Ashok, S. (2007). Optimised model for community-based hybrid energy system.*Renewable energy*, *32*(7), 1155-1164.

[3] Ekren, O., & Ekren, B. Y. (2010). Size optimization of a PV/wind hybrid energy conversion system with battery storage using simulated annealing. *Applied Energy*, *87*(2), 592-598.

[4] Ekren, O., Ekren, B. Y., & Ozerdem, B. (2009). Break-even analysis and size optimization of a PV/wind hybrid energy conversion system with battery storage–a case study. *Applied Energy*, *86*(7), 1043-1054.

[5] Dudhani, S., Sinha, A. K., & Inamdar, S. S. (2006). Assessment of small hydropower potential using remote sensing data for sustainable development in India. *Energy Policy*, *34*(17), 3195-3205.

[6] Powell, W. B., George, A., Simão, H., Scott, W., Lamont, A., & Stewart, J. (2012). SMART: A Stochastic Multiscale Model for the Analysis of Energy Resources, Technology, and Policy. *INFORMS Journal on Computing*, *24*(4), 665-682.

[7] Mallah, S., & Bansal, N. K. (2010). Allocation of energy resources for power generation in India: business as usual and energy efficiency. *Energy Policy*,*38*(2), 1059-1066.

[8] Cormio, C., Dicorato, M., Minoia, A., & Trovato, M. (2003). A regional energy planning methodology including renewable energy sources and environmental constraints. *Renewable and Sustainable Energy Reviews*, *7*(2), 99-130.

[9] NEMS (2003), The National Energy Modeling System: An Overview 2003, Technical report, Energy Information Administration, Office of Integrated Analysis and Forecasting, U.S. Department of Energy, Washington, DC 20585.

[10] Loulou, R., Goldstein, G. & Noble, K. (2004), Documentation for the MARKAL Family of Models, Technical report, Energy Technology Systems Analysis Programme.

[11 Fishbone, L. G., & Abilock, H. (1981). Markal, a linear-programming model for energy systems analysis: Technical description of the bnl version. *International journal of Energy research*, *5*(4), 353-375.

[12] Lamont, A. D. (2008). Assessing the long-term system value of intermittent electric generation technologies. *Energy Economics*, *30*(3), 1208-1231.

[13] International Energy Agency. (2012). *World Energy Outlook, 2012*. OECD/IEA.

[14] Ahn, S. J., & Graczyk, D. (2012). Understanding Energy Challenges in India: Policies, Players, and Issues. *Paris: IEA.*

[15] Bhattacharya, S. C., & Jana, C. (2009). Renewable energy in India: Historical developments and prospects. *Energy*, *34*(8), 981-991.

[16] Suganthi, L., & Williams, A. (2000). Renewable energy in India—a modelling study for 2020–2021. *Energy policy*, *28*(15), 1095-1109.

[17] Singh, P., & Jain, S. K. (2002). Snow and glacier melt in the Satluj River at Bhakra Dam in the western Himalayan region. *Hydrological sciences journal*,*47*(1), 93-106.

[18] Sharma, N. K., Tiwari, P. K., & Sood, Y. R. (2012). Solar energy in India: Strategies, policies, perspectives and future potential. *Renewable and Sustainable Energy Reviews*, *16*(1), 933-941.

[19] http://pmindia.nic.in/speech-details.php?nodeid=1307

[20] Kumar, A., Kumar, K., Kaushik, N., Sharma, S., & Mishra, S. (2010). Renewable energy in India: Current status and future Potentials. *Renewable and Sustainable Energy Reviews*, *14*(8), 2434-2442.

[21] MNRE (Ministry of New and Renewable Energy), http://www.mnre.gov.in/mission-and-vision-2/achievements/

[22] Førsund, F. (2007). Hydro Power Economics. International Series in Operations Research and Management Science.

[23] Barlev, D., Vidu, R., & Stroeve, P. (2011). Innovation in concentrated solar power. *Solar Energy Materials and Solar Cells*, *95*(10), 2703-2725.

[24] Mousazadeh, H., Keyhani, A., Javadi, A., Mobli, H., Abrinia, K., & Sharifi, A. (2009). A review of principle and sun-tracking methods for maximizing solar systems output. *Renewable and Sustainable Energy Reviews*, *13*(8), 1800-1818.

[25] Parida, B., Iniyan, S., & Goic, R. (2011). A review of solar photovoltaic technologies. *Renewable and Sustainable Energy Reviews*, *15*(3), 1625-1636.

[26] Sanoh, A., Kocaman, A. S., Kocal, S., Sherpa, S., & Modi, V. (2014). The economics of clean energy resource development and grid interconnection in Africa. *Renewable Energy*, *62*, 598-609.

[27] Wu, F., Varaiya, P., Spiller, P., & Oren, S. (1996). Folk theorems on transmission access: Proofs and counterexamples. *Journal of Regulatory Economics*, *10*(1), 5-23.

[28] Tejada-Guibert, J. A., Johnson, S. A., & Stedinger, J. R. (1995). The value of hydrologic information in stochastic dynamic programming models of a multireservoir system. *Water Resources Research*, *31*(10), 2571-2579.

[29] Dharmadhikary, S. (2008). *Mountains of concrete: Dam building in the Himalayas* (No. id: 1815).

[30] Liang, X., Lettenmaier, D., Wood, E., & Burges, S. (1994). A simple hydrologically based model of land surface water and energy fluxes for general circulation models. Journal of

Geophysical Research, 99(14), 415–14.

[31] Liang, X., Wood, E. F., & Lettenmaier, D. P. (1996). Surface soil moisture parameterization of the VIC-2L model: Evaluation and modification. *Global and Planetary Change*, *13*(1), 195-206.

[32] Singh, P., & Jain, S. K. (2003). Modelling of streamflow and its components for a large Himalayan basin with predominant snowmelt yields. *Hydrological sciences journal*, *48*(2), 257-276.

[33] http://www.cea.nic.in/executive_summary.html (2013)

[34] Statistical Abstract of Delhi (2012), Directorate of Economics and Statistics

[35] http://www-01.ibm.com/software/info/ilog/

[36] Adhikari, B. R. (2009). Tehri Dam: An Engineering Marvel. *Hydro Nepal: Journal of Water, Energy and Environment*, *5*, 26-30.

[37] Rueda, F., Moreno-Ostos, E., & Armengol, J. (2006). The residence time of river water in reservoirs. *Ecological Modelling*, *191*(2), 260-274.

[38] http://www.hydrocoop.org/publications/2.12.1.1.article.pdf

[39] Keller, A. A., Sakthivadivel, R., & Seckler, D. W. (2000). Water scarcity and the role of storage in development (Vol. 39)

[40] OpenEI Transparent Cost Database. Accessed 12/10/2013

[41] Gutman, R., Marchenko, P. P., & Dunlop, R. D. (1979). Analytical development of loadability characteristics for EHV and UHV transmission lines. *Power Apparatus and Systems, IEEE Transactions on*, (2), 606-617.

[42] Hao, J., & Xu, W. (2008, October). Extended transmission line loadability curve by including voltage stability constrains. In Electric Power Conference, 2008. EPEC 2008. IEEE Canada (pp. 1-5). IEEE.

[43] Bergen, A. R. (2000). Vijay Vittal Power System Analysis.

[44] Ekholm, T., Krey, V., Pachauri, S., & Riahi, K. (2010). Determinants of household energy consumption in India. *Energy Policy*, 38(10), 5696-5707.

Chapter 4

Modelling of a Hybrid Energy System with Pumped Hydro Storage

4.1 Introduction

Current supply for electricity generation mostly relies on fossil fuels. However, fossil fuels are finite and their combustion causes global warming and health hazards. To reduce the role of fossil fuels and ease the concerns on the electricity generation, energy models which involve clean and renewable energy sources are necessitated. The International Energy Agency (IEA) estimates that nearly 50 percent of global electricity supplies will have to come from renewable energy sources in order to achieve a 50 percent reduction of global CO_2 emissions by 2050.

Transition to renewable sources is inevitable. However, renewable sources present important challenges while integrating them into the power systems: i) Renewable sources are *intermittent*; ii) They are heavily dependent on the *spatial location*. Intermittency causes limited control on power output because of variability and partially predictability of the renewable sources such as solar and wind and dependence on the spatial location causes a mismatched between potential of renewable energy generation and where the energy will be ultimately consumed. Delucchi and Jacobson argue that it is possible to overcome the difficulties working with renewables and show that it is technologically and economically feasible to meet the 100 percent of the world energy demand by wind, water and solar [1,2].

To mitigate the intermittency of renewable sources, there are several ideas proposed to design and operate cost efficient and reliable renewable energy systems. Designing hybrid systems which operate as a combination of alternative resources, using energy storage and long distance transmission lines and employing demand response programs help reduce intermittency of the renewables and allow the grid accommodate more variation on both supplies and demand. The technology improvements in transmission lines provide more geographic aggregation which smoothes the variability of intermittent sources over large distances. This helps design more efficient hybrid systems reducing the storage costs and generate more dispatchable (controllable) power.

As generation and distribution of energy resources are becoming more complex, there is an ever-growing need for mathematical optimization to design the worldwide competitive energy systems and to provide a formal framework that enables systematic and clear decision-making in energy operations. Here, we imagine a country in the future where hydro and solar are the dominant sources and fossil fuels are only available in minimal form (e.g. in the shape of diesel generators). In this country, we identify candidate locations for pumped hydro power stations and aggregated demand point locations such as the cities or the states in which solar energy production is possible. We then mathematically model a hybrid energy generation and allocation system including long distance transmissions and pumped hydro storage (water is pumped from lower reservoir to upper reservoir during periods of low demand to be released for generation when demand is high).

Aim of the model is to determine optimal sizing of infrastructure needed to match demand and supply in a most reliabile and cost effective way. With this model, we are for the first time combining three important concepts which help reduce the intermittency of renewable sources: hybrid systems, pumped hydro storage with two-level reservoirs and long distance transmission in regional or national level. Similar to Chapter 3, the model will be presented with a case study of India and we compare the conventional hydro power system discussed in Chapter 3 and the pumped hydro power systems in terms of reliability and cost efficiency. Our model helps assess how efficiently solar energy in addition to pumped hydro system utilizing high hydro power potential in Himalaya Mountains could meet the growing electricity demand if fossil fuels could be almost completely phased out from electricity generation.

4.2 Background

There is an ample amount of literature dealing with problems which are relevant to our study such as planning of hybrid energy systems, long-term energy investment planning problems, water reservoir management problem. In Section 3.2.1 of Chapter 3, we explain the developments in these problems and how our problem can be differentiated from the local level hybrid systems [3-5] and here we discuss how the stochastic nature of this problem and water reservoir management have been studied in the previous studies.

So-called energy-technology management problems are also studied extensively in the literature [6-13]. These problems usually differ from each other by the technologies they include, uncertainties present in the model or the objective of the models. Stoyan et al. (2011) [12] uses a scenario based approach to consider uncertainties and proposes stochastic mixed-integer model which minimizes cost and emission levels associated with energy generation while meeting energy demand of a given region. Powell et al. (2010) [13] addresses the problem of modeling energy resource allocations with long term investment strategies for new technologies using an approximate dynamic programming approach.

The literature on water reservoir management is fairly broad [13-22]. There is a variety of models that aim to model specifically hydroelectric power generation [20 - 22]. Some of these work use dynamic programming technique and they propose different alternatives to overcome the "curse of dimensionality" which is a major drawback of dynamic programs [18]. Pereira and Pinto (1991) [19] provide one of the first attempts of stochastic dual dynamic programming approach applied to large, multi-reservoir hydropower plant. Jacobs et al. (1995) [20] uses the concept of scenario trees to include information related the uncertain exogenous variables in hydroelectric generations systems. Lall et al. (1981) [14] proposes a non-linear model for planning more generalized regional water-energy systems. Lall et al. (1988) [13] describes a simulation optimization methodology proposed for multipurpose reservoir systems. Their model differs from others by not including monthly mass balance equations for reservoirs but they require presumption of monthly demand fractions for hydropower, municipal and irrigation demands. A common practice in reservoir management problem is to use a monthly time increment and as also discussed in Chapter 4, it is not a suitable approach when hydropower generation is combined with the renewable sources to accurately capture the diurnal variability.

Another body of literature can be grouped under modeling of energy storage to handle either variability or uncertainty of supply sources and demand [23-25]. Castronuovo et al. (2004) [23] takes the hourly variability of wind into account to optimize the daily operation of a wind-hydro power plant. They determine a set of scenarios for wind power generation and solve a linear program for each scenario individually providing an average and a range of output. Brown et al (2008) [24] uses a fuzzy clustering approach to determine hourly load and renewable generations scenarios for a day and minimizes the expected daily cost of operation and amortization of investments. The pumped hydro systems in these models include only one elevated reservoir which is expected to have a daily cycle, storing energy during off-peak hours to be used during peak hours. Since the pump/generator operations is supposed to be on a daily basis, the planning horizon is chosen to be one day, assuming that beginning and ending balances in the reservoir will be same. In a more recent study, Kuznia et al. (2012) [25] proposes a stochastic mixed integer programming model for a hybrid power system design problem, including renewable energy generation, storage device, transmission network and thermal generators. They generate daily scenarios with hourly time increments taking samples from the different seasons of a year. An important assumption here is that their storage type also has daily cycle. They show that this problem is a special case of capacitated lot sizing problem which is in general a NP-hard problem. To solve the stochastic mixed integer program, they propose a Benders' decomposition algorithm.

In summary, although there is vast literature on the direct impact of various clean energy technologies, few investigations involve large scale analysis of managing such technologies to meet future energy demands [12]. Here, we propose the first hybrid model with pumped hydro storage with two-level reservoir designed for national/regional energy planning considering spatial locations and transmission lines where the storage cycle can be anywhere between a few hours to a year.

4.3 Problem Statement

As in Chapter 3, we are interested in optimally sizing the infrastructure of a hybrid system which combines hydro and solar energy and transmission lines between basins and demand points. To mitigate the volatility of the supply and demand, we still use reservoirs as "water storage" with one crucial difference. Instead of having one reservoir, we have two-level of reservoirs and water can be pumped from lower reservoir to upper reservoir during periods of low demand to be released for generation when demand is high. A schematic illustration of our hybrid system with pumped hydro storage is given in Figure 4.1. In pumped hydro energy storage (PHES), the generator and water turbine can operate as a motor and pump. The excessive solar energy can be transmitted to hydropower stations with bi-directional transmission lines and can be used to pump water stored in the lower reservoir to upper reservoir. Unlike the hybrid system with conventional hydro storage which is described in the previous chapter, here solar energy can be stored and this increases the flexibility while reducing the intermittency of solar and stream flow. The goal is to minimize the infrastructure cost of power stations (reservoirs, generator, and solar panel area), transmission lines and the cost of an expensive back-up source (e.g. diesel generators) that will be used when renewable sources are not available.

Reservoirs in PHES effectively transfer energy from low use periods to peak use periods, allowing the system operate based on demand load while maintaining high system reliability. Thus, these systems can increase the flexibility of hydro systems even more compared to conventional systems, however they still do face important limitations that other hydro systems have. Hydro installations are geographically limited and in certain cases adding additional hydro may not be an option. They are only feasible for locations that have sufficient water availability and are capable of having large reservoirs at different heights. These geographical challenges increase the importance of transmission lines. Different technologies (alternating current (AC) and direct current (DC)) are available that can facilitate the routing of power from one location to another in a controllable fashion. For connection of remote renewables, high voltage direct current (HVDC) technology is especially well suited due to low losses and high controllability when compared to AC. In this problem, possible network flow directions from sources to demand points are prescribed with dedicated lines and designed as a point-to-point topology. Here we neither model the grid itself nor consider real power flow equations and phase angle differences. We assume that power flows over lines can be independently assigned. This representation of power flows, which captures point-to-point movements without explicitly modeling the grid, is a common approximation made in policy studies [6].

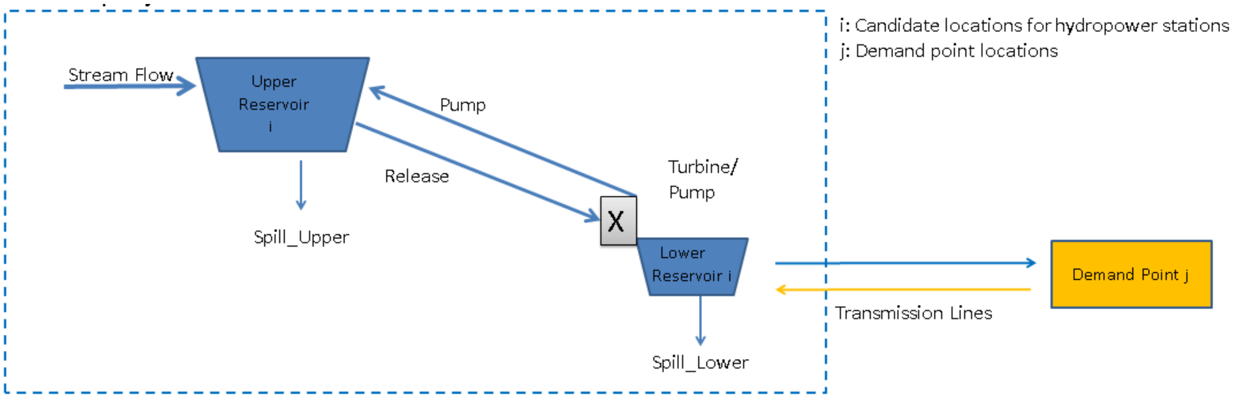


Figure 4.1 | A schematic illustration for hybrid system with pumped hydro storage. There are two-level of reservoirs and water can be pumped from lower reservoir to upper reservoir during periods of low demand to be released for generation when demand is high.

We use the same data sets (basins and demand point locations, solar radiation, stream flow and demand) that we used in the previous chapter in order to be able to see easily the advantage of a pumped hydro system over a conventional hydro system. As in the hybrid system with conventional hydro storage, it is necessary to model energy supply and demand with hourly time periods for at least one year planning horizon to accurately capture both the hourly and seasonal variability of the sources. Here, as the solar energy can be also stored and there is significant solar radiation variability in hourly level, using hourly time increments in addressing this problem becomes even more important. An approach that avoids capturing every time increment over a year by simply sampling different time periods (e.g. different time of the years and time of the days) fails to accurately model the storage. Moreover, modeling reservoir systems is especially more complicated than modeling other storage types such as a battery which stores energy during the day and releases at night. However, in a realistic reservoir system, we may put water in reservoir storage in September so it can be used couple of months later in the future. In addition, we also need to include the mass balance equations of the lower reservoir and power flows from demand points to hydro stations and deal with a larger scale problem compared to conventional system.

In terms of capturing the uncertainty, we are following the powerful and highly acceptable approach in similar problems in the literature using a scenario approach [24-26]. Here, we present a scenario based static model with multiple time periods that are coupled by storage. By scenario approach, a set of prototype 1-year series with 3 hourly time increments are determined as a particular realization of the uncertain data such as stream flows. A drawback of a scenario-based approach is the fact that scenarios are generated in advance, and this limits their ability to capture the interaction between decisions and exogenous events. We assume that effect of this drawback can be minimized during the real-time operations of the power systems. For example, in case of quite rainy season which is not foreseen and captured by scenarios, water in the reservoirs can be controlled to be prepared for the season.

4.4 Methodology

In this section, we first provide our stochastic linear program formulation for the hybrid system with pumped hydro storage. Then, we describe our methodology to determine the scenarios which will be used to determine our scenarios.

4.4.1 Problem Formulation

As we have discussed above, to formulate and solve the described problem, we propose a two-stage stochastic linear programming approach where uncertainties in the input data will be facilitated in the form of scenario realizations. Table 4.1, 4.2 and 4.3 summarize the indices, parameters and variables used in the model.

Table 4.1| Indices for parameters and decision variables

i:	hydropower generation point 1,…,I, with a total of I locations
j:	demand (solar power generation) point 1,…,J, with a total of J points
t:	time period 1,…,T, with a total of T periods
ω:	scenarios 1,…, Ω, with a total of Ω scenarios

Table 4.2| Parameters for model

n:	length of time periods
d:	dimensionless annualization parameter used to express the investment cost on a yearly basis
l_{ij}:	percentage of power loss while transmitting electricity from hydropower generation point i to demand point j.
g:	acceleration
h_i:	height of the reservoir in hydropower generation point i
α:	efficiency of hydropower station
γ:	efficiency of solar panels
C_{Si}:	unit cost of reservoir capacity in hydropower generation point i
C_{PGi}:	unit cost of generator capacity in hydropower generation point i
C_{Mj}:	unit cost of solar array in demand point j
C_{Tij}:	unit cost of transmission line capacity between hydropower generation point i and demand point j
μ_j:	unit cost of generating electricity using diesel generator (i.e. penalty for mismatched demand in demand point j)

p_ω:	weight of scenario ω, where $\sum_{\omega=1}^{\Omega} p_\omega = 1$and $p_\omega \geq 0$

Table 4.3| Variables of the model

Exogenous Variables:

$W_i^{\omega t}$:	water runoff to hydropower generation point i in period t in scenario ω
$N_j^{\omega t}$:	solar radiation in point j in period t in in scenario ω
$D_j^{\omega t}$:	demand in point j at time t in in scenario ω

State / Decision Variables:

$SU_i^{\omega t}$: water stored in the upper reservoir in hydropower generation point i at the end of period t in scenario ω

$SL_i^{\omega t}$: water stored in the lower reservoir in hydropower generation point i at the end of period t in scenario ω

$Z_j^{\omega t}$: mismatched demand in demand point j in period t in scenario ω

$T_{ij}^{\omega t}$: electricity sent from hydropower generation point i to demand point j in period t in scenario ω

$T_{ji}^{\omega t}$: electricity sent from demand point j to hydropower generation point i in period t in scenario ω

$LU_i^{\omega t}$: water spilled from upper reservoir in hydropower generation point i in period t in scenario ω

$LL_i^{\omega t}$: water spilled from lower reservoir in hydropower generation point i in period t in scenario ω

$V_j^{\omega t}$: solar energy internally used in point j in period t in scenario ω

$R_i^{\omega t}$: water released from upper reservoir in hydropower generation point i in period t in scenario ω

$P_i^{\omega t}$: water pumped from lower reservoir to upper reservoir in hydropower generation point i in period t in scenario ω

$SUmax_i$: active upper reservoir capacity in hydropower generation point i

$SLmax_i$: lower reservoir capacity in hydropower generation point i

M_j: size of solar panels at demand point j

$PGmax_i$: generator size in hydropower generation point i

$Tmax_{ij}$: maximum energy transmitted from hydropower generation point i to demand point j

---

4.4.1.1 Objective Function

Similar to the Chapter 3, the objective of the model is to minimize the sum of the investment costs and expected penalty cost for the mismatched demand. Unit costs of investments are assumed to be the constant incremental amount of installing capacities and indexed by the location so that different costs parameters can be used for different locations. Objective function has five components:

i) Cost of Reservoirs:

$$O_1 = \sum_i C_{S_i} * (SUmax_i + SLmax_i)$$

ii) Cost of Hydropower Generators:

$$O_2 = \sum_i C_{PGi} * PGmax_i$$

iii) Cost of Solar Power Stations:

$$O_3 = \sum_j C_{Mj} * M_j$$

iv) Cost of Transmission Lines:

$$O_4 = \sum_i \sum_j C_{Tij} * \ Tmax_{ij}$$

v) Expected Cost of Mismatched Demand:

$$O_5 = \sum_{jt\omega} p_\omega * Z_j^{\omega t} * \mu_j$$

Objective function can be stated as:

min $(O_1+O_2+O_3+O_4)*d+O_5$

4.4.1.2 Constraints

The equality and inequality constraints of the problem are stated below:

(1) $SU_i^{\omega t} \leq SUmax_i$ $\quad \forall\, i, t, \omega$

(2) $SL_i^{\omega t} \leq \ SLmax_i$ $\quad \forall\, i, t, \omega$

(3) $SU_i^{\omega t} = SU_i^{\omega(t-1)} + W_i^{\omega t} + P_i^{\omega t} - R_i^{\omega t} - LU_i^{\omega t}$ $\quad \forall\, i, t{:}\, t > 1, \omega$

(4) $SU_i^{\omega 1} = SUmax_i + W_i^{\omega 1} + P_i^{\omega 1} - R_i^{\omega 1} - LU_i^{\omega 1}$ $\quad \forall\, i, \omega$

(5) $SU_i^{\omega T} = SUmax_i$ $\quad \forall\, i, \omega$

(6) $SL_i^{\omega t} = SL_i^{\omega(t-1)} + R_i^{\omega t} - P_i^{\omega t} - LL_i^{\omega t}$ $\quad \forall\, i, t, \omega$

(7) $SL_i^{\omega 1} = R_i^{\omega 1} - P_i^{\omega 1} - LL_i^{\omega 1}$ $\quad \forall\, i, t{:}\, t > 1, \omega$

(8) $max\{f_{Gi}(R_i^{\omega t}), f_{Pi}(P_i^{\omega t})\} \leq PGmax_i * n$ $\quad \forall\, i, t, \omega$

(9) $\sum_j T_{ij}^{\omega t} = f_{Gi}(R_i^{\omega t})$ $\quad \forall\, i, t, \omega$

(10) $f_{Pi}(P_i^{\omega t}) = \sum_j T_{ji}^{\omega t} * \left(1 - \mathrm{l}_{ji}\right)$ $\quad \forall\, i, t, \omega$

(11) $max\left\{T_{ij}^{\omega t}, T_{ji}^{\omega t}\right\} \leq Tmax_{ij} * n$ $\quad \forall\, i, j, t, \omega$

(12) $V_j^{\omega t} + \sum_i T_{ji}^{\omega t} \leq f_{Sj}\left(M_j\right)$ $\quad \forall\, j, t, \omega$

(13) $Z_j^{\omega t} = D_j^{\omega t} - V_j^{\omega t} - \sum_i T_{ij}^{\omega t} * \left(1 - \mathrm{l}_{ij}\right)$ $\quad \forall\, j, t, \omega$

(14) $SU_i^{\omega t}, SL_i^{\omega t}, SUmax_i, SLmax_i, PGmax_i, P_i^{\omega t}, R_i^{\omega t}, LU_i^{\omega t}, LL_i^{\omega t}, M_j$

$$T_{ij}^{\omega t}, T_{ji}^{\omega t}, Tmax_{ij}, V_j^{\omega t}, Z_j^{\omega t} \geq 0 \qquad \forall\, i, j, t, \omega$$

In the hybrid energy system with pumped hydro storage, we have thirteen constraints as opposed to nine in the conventional system in Chapter 3. This is as a result of including the mass balance equations of the lower reservoir and power flows from demand points to hydro station. We explain each one of the updated constraints here. The constraints in (1) and (2) ensure that water stored in the reservoirs is limited by the size of the reservoir at each time period of every scenario. Constraints in (3-5) represent the mass balance equations in reservoirs. Constraint in (3) couples the upper reservoir levels between subsequent time periods. In (4) and (5), beginning and ending balance of reservoirs are set. Here, we again assume that operations begin and end with full upper reservoirs. In the model, each scenario starts in September, the end of Monsoon season in India, and continues for a year. Thus, it is quite reasonable to assume that reservoirs are full at this time of the year. Constraint in (6-7) couples the lower reservoir levels between subsequent time periods. Constraint in (8) ensures that generated and pumped energy defined by the functions $f_{Gi}(R_i^{\omega t})$ and $f_{Pi}(P_i^{\omega t})$ respectively, are limited by the generator/pump capacity in every time period of every scenario where $f_{Gi}(R_i^{\omega t}) = R_i^{\omega t} * g * h_i * \alpha$ and $f_{Pi}(P_i^{\omega t}) = P_i^{\omega t} * g * h_i * 1/\alpha$. Constraint can easily be linearized by reformulating it in the form of $f_{Gi}(R_i^{\omega t}) \leq PGmax_i * n$ and $f_{Pi}(P_i^{\omega t}) \leq PGmax_i * n$. Constraint in (9) ensures that energy transmitted to demand points from a hydropower generation is equal to generated energy in that hydropower generation point in period t in scenario ω and constraint. Likewise, constraint in (10) provides that pumped energy cannot be greater than the total amount of energy sent from demand points in in period t in scenario ω. Constraint in (11) ensures that transmitted energy is limited by the transmission line capacity. In pumped hydro system, since there is also power flow from demand points, transmission lines are bi-directional and should be size based on the flow in both directions. Constraint in (12) ensures that solar energy internally used in a demand point j and the total energy sent from point j to hydropower stations should be less than the amount of solar energy generated in demand point j in time period t in scenario ω. Energy generated in solar power stations is defined by the function $f_{Sj}(M_j)$ where $f_{Sj}(M_j) = N_j^{\omega t} * M_j * \gamma$. Constraint in (13) ensures that demand $D_j^{\omega t}$ is met by sum of the energy transmitted from hydropower generation points, energy generated in solar power stations and energy generated using diesel generators within demand point j in time period t in scenario ω. Since we don't have an operational cost for each pumping and releasing operation in our objective function, to prevent the pumping and releasing happening at the same time especially for multi-basin and multi–demand point cases, we also include a binary decision variable, I_p , into our model which makes our model a mixed integer program (MIP). The two sets of constraints, $P_i^{\omega t} \leq I_p * M$ and $R_i^{\omega t} \leq (1 - I_p) * M$ are, then, added for $\forall\, i, t, \omega$ where M is a very large number. The same idea can also be used to prevent the simultaneous bidirectional flows in the transmission lines. As in Chapter 3, again for operational purposes, it is important to keep as much water as possible in the upper reservoirs although full kept amount will never be used. Here, since for one scenario the policy is "anticipatory" of what's happening in the future, the system spills the water that would not be needed. Although it would also have been optimal to keep water as much as possible in the upper reservoir, the solver can choose the solution with less water. For this reason, in the objective function we add another term: $\sum_{it\omega} SU_i^{\omega t} * \mathcal{E}$ where $\mathcal{E}$ is a very small amount. This term tilts the balance so that the solver will choose the option with more water in the upper reservoir. When we report the final cost, we omit this term.

4.4.2 Scenario Determination

Our model consists of three exogenous variables: stream flow, solar radiation and demand. As discussed in Chapter 3, time series data for stream flows for the years between 1951 and 2004 is obtained from the VIC land surface model [27, 28]. We decompose the time series using an additive model, as the seasonal fluctuations are roughly constant in size over time and do not seem to depend on the level of the time series. The components of the additive model are presented in Figure 4.3a. We observe that after 1970, time series has a more fluctuating trend and we focus on this part of the time series while determining our scenarios. To be able to find the best representative years of the time series, we first divide a year into six seasons, each including two months: autumn (end of monsoon), prewinter (post monsoon), winter, spring, summer, monsoon. For each season of the years between 1971 and 2004, we calculate the mean, standard deviation and coefficient of variation. Our goal is to determine the years in which 'wettest', 'driest' and the most 'variable' seasons are observed. To find the wet and dry seasons we use the mean of the stream flows within a month as our metric. For example, in Bahagirathi River, the wettest autumn season is observed in 1978 with the mean flow of 8.17x10-3 km3 whereas, the direst autumn season is observed in 1974 with the mean flow of 8.45 x10-4 km3. To find the year in which the most variable autumn season is observed, we use coefficient of variation as our decision metric. Coefficient of variation represents the ratio of standard deviation to the mean and is a widely used metric to see the variability of the data sets with different mean.

Time series of climate and meteorological variables are frequently serially dependent, reflecting the strong persistence of the meteorological and climate phenomena. For example, a rainfall effect can be seen in multiple consecutive periods. The autocorrelation function of stream flow up to 50 lags presented in Figure 4.2 also supports this. The idea of taking yearly samples from the time series also helps keep the dynamics within a year. Using the approach described here, we determined 13 years as representatives of the time series in terms of the extreme and the most variable cases.

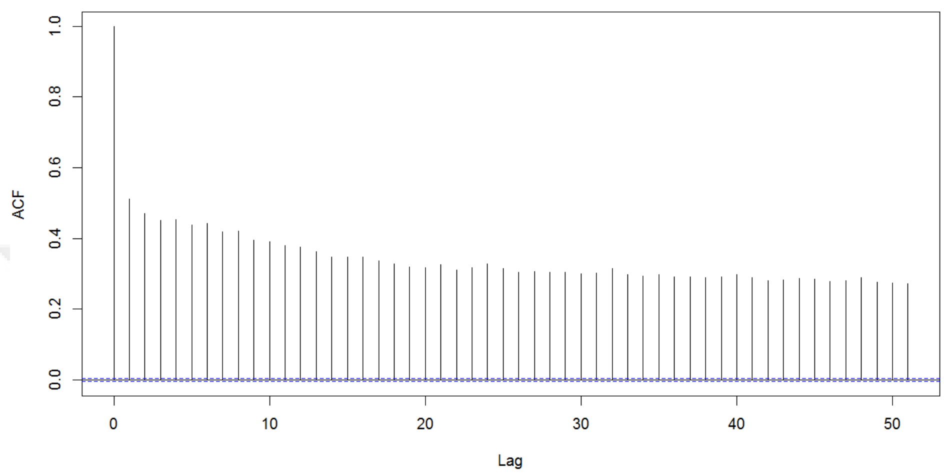


Figure 4.2| Autocorrelation function for the stream flow data of Bhagirathi River (1970-2004). Time series of climate and meteorological variables are frequently serially dependent, reflecting the strong persistence of the meteorological and climate phenomena. For example, a rainfall effect can be seen in multiple consecutive periods.

Same procedure is also applied to 6 years of solar radiation data of Delhi. Since solar radiation data did not show a significant variation between years, we take one year of data (2002) as a representative solar radiation data. Additive decomposition of solar radiation time series can be seen in Figure 4.3b. We also use one demand scenario which is the yearly demand load curve described in Chapter 3. Therefore, we have 13 scenarios that we can use with our model.

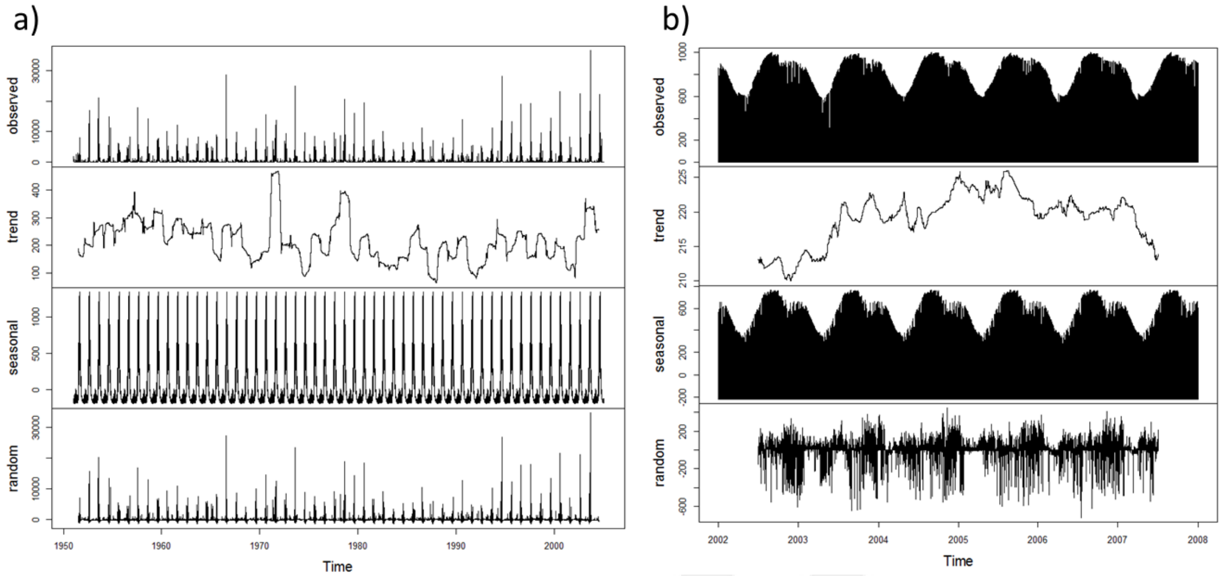


Figure 4.3| Decomposition of time series data. The seasonal fluctuations are roughly constant in size over time and do not seem to depend on the level of the time series. **a,** The components of the additive model for stream flow. **b,** The components of the additive model for solar radiation.

Table 4.4 | Scenarios determined from time series of stream flows

	Autumn (End of Monsoon)		**PreWinter (Post Monsoon)**		**Winter**		**Spring**		**Summer**		**Monsoon**	
	Mean	CV	Mean	CV	Mean	CV	Mean	CV	Mean	CV	Mean	CV
1970-1971	3678.1	1.1	639.9	0.4	391.3	2.0	715.2	2.8	4304.4	1.5	18120.7	1.0
1974-1975	844.8	0.5	229.5	0.9	663.5	2.7	609.3	1.8	884.1	3.4	7060.4	1.3
1977-1978	5700.1	0.9	913.0	1.1	454.0	1.7	1134.2	2.6	1210.8	1.4	12493.3	1.5
1978-1979	8165.8	1.7	927.5	0.8	1009.9	3.0	1050.3	1.7	826.4	1.9	6612.4	1.9
1979-1980	1228.4	0.9	251.6	0.4	339.5	3.4	314.5	2.6	845.0	1.5	10121.0	1.3
1981-1982	1210.7	0.8	529.9	3.0	586.5	3.1	907.1	2.0	641.1	1.5	4527.8	1.8
1985-1986	6250.0	1.0	1272.7	1.0	545.1	1.8	875.4	1.4	1184.1	2.5	7413.5	1.3
1986-1987	1814.8	0.8	596.4	1.2	494.0	2.2	747.0	1.6	1258.7	2.4	1613.3	2.7
1987-1988	1253.7	1.6	179.8	0.9	130.6	1.8	422.3	2.2	929.9	1.6	7444.8	1.1
1989-	3565.9	2.7	613.7	1.2	360.5	1.9	718.5	1.7	1297.8	1.6	7741.3	1.4

1990												
1990-1991	3261.1	0.9	761.0	1.2	1099.3	2.8	700.0	1.7	830.2	1.6	2719.1	2.0
1994-1995	1904.6	0.8	389.6	0.4	385.7	2.4	517.4	1.6	396.5	5.2	6392.1	1.8
2001-2002	1224.9	1.3	241.8	0.4	697.8	3.0	741.2	2.5	302.4	2.0	6075.1	2.9

4.4.3 Heuristic Approach

The scenario based linear program described in Section 4.4.1 can be a large scale program depending on the number of basins, demand points and the scenarios. Although number of constraints and variables increases linearly with the number of scenarios, solution time of the linear programming solvers is not linearly related to number of variables or constraints.

Given some characteristics of the data, the number of scenarios in the problem could be limited and then solved optimally using CPLEX 12.5 [29]. However, one might probably still need a heuristic approach when large scale instances of the problem cannot be solved in full space. With having short time periods in hourly level (3 hours), the number of time periods in a planning horizon can be high (2920 in one year planning horizon). This increases the problem size very quickly, especially when the scenarios are included into the problem to take the uncertainty into account. The reason why such short time increments (fine resolution) are used here is that solar radiation has daily cycle and the hourly variations are needed to be captured. Using longer timer increments has the effect of averaging out the peaks which are crucially important while designing more controllable power systems [30].

We have started with running the individual 1-year scenarios separately and the results are summarized in Figure 4.4. From these results, we can clearly conclude that only the upper reservoir size varies significantly for each individual scenario and the other components of the power system especially solar panel area and lower reservoir size do not vary a lot between scenarios. Next, we have developed our heuristic algorithm based on the fact that solar panel area and lower reservoir size are mostly dependent on solar radiation which does not significantly change between years.

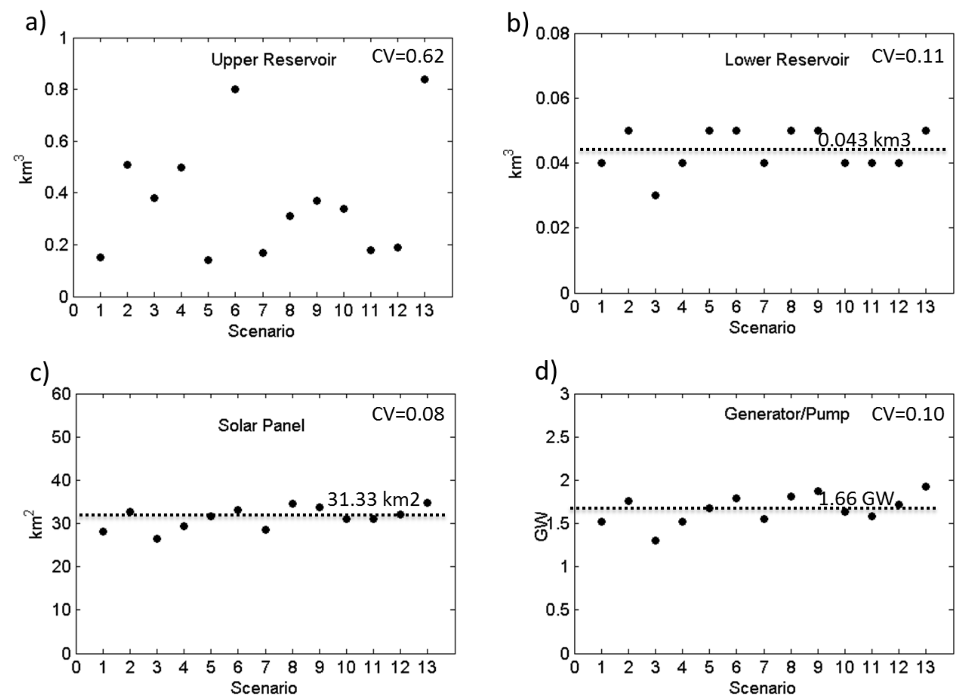


Figure 4.4| Results for the first step of the heuristic algorithm (each scenario is run individually). **a**, Only the upper reservoir size varies significantly for each individual scenario with 0.62 coefficient of variation (CV) . **b-d** the other components of the power system especially solar panel area, lower reservoir size do and generator capacity not vary a lot between scenarios (Coefficient of variations: 0.08, 0.11, 0.10 respectively) .

Our heuristic algorithm involves three steps. Running each scenario separately and estimating the solar panel area, lower reservoir size and amount of pumped water in each time period is the first step. In the second step, we refine our solution and run each scenario again with the fixed solar panel area and lower reservoir to have a more accurate estimate of how much water is pumped in a day/week with the specified variables. Then, in the third step, we combine all the scenarios together and run the multi-scenario model with a higher resolution (one day).

An important feature of our heuristic is that once we estimate the “solar energy related” components of the model such as solar panel area, lower reservoir and amount of energy pumped to upper reservoir for each scenario, we can change the resolution of our problem and run our model with computationally supportable longer time increments (one day or one week) to estimate the size of the upper reservoir. This is appropriate as upper reservoir size is highly dependent on the stream flows of various years (as can be seen in Figure 4.4 a) and it is the parameter of interest between different years.

We first test this idea for each scenario individually, without combining all scenarios together. We run each scenario with daily resolution with fixed (estimated) lower reservoir size, solar panel area and amount of pumped water to upper reservoir. Then we superimpose the results of the 3-hourly and daily runs, and observe a remarkable match. An example (Bhagirathi River and Delhi with the stream flow data of 1974 and 1980) is illustrated in Figure 4.5.

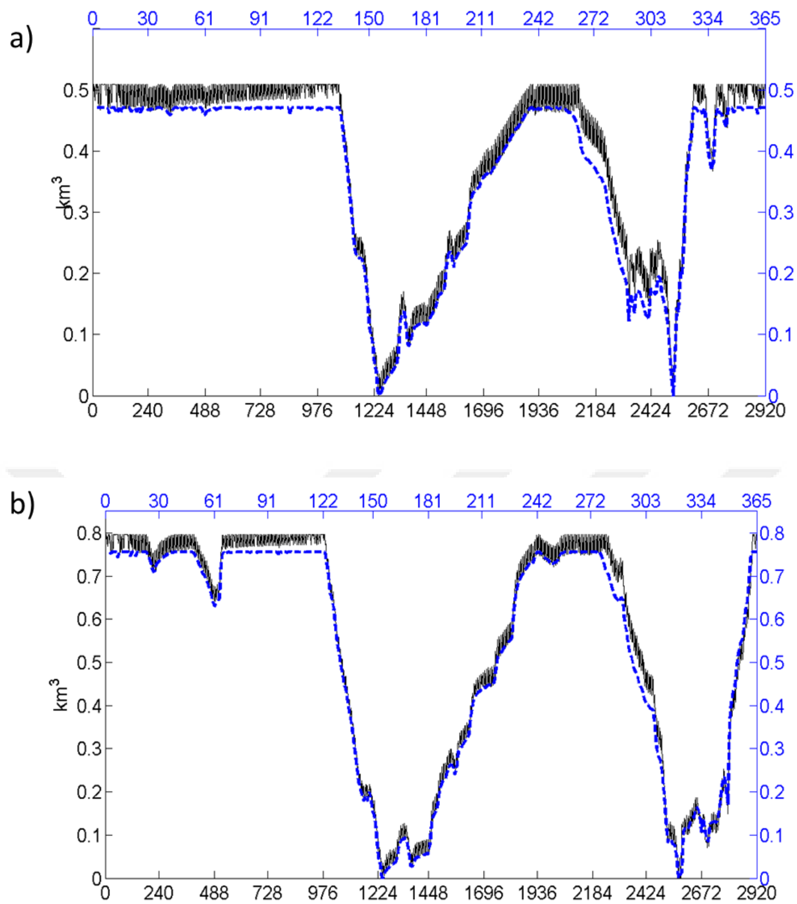


Figure 4.5| Superimposed results for 3hourly and daily run for a scenario**.** Once the "solar energy related" components of the model are estimated, the resolution of our problem can be changed (one day or one week) to estimate the size of the upper reservoir. **a**, 3hourly and daily individual run of year 1974. **b**, 3hourly and daily run of year 1981.

In summary, the main point of our heuristic is to estimate the 3-hourly fluctuations due to pumped water and then calculate the blue curve in Figure 4.5, which is determined mostly by seasonal variations of stream flow based on the estimated values (i.e. daily resolution is enough). It should be noted that our heuristic algorithm can underestimate the upper reservoir size. This is due to the fact that, we estimate some decision variables of our model, thus relax some constraints based on these variables. We should keep in mind that this means increasing the feasibility region (solution space) of our model and may lead to a lower cost than the optimal cost for our original problem. The intuitive way to explain this would be the following; when we increase the time period of the model (resolution) to one day and estimate the total pumped water within a day, we lose some details of the information such as the time pump operations peak within a day. Then, the system decides about release amount based on the aggregated pumped water which may provide more flexibility (lower objective for a minimization problem) to the system.

In order to see the effect of relaxing the constraints, we tested the performance of our heuristic algorithm by comparing it with the optimal solution of a *single basin-single demand point* case run with the scenarios described in the previous section. Table 4.5 summarizes the results for each investment variable. There is indeed a smaller size upper reservoir; however the values are in the same order and we can take the reservoir size of the heuristic result as the lower bound for the optimal solution. Moreover, using these solutions as initial points for solvers often leads to further improvements. We note that since the algorithm involves decomposition of each scenario and running them individually, it is suitable for implementation in parallel processors which results in a shorter running time and efficient memory usage.

Table 4.5 | Heuristic results compared optimal solution

	Heuristic	Optimal
Upper Reservoir	0.134	0.165
Lower Reservoir	0.045	0.042
Pump/Generator	1.600	1.624
Solar Panel	31.356	30.894
Transmission Line	1.600	1.624

Finally, as a validation of our scenario approach we perform a simulation analysis and run the data for all years (including the ones not chosen as a scenario here) fixing the investment decision variables to optimal values presented in Table 4.5. Figure 4.6 presents this result. Red dots are the years chosen as scenarios and blue dots are the rest of the years. We observe that the unit cost of the system varies between the bounds determined the scenarios. This result supports the fact that scenarios are chosen to represent extreme event years such as most variables, most dry, wettest etc.

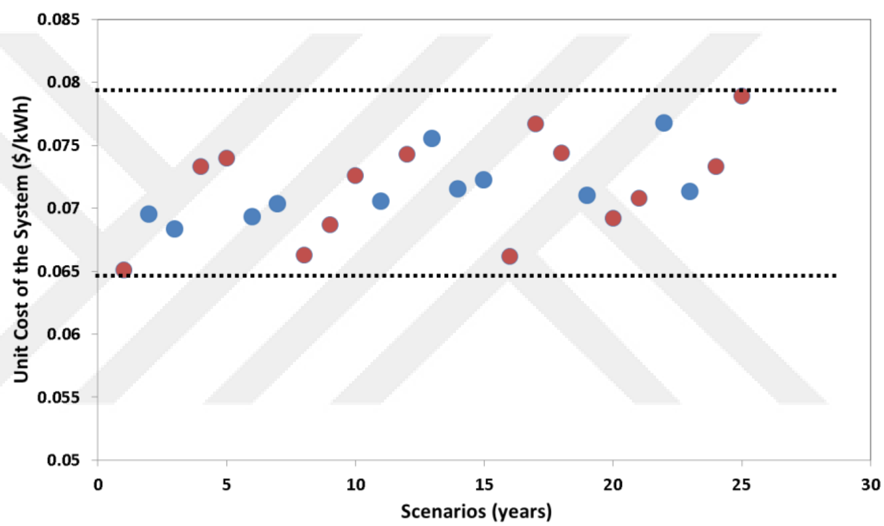


Figure 4.6| The data for each year is used to run the model fixing the investment decision variables to the values in the optimal solution to observe how the diesel usage changes over different years. Red dots are the years chosen as scenarios and we observe that the unit cost of the system varies between the bounds determined by the years chosen as scenarios. This result supports the fact that scenarios are chosen to represent extreme event years such as most variables, most dry, wettest etc.

4.5 Results

The goal of Chapter 3 and 4 of this thesis is to provide mathematical models that demonstrate how combining alternative sources (hybrid systems), energy storage and long distance transmission lines would handle the limited controllability of renewable sources. In Chapter 3, we showed how we can combine variety of clean energy sources in a conventional system to obtain a cost efficient and reliable. As explained in Section 4.3 and 4.4, in Chapter 4 we add a pump component to make the system further clean energy friendly. That's why we present our results here compared to the systems in Chapter 3 and talk about the enhancements due to the pumping component. Results in Chapter 3 may be reviewed together with these results.

Our linear program described in section 4.3 is implemented by following the approach in section 4.4, solved using CPLEX and the results are presented in a similar fashion with Chapter 3. We have multiple case studies for India to emphasize the different aspects of our results. We first present our results with a *single basin, single demand point* case study to illustrate the basic information that our pumped hydropower model can provide. Next, we introduce more basins and demand points to explain how the transmission and storage improvements can aid the integration of intermittent renewables as geographic aggregation smoothes the variability of the stream flows and solar radiations and how integration of additional demand points can change the balance in the system. Then, we calculated the results of our algorithm for all the data we have for India (8 basins and 10 demand points mostly in Northern India and Bengal) and observe how our model works in a real time complicated nationwide network. Cost parameters that are used in analysis are the same as in Chapter 3. We assess the sensitivity of the system in terms of cost parameters in the next section.

4.5.1 Case Study I: Single Basin, Single Demand Point Case

This is the section where we show the benefits of having a lower reservoir and pump some of the water to save the excessive solar power in form of potential energy. We use a simple *single basin, single demand point* case for this purpose. We first illustrate the water stored in the upper and lower reservoirs together with components that determine the upper reservoir level. Then, we discuss a one week zoomed operation balance in reservoir and demand points. We finally compare the results with the conventional system in Chapter 4 including how we should consider the residence time of the reservoirs in this system.

Data used in Chapter 3 for illustration of *single basin, single demand point* case is also used here in order to be able to show the benefits of pumped hydro model explicitly and compare the results with conventional hydropower system case easily. The sample one year data with 13.3 km^3 annual inflow is taken from the time series data of stream flow for Bhagirathi River calculated by VIC land surface model. Solar radiation data of Delhi for the year 2002 is collected from National Renewable Energy Laboratory website. 3-hourly demand data for Delhi is estimated as described in previous chapter. To be able to scale supply potential and demand, a low demand profile for Delhi is created by dividing actual demand by four and results are presented based on the scaled data. The input data used in this section is shown in Figure 4.

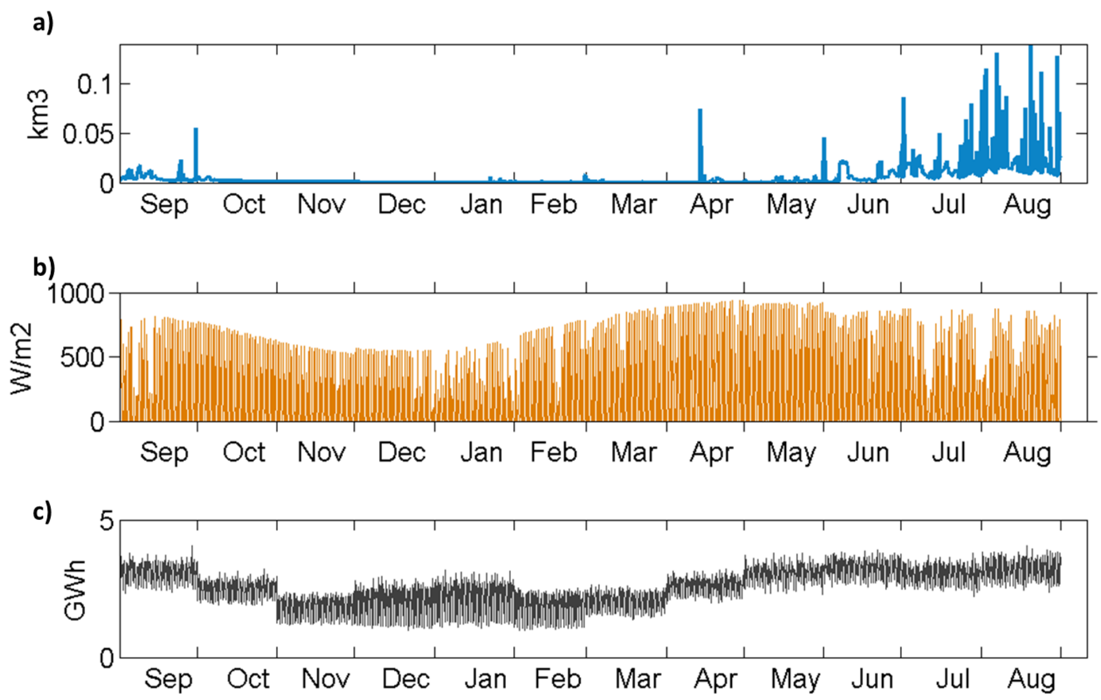


Figure 4.7 | Input data used in the model **-** Data used in Chapter 3 for illustration of single basin, single demand point case is also used here in order to be able to show the benefits of pumped hydro model.

Figure 4.8a-b shows the water level stored in the upper and lower reservoirs during the operation for one year. We start and end the operations with full upper reservoirs due to the fact that our planning horizon starts in September (almost the end of the Monsoon season). The fluctuations in the upper and lower reservoirs support a daily cycle in pumped hydro operations. As can be seen in Figure 4.8, there is no water stored in the lower reservoir in the Monsoon season since there is no need to pump water to upper reservoir. We can also conclude that water stored in the lower reservoir (water to be pumped with the extra solar power) during winter season closely follow the solar radiation curve (Figure 4.7b). During the dry season (~ between November and March), upper reservoir is highly utilized: Since the model “see” that rainy season (high inflow) coming up (violating nonanticipativity condition), water is stored in dry season and is used starting from March. In addition, as there is also high solar radiation in spring-summer period that can be used to pump more water, we observe that there is more water in the lower reservoir until Monsoon seasons starts again. This result can be verified from the components of the mass balance equations of the upper reservoir such as inflow, released, pumped and spilled water which are presented in Figure 4.9. The pumped water to upper reservoir is so high that the pumped amount is limited by the generator size as can be understood from the flatness between February and June.

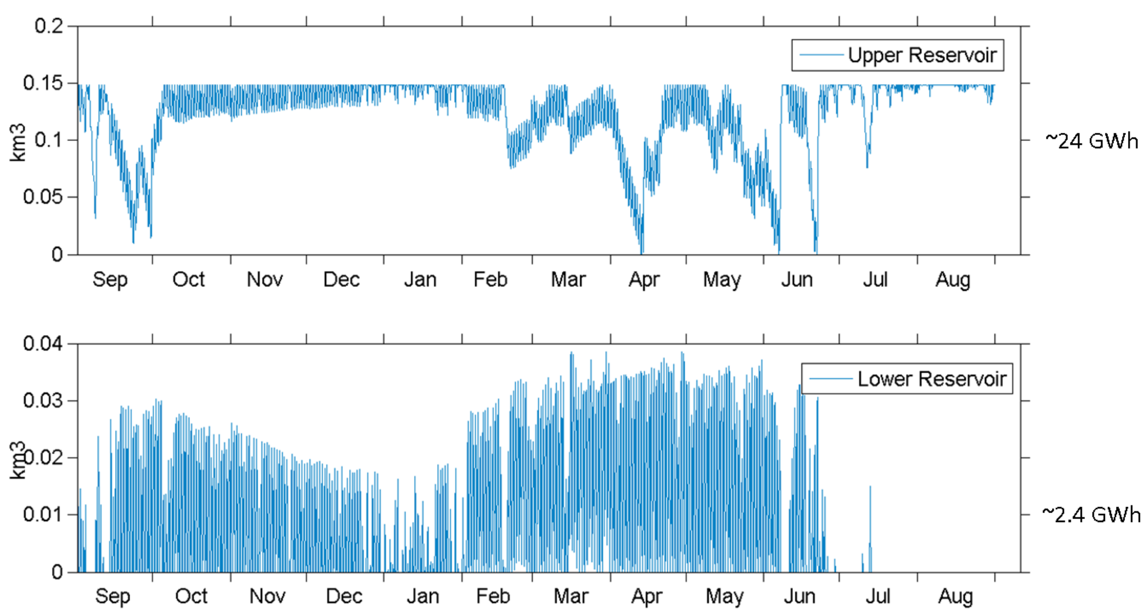


Figure 4.8| Water stored in the upper and lower reservoirs – The upper reservoir is assumed to be full at the start and the end of the cycle as the planning horizon starts in September which is close to end of the Monsoon season. There is no water stored in the lower reservoir in the Monsoon season since there is no need to pump water to upper reservoir. (0.01 km3 ~ 2.4 GWh)

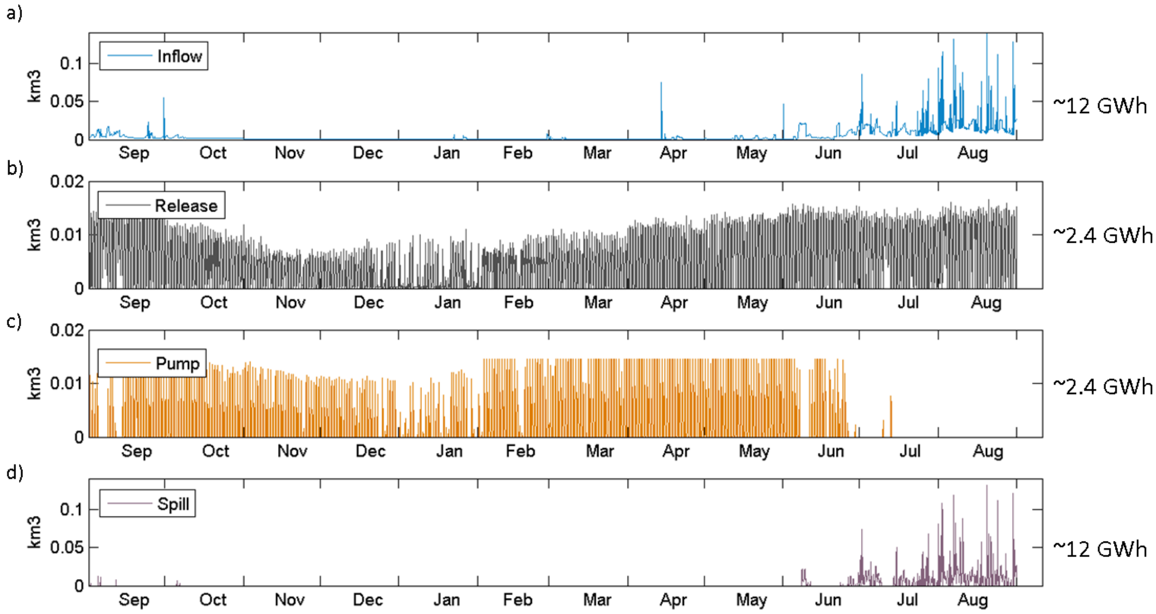


Figure 4.9| Flows from/to the upper reservoir level - Water is stored in dry season (November to March) and is utilized starting from March. As there is high solar radiation in spring-summer period that can be used to pump more water, lower reservoir utilization is increased until Monsoon seasons starts again. (0.01 km^3 ~ 2.4 GWh)

Furthermore, we now zoom into one week in March and see the balance between solar, hydro and diesel components to meet the demand and the results are shown in Figure 4.9. Figure 4.10a illustrates that solar energy is enough to meet the demand during the day and hydro becomes effective when the Sun does not shine (at night). Hydro and solar together satisfies the demand during the week. Figure 4.10b shows how the reservoir operates for the same week. Due to the inflow observed in the first day of the week, extra solar energy generated in the demand point is not needed to pump to upper reservoir in the first two days.

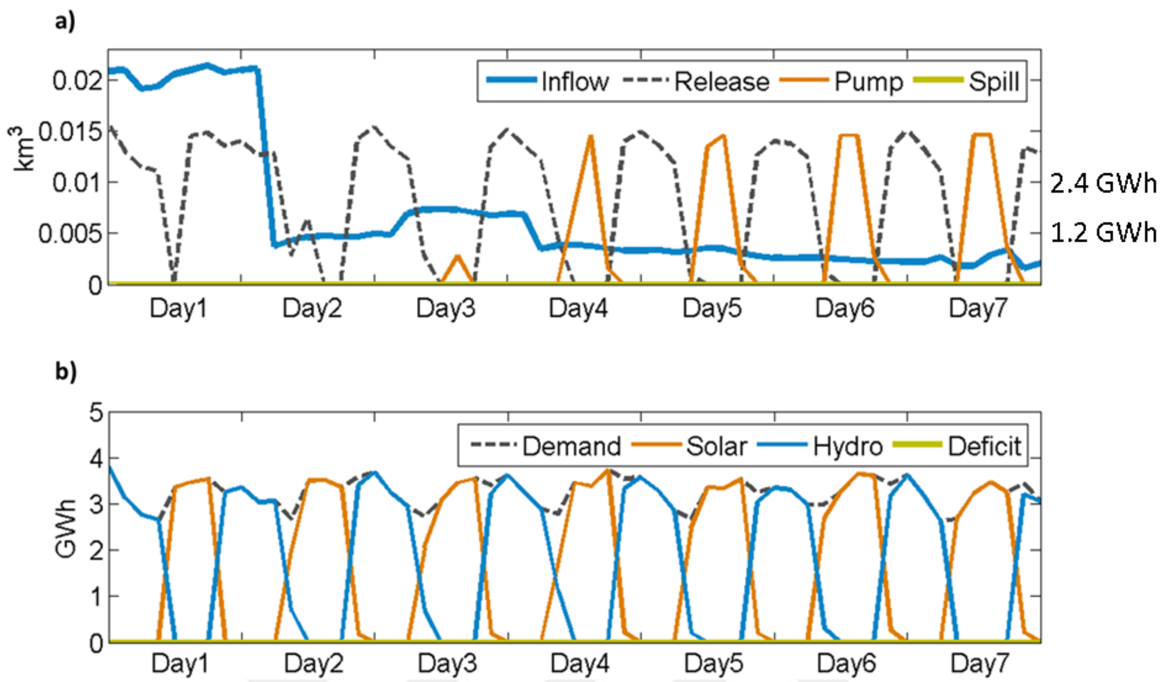


Figure 4.10 | One week operation balance in reservoir and demand points – **a,** Solar energy is enough to meet the demand during the day and hydro becomes effective at night. Deficit is zero during all week. **b,** Due to the inflow observed in the first day of the week, extra solar energy generated in the demand point is not necessary to pump to upper reservoir in the first two days. No spill is observed during all week.

Next, the detailed distribution of alternative sources to meet the demand is summarized in Table 4.6. 47% of the demand is met by hydropower. 25% of this is directly generated from the inflow to the reservoir whereas, 22% is generated using the pumped water. Solar energy generated 84 percent of the annual demand, of which 45% is directly used as “internal” solar energy and 28% is sent to reservoirs to be stored. The 6% difference between the hydro energy generated from the pumped water and solar energy sent to reservoirs stems from the generator and pump efficiencies.

Table 4.6| Distribution of resources to meet the demand

		GWh	% of Demand
Hydro Production	Inflow	1864.6	25%
	Solar	1652.4	22%
	SUM	3517.0	47%
Solar Production	Internal	3402.6	45%
	Pump	2133.7	28%
	Spill	779.2	10%
	SUM	6315.5	84%
Diesel		583.8	8%
Demand		7503.3	100%

As we now have the results for the pumped hydro system for the same case study discussed in previous chapter, we can expand the comparison of alternative systems results to include pumped hydro system. Figure 4.11 clearly shows that compared to other hybrid systems the system with pumped hydro storage is the most cost efficient design with much lower unit cost. Having two-level reservoir with a pump system reduces the intermittency effect of renewable sources and the system can utilize more clean energy. In particular, total hydro production increases to 47% (it was 29% in conventional system in Chapter 3) with even smaller reservoir size compared to conventional system. With an addition of 0.04 km^3 lower reservoir, upper reservoir size is reduced to 0.15 km^3 (0.29 km^3 in the conventional system). Further details about the efficiencies, capacity factors and unit cost of the systems are summarized below in Table 4.7.

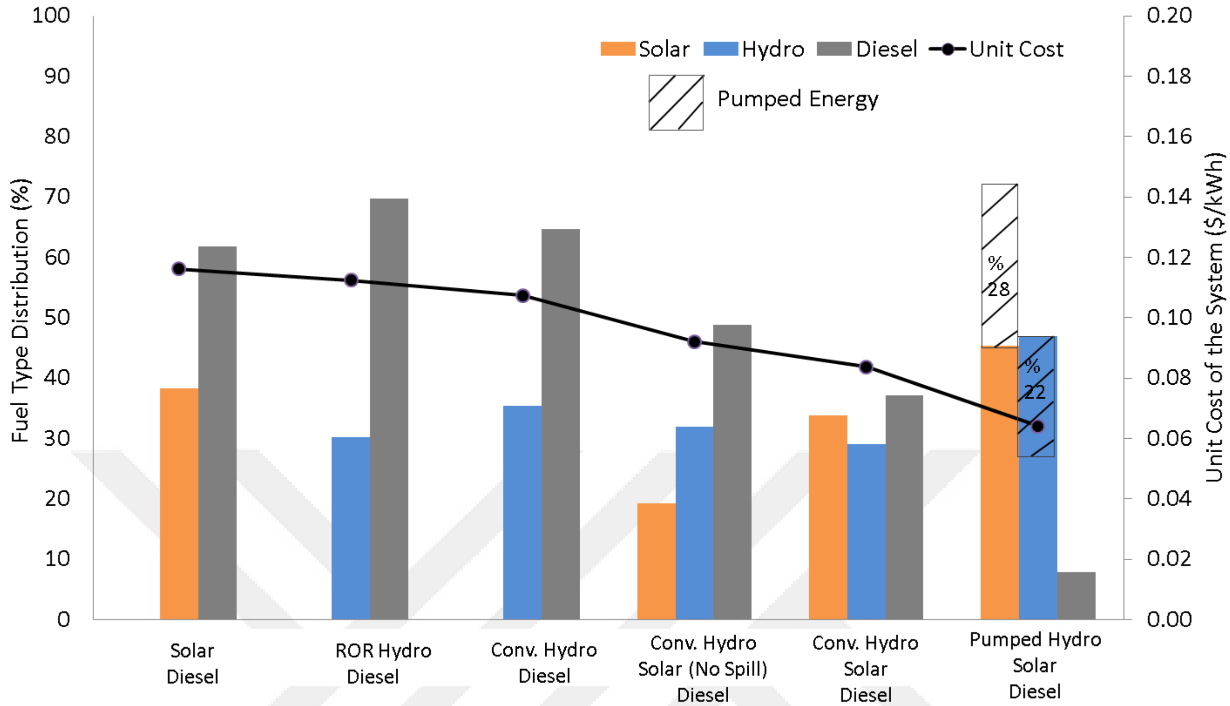


Figure 4.11| Alternative Technologies - Compared to other hybrid systems, the system with pumped hydro storage is the most cost efficient design with much lower unit cost. Having two-level reservoir with a pump system reduces the intermittency effect of renewable sources and the system can utilize more clean energy. In particular, total hydro production increases with even smaller reservoir size compared to conventional system. Shaded area in the solar power bar is transferred to hydro power.6% is lost due to generator and pump efficiencies.

Table 4.7| Results of the pumped hydro system compared to conventional system

	Conventional Hydro		Pumped Hydro	
Investments				
Upper Reservoir Size(km^3/~GWh)	0.29/~70		0.15/~36	
Lower Reservoir Size(km^3/GWh)	N/A		0.04/~9	
Pump/Generator Size(GW)	1.06		1.50	
Solar Panel Size(km^2/~GW)	12.50/~0.30		28.18/~0.68	
Production	GWh	% Demand	GWh	% Demand
Solar (Internal)	2533	33.76	3224	45.35
Hydro	2184	29.11	3695	46.87
Diesel	2786	37.13	584	7.78
Total	7503	100	7503	100

Diesel		
Average Diesel Usage (GW)	0.32	0.06
Peak Diesel Usage (GW)	1.29	1.08

Demand	
Average Demand (GW)	0.86
Peak Demand (GW)	1.35
Load Factor	0.63

Solar		
Efficiency	90%	88%
Average Solar Production (GW)	0.32	0.72
Peak Solar Production (GW)	1.41	3.18
Capacity Factor	0.23	0.23

Hydro	Turbine	Pump	Turbine
Average Hydro Production (GW)	0.25	0.21	0.40
Peak Hydro Production (GW)	1.06	1.50	1.50
Capacity Factor	0.24	0.14	0.27

Transmission	Hydro	Hydro	Solar
Average Transmission (GW)	0.25	0.40	0.24
Peak Transmission (GW)	1.06	1.50	1.50
Capacity Factor	0.23	0.27	0.16

Unit Cost		
Unit Cost of Hydro ($/kWh)	0.04	0.02
Unit Cost of Solar ($/kWh)	0.05	0.05
Unit Cost of Diesel ($/kWh)	0.15	0.15
Unit Cost of the System ($/kWh)	0.08	0.06

We now compare the pumped hydro system with the conventional hydro system and see that how the residence time of the reservoir went down (water has been cycled more), how the solar contributed both to hydropower (pumping) and to internal usage (larger panels) and as a result diesel usage decreased to 8% from 37%.

In pumped hydro system, there are two level reservoirs in the system with more complicated working principal than the conventional system, we revisit the term, “residence time of the reservoir” that is introduced in Chapter 4. The operation of the upper reservoir in our pumped hydro system is no different than the reservoir in the conventional hydro system. As pumped water can only be considered as an additional inflow to the upper reservoir, we do not need to change our definition of residence time for the upper reservoir. However, we need to define the residence time for the lower reservoir which practically has a quite different operation cycle compared to the upper reservoir. To estimate the average time a water molecule spends in the lower reservoir, we again rely on the conservation of mass principle and estimate the residence time of the reservoir by dividing the volume of the reservoir by the rate by which water either gets filled or depleted the reservoir [31]. Here, the rate by which the water is pumped to upper reservoir represents the rate by which the water exits the lower reservoir. Therefore, residence time of upper and lower reservoirs, $\boldsymbol{RT\_upper_i^{\omega}}$ and $\boldsymbol{RT\_lower_i^{\omega}}$ respectively, in hydropower station i under scenario $\boldsymbol{\omega}$ can be calculated by the formulas below:

$$\boldsymbol{RT\_upper_i^{\omega} = SUmax_i * 365 / \sum_t R_i^{\omega t}}$$

$$\boldsymbol{RT\_lower_i^{\omega} = SLmax_i * 365 / \sum_t P_i^{\omega t}}$$

where $\mathrm{SUmax_i}$ and $\mathrm{SLmax_i}$ are the size of the upper an d lower reservoirs respectively , $\sum_t \mathrm{R_i^{\omega t}}$ represents the total amount of water released from the upper reservoir i and $\sum_t \mathrm{P_i^{\omega t}}$ represents the total amount of pumped water from lower reservoir to upper reservoir in a year (our time horizon) under scenario ω. Using the formula above, residence times of the reservoirs are estimated to be 3.6 and 2 days for upper and lower respectively. We note that in Chapter 3, residence time of the reservoir in conventional hydro system with the same demand data is calculated as 11.7 days. This means we can cycle water a few times and generate more power from renewable sources. In Table 4.7, detailed results of the pumped hydro system are given compared to conventional hydro system. Results show that with an additional 0.04 km^3 lower reservoir, a smaller upper reservoir (almost the half of the one in conventional system (0.15 km^3 vs 0.29 km^3) is needed. The flexibility that pumped hydro system brings combining hydro and solar potentials provides reduced residence time of upper reservoir. Table 4.7 shows that introducing the pump to the system reduced the diesel usage from 37 percent to 7 percent and decreased the unit cost of the overall system.

Another interesting result demonstrated in Table 4.7 and Figure 4.12 is the increase in the percentage of internally used solar energy in pumped hydro system, compared to conventional system. This is mainly due to the fact that it is expected that solar panel area in pumped hydro system is larger than the area in conventional system as the role of solar energy in pumped system is twofold: internal solar and pumped solar. In addition, in the pumped hydro system, solar energy is usually transmitted to hydro stations for pumping for two consecutive time periods (total of 6 hours) in one day on average. However, solar radiation is available for more time periods and as the solar panel area is larger, the energy generated during the day when there is no pumping, is being used internally. Therefore increased solar panel area also contributes the internally used solar energy. In particular, in Table 4.7, we see that 28% of solar generated in the pumped hydro system is being pumped and 45% is being used internally within demand points whereas, in conventional system, the internally used solar energy is 33%. Figure 4.12 compares conventional and pumped hydro systems in terms of the solar energy produced in one day. It can be seen that total production is scaled up by increased solar panel area as the solar radiation curve is same. Solar energy used internally for the 4^{th} and 5^{th} time periods are the same as extra solar energy is spent for pumping, however shaded area in 3^{rd} and 6^{th} time periods in Figure 4.12b represents the parts which explain the difference between solar internal percentages in conventional and pumped systems. This consistently supports the explanation above. The fact that we still have some spill in the solar power in the pumped hydro system is due to optimized size in the high demand period as explained in Chapter 3. Another way to see the difference is presented in Figure 4.13. Internally used solar energy for one week in November is presented.

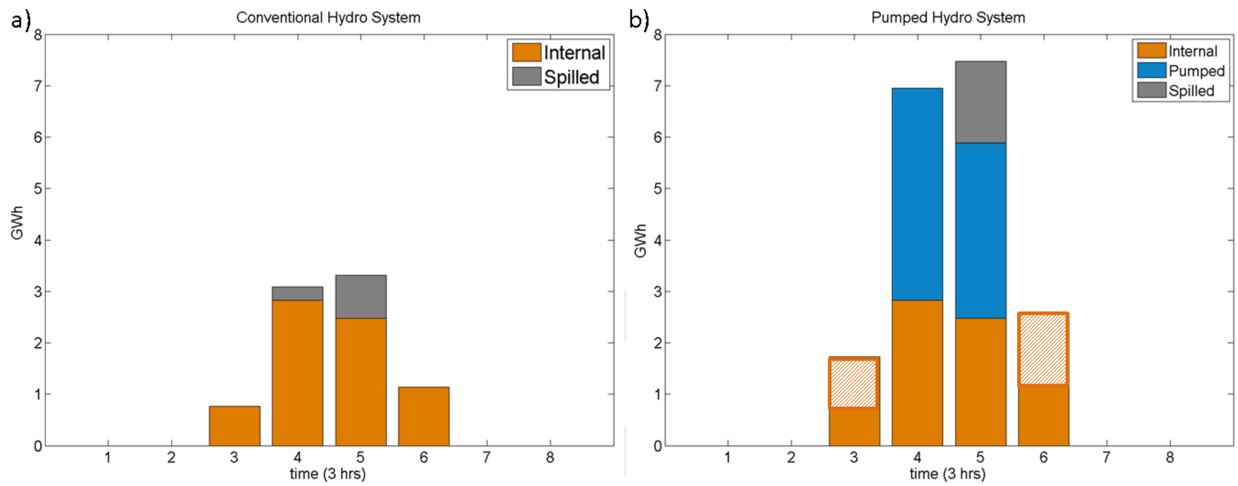


Figure 4.12 | Comparisons of solar production profile of one day in November for conventional **(a)** and pumped hydro systems **(b) -** Total production is scaled up by increased the solar panel area as the solar radiation curve is same. The role of solar energy in pumped system is twofold: internal solar and pumped solar. Solar energy used internally for the 4th and 5th time periods are the same between two systems as extra solar energy is spent for pumping, however shaded area in 3rd and 6th time periods in (**b)** represents extra solar internal.

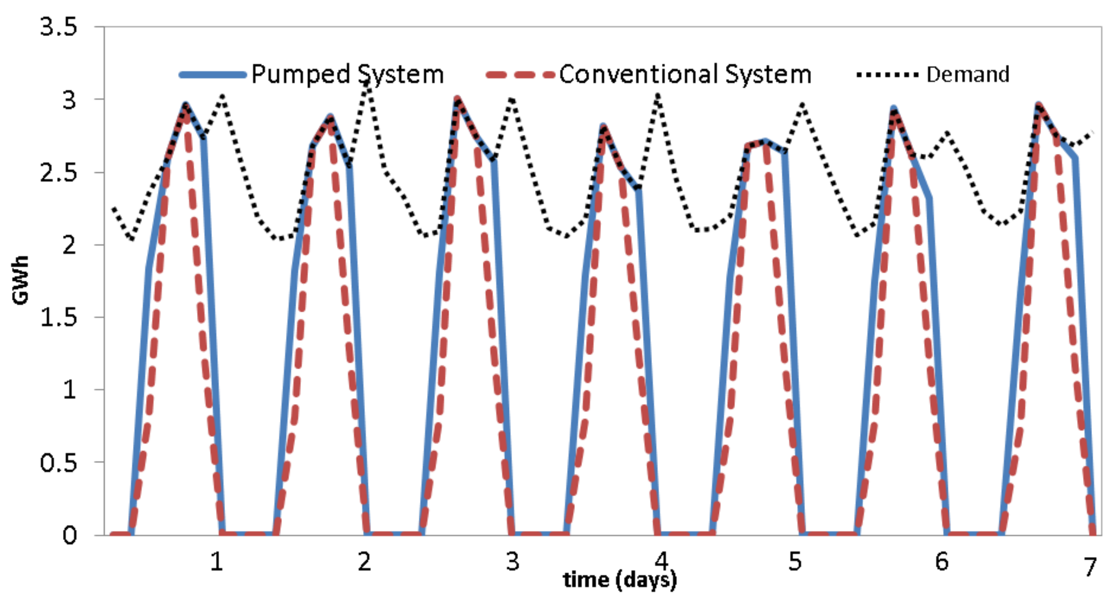


Figure 4.13 | Comparisons of internally used solar energy for one week in November. The area under curves explains the difference between solar proportions to meet the demand in Table 4.7.

4.5.2 Case Study II: Multi Demand Points, Multi Basins System

In Chapter 3 and 4.5.1, we have shown that combining variety of resources gives us higher cost efficiency (Figure 4.11) and how a pumped hydro storage system increases the flexibility of intermittent sources in matching supply and demand. Now, in this section we provide results for the effect of long distance transmission lines between resources which are expected to have different variability. As individual systems are connected with each other through transmission lines and work as a single system, the variability and intermittency of renewable sources is expected to be smoothed out. To demonstrate this effect, we sequentially add more basins and demand points to the system that we discussed in previous section.

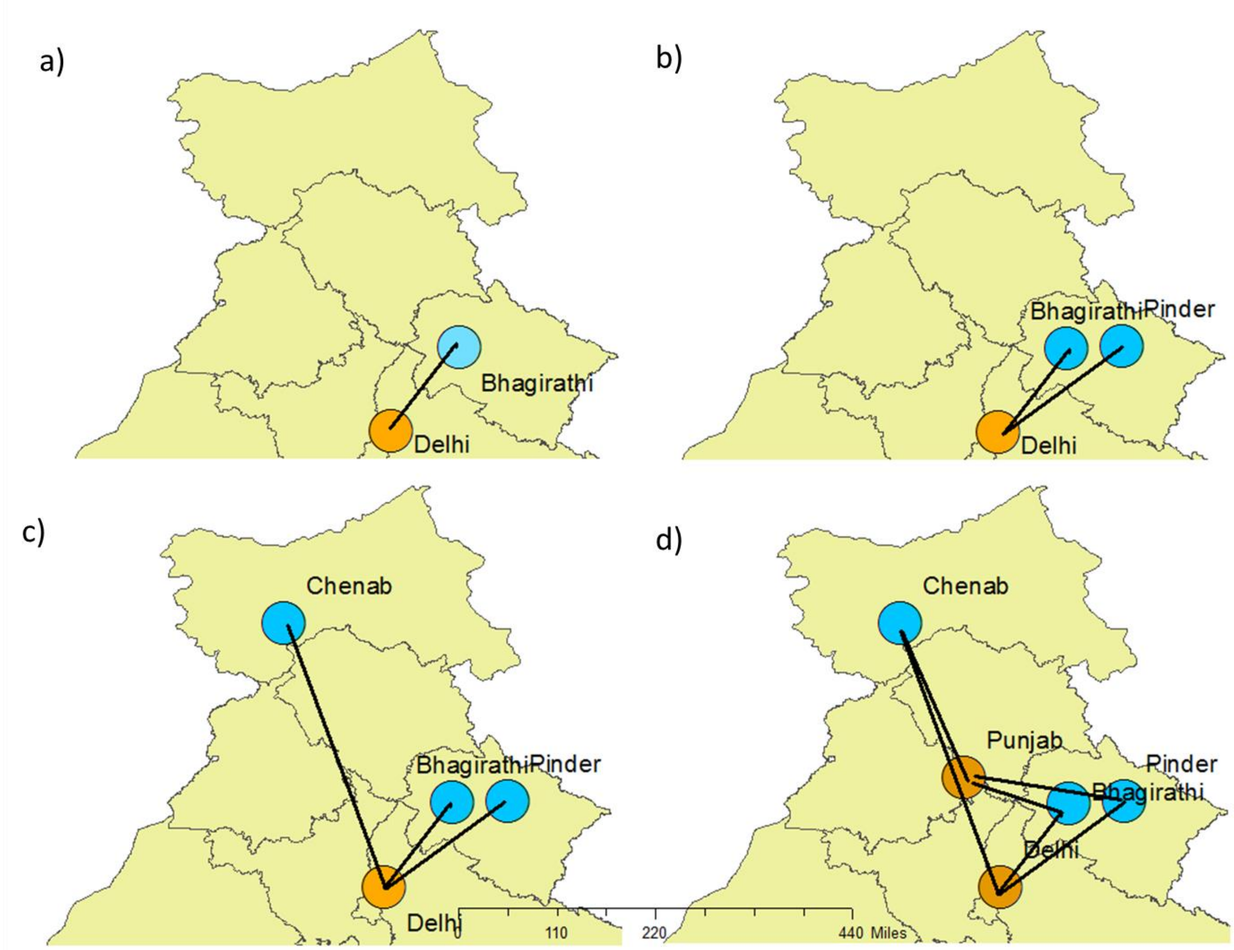


Figure 4.14 | Multi-basins, Multi-demand Points Cases – **a,** is the system discussed in Section 5.5.1. **b,** introduces an additional basin (Pinder River with 6.2 km3 annual inflow) to the system and reduces the unit cost. **c,** Higher flexibility from the variety of resources is observed when third basin (Chenab River with 45.5 km^3) is included. **d,** Another demand point (Punjab - 10000GWh annual demand in the calculations) is added and observed the balances between reservoirs change when the demand increases.

We start with the system in Figure 4.14a. With 13.3 km^3 annual inflow of Bhagirathi River and 7500 GWh demand size (one fourth of actual demand of Delhi) is discussed in detail in Section 5.5.1. Then, we first add Pinder River with 6.2 km^3 annual inflow to the system (Figure 4.14b). Results showing the comparison between two cases are summarized in Table 4.8.

The immediate result is the reduced unit cost of combined system (Table 4.8). In further details, one important point is that although the distance between Pinder and Delhi is longer than the distance between Delhi and Bhagirathi (Figure 4.14b), it is still cost efficient to reduce the size of a shorter transmission lines between Delhi and Bhagirathi and add another transmission line for 0.21 GW between Delhi and Pinder. Moreover, total generator size in latter case is smaller than the *single basin-single demand point* case. This indicates that combining two sources with different variability increases the flexibility of obtaining cheaper systems.

Table 4.8 | Bhagirathi and Pinder Rivers combined to meet Delhi's demand

	Pinder + Bhagirathi		Bhagirathi
Investment Size	Pinder	Bhagirathi	
Upper Reservoir Size(km^3/~GWh)	0.032/7.6	0.093/22.4	0.148/36.6
Lower Reservoir Size(km^3/~GWh)	0.004/1.0	0.028/6.8	0.0385/9.4
Pump/Generator Size(GW)	0.216	1.139	1.499
Transmission Line to Delhi (GW)	0.216	1.139	1.500
Solar Panel Size(km^2 /~GW)	24.98/0.60 GW		28.18/0.68

		GWh	% Demand	GWh	%Demand
Hydro Production I Devsari	Inflow	577.4	8%		
	Solar	105.1	1%		
	SUM	682.5	9%		
Hydro Production II Tehri	Inflow	1758.8	23%	1864.6	25%
	Solar	1180.0	16%	1652.4	22%
	SUM	2938.9	39%	3517.0	47%
Solar Production	Internal	3291.9	44%	3402.6	45%
	Pump	1659.5	22%	2133.7	28%
	Spill	645.5	9%	779.2	10%
	SUM	5597.0	75%	6315.5	84%
Diesel		590.0	8%	583.8	8%
Demand		7503.3	100%	7503.3	100%

Diesel		
Average Diesel Usage (GW)	0.07	0.06
Peak Diesel Usage (GW)	0.91	1.08

Demand	
Average Demand (GW)	0.86
Peak Demand (GW)	1.35
Load Factor	0.63

Solar		
Efficiency	0.885	0.880
Average Solar Production (GW)	0.64	0.72
Peak Solar Production (GW)	2.82	3.18
Capacity Factor	0.23	0.23

Hydro Production I (Pinder)	Pump	Turbine	
Average Hydro Production (GW)	0.01	0.08	
Peak Hydro Production (GW)	0.216	0.216	
Capacity Factor	0.06	0.36	

Hydro Production II (Bhagirathi)	Pump	Turbine	Pump	Turbine
Average Hydro Production (GW)	0.15	0.34	0.21	0.40
Peak Hydro Production (GW)	1.14	1.14	1.50	1.50
Capacity Factor	0.13	0.29	0.14	0.27

Unit Cost		
Unit Cost of Hydro ($/kWh)	0.016	0.019
Unit Cost of Solar ($/kWh)	0.054	0.054
Unit Cost of Diesel ($/kWh)	0.150	0.150
Unit Cost of the System ($/kWh)	0.059	0.064

Higher flexibility from the variety of resources can be observed with further clarity when Chenab River with 45.5 km^3 is included in the system (Figure 4.14c). Table 4.9 summarizes the results and shows that Chenab River with a high stream flow potential reduced the unit cost of the system by around 42% (from 59 cents to 34 cents). Interestingly, due to the high potential of the river, a conventional hydropower station is found to be sufficient for Chenab River. In addition, the new source has also affected to Pinder Hydropower Station and a pumped hydro system is no longer needed. In comparison with the previous case, the sum of the reservoirs, generators and transmission sizes are all reduced. As more potential is added to the system, hydro power got particularly cheaper reducing the proportion of solar in meeting the total demand. Next, it should be noted that when all cases are compared, we observe that proportion of diesel is about 8% and could not be reduced. This is because of the fact that although basins have different potentials with short term variability, they have similar seasonal variability. In winter, the stream flow potential and solar radiation decreases and diesel usage becomes compulsory to provide 100% reliability.

Table 4.9 | Chenab, Pinder and Bhagirathi River Systems

	Chenab + Pinder + Bhagirathi			Pinder+ Bhagirathi	
Investment Size	Chenab	Pinder	Bhagirathi	Pinder	Bhagirathi
Upper Reservoir Size(km^3/~GWh)	0.010/2.5	0.005/1.2	0.019/4.5	0.032/7.6	0.093/22.3
Lower Reservoir Size(km^3/~GWh)	0.000	0.000	0.009/2.2	0.004/1.0	0.028/6.8
Pump/Generator Size(GW)	0.665	0.103	0.355	0.216	1.139
Transmission Line to Delhi (GW)	0.665	0.103	0.355	0.216	1.139
Solar Panel Size(km^2/~GW)	8.82/~0.21			24.98/~0.60	

Production		GWh	% Demand	GWh	% Demand
Hydro Production I Chenab	Inflow	3619.3	48%		
	Solar	0.0	0%		
	SUM	3619.3	48%		
Hydro Production II Pinder	Inflow	349.0	5%	577.4	8%
	Solar	0.8	0%	105.1	1%
	SUM	349.7	5%	682.5	9%
Hydro Production III Bhagirathi	Inflow	1026.4	14%	1758.8	23%
	Solar	73.3	1%	1180.0	16%
	SUM	1099.7	15%	2938.9	39%
Solar Production	Pump	95.6	1%	1659.5	22%
	Spill	44.9	1%	645.5	9%
	SUM	1976.9	26%	5597.0	75%
Diesel		598.3	8%	590.0	8%
Demand		7503.3	100%	7503.3	100%

Diesel		
Average Diesel Usage (GW)	0.07	0.07
Peak Diesel Usage (GW)	0.83	0.91

Demand	
Average Demand (GW)	0.86
Peak Demand (GW)	1.35
Load Factor	0.63

Solar		
Efficiency	98%	88%
Average Solar Production (GW)	0.23	0.72
Peak Solar Production (GW)	0.99	3.18
Capacity Factor	0.23	0.23

Hydro Production I (Chenab)	Pump	Turbine	
Average Hydro Production (GW)	0.00	0.41	
Peak Hydro Production (GW)	0.67	0.67	
Capacity Factor	0.00	0.62	

Hydro Production II (Pinder)	Pump	Turbine	Pump	Turbine
Average Hydro Production (GW)	0.00	0.04	0.01	0.08
Peak Hydro Production (GW)	0.10	0.10	0.22	0.22
Capacity Factor	0.00	0.39	0.06	0.36

Hydro Production III (Bhagirathi)	Pump	Turbine	Pump	Turbine
Average Hydro Production (GW)	0.01	0.13	0.15	0.34
Peak Hydro Production (GW)	0.36	0.36	1.14	1.14
Capacity Factor	0.03	0.35	0.13	0.29

Unit Cost		
Unit Cost of Hydro ($/kWh)	0.008	0.016
Unit Cost of Solar ($/kWh)	0.049	0.054
Unit Cost of Diesel ($/kWh)	0.150	0.150
Unit Cost of the System ($/kWh)	0.034	0.059

Finally, another demand point with annual demand around 10000GWh (one half of Punjab's demand) is added to system in order to observe how the balances between reservoirs change when the demand increases (Figure 4.14d). These results are presented in Table 4.10 and it can be seen that Chenab and Pinder have now pump hydro stations. Therefore transmission lines in the system clearly distribute the renewable sources eliminating variability and intermittency.

Table 4.10 | Another demand point (Punjab) added to the system

	Delhi +Punjab			Delhi		
Investment Size	Chenab	Pinder	Bhagirathi	Chenab	Pinder	Bhagirathi
Upper Reservoir Size(km3/~GWh)	0.032/7.7	0.015/3.5	0.168/0.17	0.01/2.5	0.005/1.19	0.019/4.5
Lower Reservoir Size(km3/~GWh)	0.012/2.8	0.004/0.99	0.039/9.4	0.000	0.00/0.05	0.009/2.19
Pump/Generator Size (GW)	1.08	0.21	1.60	0.67	0.10	22.46
Transmission Line to						
Delhi (GW)	0.20	0.21	0.76	0.67	0.10	0.36
Punjab (GW)	0.88	0.00	0.98			
Solar Panel Size						
Delhi (km2/~GW)	19.06/0.46			8.82/0.21		
Punjab (km2/~GW)	22.46/0.54					

		Delhi +Punjab GWh	Delhi +Punjab % of Total	Delhi GWh	Delhi % of Total
Hydro Production I Chenab	Inflow	4988.4	27.40%	3619.3	48.24%
	Solar	189.1	1.04%	0.0	0.00%
	SUM	5177.4	28.44%	3619.3	48.24%
Hydro Production II Pinder	Inflow	741.4	4.07%	349.0	4.65%
	Solar	99.4	0.55%	0.8	0.01%
	SUM	840.8	4.62%	349.7	4.66%
Hydro Production III Bhagirathi	Inflow	2343.9	12.87%	1026.4	13.68%
	Solar	1524.3	8.37%	73.3	0.98%
	SUM	3868.2	21.25%	1099.7	14.66%
Solar Production I Delhi	Internal	2946.9	16.19%	1836.4	24.48%
	Pump	1097.9	6.03%	95.6	1.27%
	Spill	225.7	1.24%	44.9	0.60%
	SUM	4270.4	23.46%	1976.9	26.35%
Solar Production II Punjab	Internal	3510.7	19.28%		
	Pump	1243.0	6.83%		
	Spill	150.9	0.83%		
	SUM	4904.6	26.94%		

Diesel	Delhi- Punjab		Delhi
Average Diesel Usage (GW)	0.1	0.1	0.07
Peak Diesel Usage (GW)	0.9	1.3	0.83

Demand	Delhi- Punjab		Delhi
Average Demand (GW)	0.86	1.22	0.86
Peak Demand (GW)	1.35	2.15	1.35
Load Factor	0.63	0.57	0.63

Solar	Delhi- Punjab		Delhi	
Efficiency	95%	96%	98%	
Average Solar Production (GW)	0.49	0.56	0.23	
Peak Solar Production (GW)	2.15	2.45	0.99	
Capacity Factor	0.23	0.23	0.23	
Hydro Production I (Chenab)	Pump	Turbine	Pump	Turbine
Average Hydro Production (GW)	0.02	0.59	0.00	0.00
Peak Hydro Production (GW)	1.08	1.08	0.67	0.67
Capacity Factor	0.02	0.55	0.00	0.00
Hydro Production II (Pinder)	Pump	Turbine	Pump	Turbine
Average Hydro Production (GW)	0.10	0.10	0.00	0.00
Peak Hydro Production (GW)	0.21	0.21	0.10	0.10
Capacity Factor	0.46	0.46	0.00	0.00
Hydro Production III (Bhagirathi)	Pump	Turbine	Pump	Turbine
Average Hydro Production (GW)	0.20	0.44	0.01	0.00
Peak Hydro Production (GW)	1.60	1.60	0.36	0.36
Capacity Factor	0.12	0.28	0.03	0.00

4.5.3 Case Study III: Multi Demand Points, Multi Basins System II

In this section, similar to Chapter 3, we provide results for India in national level (for all basins and demand points of the India that we could collect the data for). Same data sets described in Chapter 3 is used so that effect of using a pumped hydro storage can be explained easily. Table 4.11 shows the results for proposed sizes of upper and lower reservoirs and generators for each basin.

Table 4.11 | Reservoir and Generator Sizes for the Hybrid System with Pumped Hydro Storage

Rivers (1km3 ~ 240 GWh)	**Annual Inflow (km3)**	**Upper Reservoir Size (km3)**	**Lower Reservoir Size (km3)**	**Generator Size (GW)**	**Generation (%Demand)**
Bhagirathi	13.51	1.19	0.51	22.72	**15.9%**
Pinder	6.29	0.16	0.02	0.87	1.0%
Chenab	45.57	0.52	0.05	2.79	4.1%
Marusudar	12.02	0.16	0.02	0.93	1.1%
Lohit	19.17	0.07	0.02	0.93	1.4%

Dibang	9.17	0.02	0.01	0.27	0.5%
Barak	57.71	0.35	0.15	5.77	7.4%
Siang	456.47	0.49	0.30	15.56	29.2%

Table 4.12 | Reservoir and Generator Sizes for the Hybrid System with Conventional Hydro

Rivers **(1km3 ~ 240 GWh)**	**Annual Inflow (km3)**	**Reservoir Size (km3)**	**Generator Size (GW)**	**Generation (%Demand)**
Bhagirathi	13.51	0.28	1.22	**1.0%**
Pinder	6.29	0.16	0.71	0.5%
Chenab	45.57	0.49	4.27	3.8%
Marusudar	12.02	0.23	1.19	1.0%
Lohit	19.17	0.07	1.21	1.5%
Dibang	9.17	0.02	0.40	0.6%
Barak	57.71	0.08	4.08	4.6%
Siang	456.47	0.81	23.07	30.3%

In order to analyze the comparison easily, we present the counterpart result of Chapter 3 in Table 4.11-4.12. Most surprising change observed in Bhagirathi. In the pumped hydro system, Bhagirathi has highest capacity (22 GW) and provides the 15.9% of the energy compared to 1.22GW and 1.0% in the conventional system. The reason for this change is as follows: Since the solar energy generated in demand points can also be stored with pumped hydro system, recirculation of the inflow is possible and this enhances the generation at Bhagirathi quite a lot

Table 4.13 summarizes the pumped hydro case from the demand points' side and Table 4.14 shows the results for conventional system. There is a remarkable change in the diesel usage and pumped hydro system evidently helps reduce the intermittency of stream flow and increased the percentage of hydro contribution while fulfilling demand.

Table 4.13 | Solar Panel Area and Energy Generation Percentages by Type

Demand Points	Demand (GWh)	Solar Panel Area (km2/~GW)	Energy Generation by Type (% Demand)		
			Solar	Hydro	Diesel
Delhi	30013	81/~1.9	41%	55%	**4%**
Punjab	47534	122/~2.9	35%	62%	**3%**
Uttaranchal	11357	68/~1.6	35%	60%	**5%**
Himachal Pradesh	7744	176/~4.2	53%	43%	**4%**
Uttar Pradesh	87916	202/~4.8	31%	65%	**4%**
Bihar	13774	33/~0.8	31%	61%	**8%**
West Bengal	40777	112/~2.7	36%	59%	**5%**
Jharkhand	5663	16/~0.4	34%	59%	**8%**
Assam	5162	12/~0.3	25%	67%	**8%**
Chhattisgarh	17718	49/~1.2	37%	56%	**7%**

Table 4.14 | Solar Panel Area and Energy Generation Percentages by Type for the Conventional System

Demand Points	Demand (GWh)	Solar Panel Area (km2/~GW)	Energy Generation by Type (% Demand)		
			Solar	Hydro	Diesel
Delhi	30013	49.7/~1.2	34%	36%	**31%**
Punjab	47534	75.1/~1.8	30%	46%	**24%**
Uttaranchal	12786	17.2/~0.4	25%	59%	**16%**
Himachal Pradesh	7744	13.6/~0.3	33%	46%	**20%**
Uttar Pradesh	87916	132.4/~3.2	29%	40%	**31%**
Bihar	13774	20.2/~0.5	29%	42%	**29%**
West Bengal	40777	55.0/~1.3	27%	45%	**27%**
Jharkhand	5663	8.7/~0.2	31%	41%	**27%**
Assam	5162	7.2/~0.2	25%	49%	**26%**
Chhattisgarh	17718	24.8/~0.6	30%	51%	**19%**

The major transmission lines (≥1GW) between hydropower stations and demand points can be seen in Table 4.15 (Table 4.16 for the conventional system). In the pumped hydro system, solar energy is also transmitted via bi-directional transmission lines and we observe that especially Northern India, this increases the capacity of the lines. For example, Himalach

Pradesh which is one of the states with the lowest demand has a twelve times larger solar panel area compared to conventional system. This shows that solar energy potential in Himalach Pradesh is mostly sent to Bhagirathi to be pumped up and released again to meet demand in the other states.

Table 4.15| Transmission lines for the pumped hydro system

GW_peak	Delhi	Punjab	Uttarak-hand	Himalach Pradesh	Uttar Pradesh	Bihar	West Bengal	Jharak-hand	Assam	Chattis-garh
Bhagirathi	3.4	4.6	4.9	17.0	3.3	0.4	0.6	0.3	0.0	1.1
Pinder	0.2	0.0	0.1	0.2	0.3	0.2	0.1	0.0	0.0	0.2
Chenab	0.3	1.9	0.0	0.6	0.0	0.0	0.0	0.0	0.0	0.0
Marusudar	0.0	0.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Lohit	0.0	0.0	0.0	0.0	0.0	0.0	0.9	0.0	0.0	0.0
Dibang	0.0	0.0	0.0	0.0	0.3	0.0	0.0	0.0	0.0	0.0
Barak	0.0	0.0	0.0	0.0	0.8	0.1	3.7	0.2	0.5	1.1
Siang	1.0	2.0	1.0	0.0	7.8	1.3	0.8	0.5	0.5	0.6

Table 4.16 | Transmission lines for the conventional system

GW_peak	Delhi	Punjab	Uttarak-hand	Himalach Pradesh	Uttar Pradesh	Bihar	West Bengal	Jharakhan	Assam	Chattis-garh
Bhagirathi	0.2	1.0	0.7	0.0	0.0	0.0	0.0	0.0	0.0	0.1
Pinder	0.2	0.1	0.7	0.0	0.4	0.0	0.0	0.0	0.0	0.1
Chenab	1.6	1.8	0.8	0.4	0.0	0.0	0.0	0.0	0.0	0.0
Marusudar	0.0	1.2	0.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Lohit	0.0	0.0	0.0	0.0	0.0	0.0	1.2	0.0	0.0	0.0
Dibang	0.0	0.1	0.0	0.1	0.0	0.0	0.0	0.1	0.0	0.0
Barak	0.0	0.0	1.6	0.1	0.5	0.0	1.8	0.0	0.3	0.4
Siang	2.3	3.8	0.6	0.5	9.5	1.6	2.3	0.6	0.6	1.6

In Table 4.17 and 4.18, we present average power transmitted between supply and demand points which is calculated over one year horizon. We note that these values are underestimated since pumping and generation cannot happen at the same time and we average both quantities over 24 hours in a day. Supporting the discussion above for the transmission lines for the pumped hydro system (Table 4.15 in comparison with Table 4.16), Table 4.18 also shows a significant power flow from Himalach Pradesh and Bhagirathi River to be pumped up. This point can be seen with further clarity: In Table 4.19 we present total solar production within each demand points over planning horizon to see what percentage of the solar energy is being pumped, internally used or wasted. We see that 82% of the solar energy generated in Himalach Pradesh is being pumped in Bhagirathi power station to be distributed to other demand points. These results can be analyzed together with Table 4.13 where we show the energy supply breakdown to meet the total demand (The solar components in Table 4.13 represent the internally used solar energy amount). One should keep in mind that here we allow a network only between demand and supply points. The transmission network might look different (and more complicated) if power flow between demand points is also allowed.

Table 4.17 | Average power sent from supply point to demand points (hydropower)

Supply to Demand (GW_avg)	Delhi	Punjab	Uttarak-hand	Himalach Pradesh	Uttar-Pradesh	Bihar	West Bengal	Jharak-han	Assam	Chattis-garh
Bhagirathi	1.3	1.0	0.4	0.1	1.2	0.1	0.2	0.1	0.0	0.4
Pinder	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.1
Chenab	0.1	0.9	0.0	0.2	0.0	0.0	0.0	0.0	0.0	0.0
Marusudar	0.0	0.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Lohit	0.0	0.0	0.0	0.0	0.0	0.0	0.4	0.0	0.0	0.0
Dibang	0.0	0.0	0.0	0.0	0.2	0.0	0.0	0.0	0.0	0.0
Barak	0.0	0.0	0.0	0.0	0.2	0.0	1.6	0.0	0.1	0.3
Siang	0.4	1.1	0.4	0.0	4.8	0.8	0.5	0.3	0.3	0.3

Table 4.18 | Average power sent from demand point to supply points (solar power)

Demand to Supply (GW_avg)	Delhi	Punjab	Uttarak-hand	Himalach Pradesh	Uttar Pradesh	Bihar	West Bengal	Jharak-han	Assam	Chattis-garh
Bhagirathi	0.3	0.5	0.8	3.4	0.5	0.1	0.1	0.0	0.0	0.2
Pinder	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Chenab	0.0	0.2	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0
Marusudar	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Lohit	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.0
Dibang	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Barak	0.0	0.0	0.0	0.0	0.1	0.0	0.6	0.0	0.1	0.2
Siang	0.1	0.2	0.2	0.0	0.7	0.1	0.1	0.1	0.0	0.1

Table 4.19 | Solar energy production and utilization in demand points

Demand Points	Demand (GWh)	Solar Panel Area (km2)	Solar Production (GWh)	% Distribution of Solar Production		
				Pumped	Internal	Spilled
1km2 ~ 0.024 GW						
Delhi	30013	81	18250	24%	67%	9%
Punjab	47534	122	26584	34%	63%	3%
Uttarakhand	12786	68	14869	57%	30%	13%
HimalachPradesh	7744	176	38349	**82%**	11%	8%
UttarPradesh	87916	202	43449	29%	63%	8%
Bihar	13774	33	7266	26%	59%	16%
West Bengal	40777	112	23686	32%	61%	7%
Jharakhan	5663	16	3768	33%	51%	17%
Assam	5162	12	2505	34%	51%	15%
Chattisgarh	17718	49	11845	36%	55%	9%

The results for multi basins- multi demand points case, so far are given for a deterministic case with comparison to conventional hydro system for the same data. In order to be able to consider different scenarios, we identify four scenarios which are observed as years with either highest or lowest inflow for multiple basins. The optimal solution for these scenarios are calculated and presented in Table 4.20-22.

Table 4.20 | Reservoir and Generator Sizes for the Hybrid System with Pumped Hydro Storage for Multiple Scenarios

Basin Number	Rivers	Annual Inflow (km3)	Upper Reservoir Size (km3)	Lower Reservoir Size (km3)	Generator Size (GW)
(1km3~ 240 GWh)					
1	**Bhagirathi**	13.51	0.68	0.47	20.16
2	**Pinder**	6.29	0.36	0.03	1.15
3	**Chenab**	45.57	1.00	0.06	3.29
4	**Marusudar**	12.02	0.32	0.02	0.98

5	**Lohit**	19.17	0.11	0.02	1.11
6	**Dibang**	9.17	0.02	0.01	0.37
7	**Barak**	57.71	0.25	0.13	4.94
8	**Siang**	456.47	0.48	0.31	15.97

Table 4.21 | Solar Panel Area in Demand Points for the Hybrid System with Pumped Hydro Storage for Multiple Scenarios

Demand Points (1km2~ 0.024 GW)	**Demand (GWh)**	**Solar Panel Area (km2)**
Delhi	30013	75
Punjab	47534	111
Uttarakhand	12786	46
HimachalPradesh	7744	208
UttarPradesh	87916	196
Bihar	13774	32
West Bengal	40777	110
Jharakhan	5663	15
Assam	5162	12
Chattisgarh	17718	47

Table 4.22 | Transmission Lines between Demand Points and Basins for Hybrid System with Pumped Hydro Storage for Multiple Scenarios

Rivers/ States	Delhi	Punjab	Uttarak-hand	Himalach Pradesh	Uttar Pradesh	Bihar	West Bengal	Jharak-hand	Assam	Chattis-garh
Bhagirathi	2.8	5.1	3.5	19.8	3.1	0.2	0.4	0.2	0.0	0.7
Pinder	0.0	0.0	0.4	0.3	0.4	0.2	0.1	0.0	0.0	0.4
Chenab	0.7	1.8	0.0	0.7	0.0	0.0	0.0	0.0	0.0	0.0
Marusudar	0.0	0.9	0.0	0.1	0.0	0.1	0.0	0.0	0.0	0.0
Lohit	0.0	0.0	0.0	0.0	0.0	0.0	1.1	0.0	0.0	0.0
Dibang	0.0	0.0	0.0	0.0	0.3	0.0	0.0	0.0	0.0	0.0
Barak	0.0	0.0	0.0	0.0	0.3	0.1	3.4	0.2	0.4	1.1
Siang	1.2	1.6	1.3	0.1	8.0	1.3	0.9	0.5	0.6	0.6

The major transmission lines between basins and demand points are also presented in Figure 4.15. Blue lines shows the flow which are dominated by hydropower transmission and yellow lines represent solar power dominated transmissions. In our model, transmission lines are assumed to be bidirectional and designed as a point-to-point topology. Direct Euclidean distances are calculated between source and destination points. Normally, for transmission networks, we observe multi-point topology. For example, in Figure 4.15 instead of having a low capacity dedicated transmission line between Basin 7 (Barak) and Chhattisgarh, a higher capacity transmission line could be first expanded to West Bengal and West Bengal and Chhattisgarh could be connected with another lower capacity transmission line. However, optimization of different network topologies is not the scope of this paper. Moreover, we should note that transmission lines which pass through Himalaya Mountains (e.g. Basin 8 and Delhi) may not be possible because of the elevation and other geographic constraints.

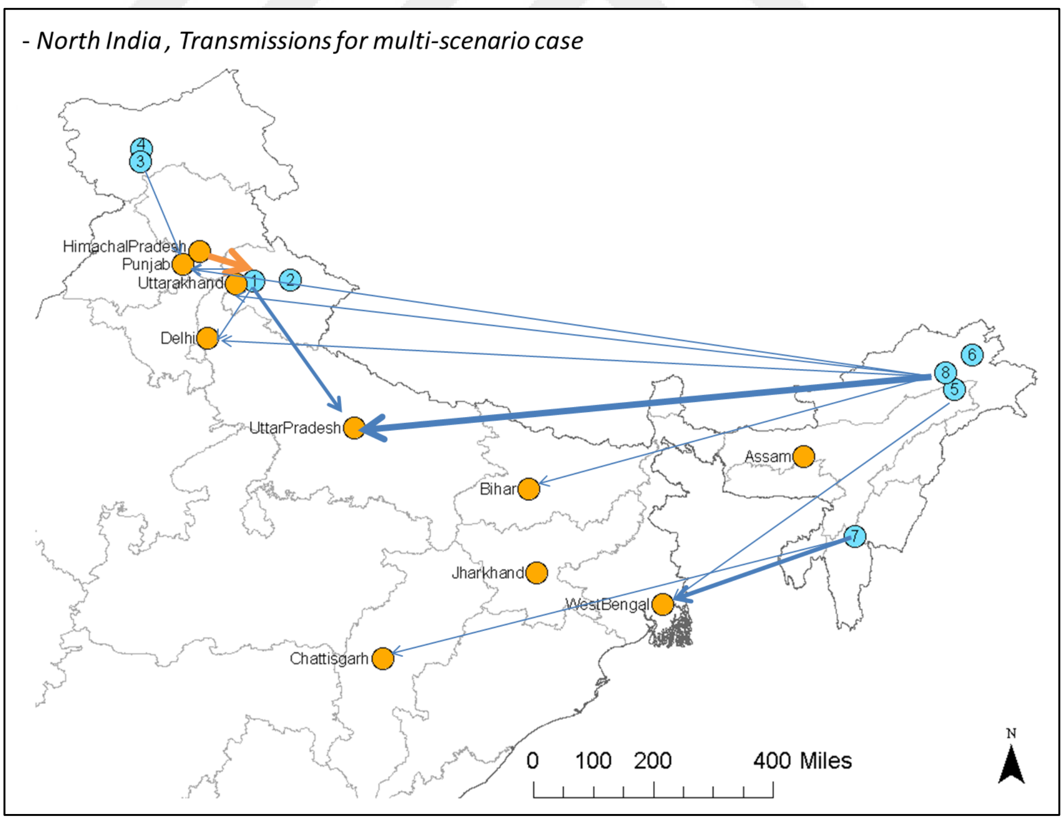


Figure 4.15 | Major transmission lines between basins and demand points. Blue lines represent hydro power dominated transmission lines. Yellow line represents the solar power dominated line between Himalach Pradesh and Bhagirathi River.

Finally, one should keep in mind that in this study we did not include lower and upper bounds for reservoir sizes or generator capacities. Other environmental and geographic constraints which are specific to basins are not also in the scope of this thesis. These results are only based on the stream flow potentials of the proposed basins.

4.6 Discussion

4.6.1 Sensitivity Analysis

Sensitivity analysis of the objective function can provide meaningful insight about the ways in which the optimal solution of the problem changes in response to small changes in the cost parameters. In Table 4.23, each investment variable, its reduced cost and the range over which its objective function coefficient can vary without forcing a change in the optimal basis is displayed. We see that we have very small ranges for each variable in which our current optimal solution remains the same.

Table 4.23| Sensitivity Analysis of Cost Parameters

	Unit Cost	Down	Up	Sensitivity Ranges
Upper Reservoir	1\$/m^3	-0.62%	9.73%	0.994-1.097
Lower Reservoir	1\$/m^3	-1.71%	1.29%	0.983-1.029
Pump/Generator	500$/kW	-0.05%	0.32%	499.75-501.62
Solar Panel	150\$/m^2	-0.46%	0.03%	149.30-150.04
Transmission Line[6]	254 $/kW	-0.16%	0.23%	252.59-253.59

Then, we perfomed another analysis to see the sensitivity of diesel usage and amount of pumped energy as the reservoir sizes change. The results of this analysis is presented in Figure 4.16. In Figure 4.16a we run our model fixing the lower reservoir size at different levels. Until

[6] Unit cost of transmission depends on the distance between points. This unit cost is calculated based on the distance between Delhi and Bhagirathi

the optimal size of lower reservoir, amount of pumped energy increases rapidly while diesel usage decreases. After optimal value we see that there is no gain from increasing the reservoir size.

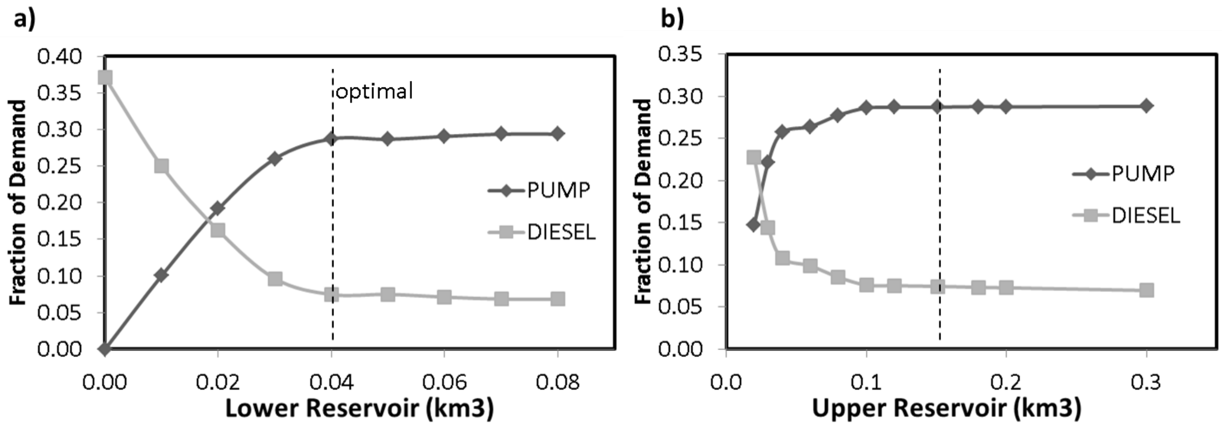


Figure 4.16 | Change in the pumped energy and diesel usage as reservoir sizes are fixed at different values. **a**, Lower Reservoir. **b**, Upper Reservoir.

In Figure 4.17, we show the marginal effect of increasing upper reservoir size. In Figure 4.16a, the change in the hydropower production with increasing reservoir size is presented. When upper reservoir size is zero, the hydro system works as run-of-the-river systems and more than 20% of the demand is met by hydro. Including a pumped storage system increases hydro's proportion to 47%. However, the marginal effect of increasing the upper reservoir size decreases. Here, it is possible to conclude that utilization of the reservoir decreases as the reservoir size increases. Another way to show the marginal effect of increasing reservoir size is presented in Figure 4.17b where we analyze the unit cost as the upper reservoir size increases. We see that when reservoir is smaller, the effect of increasing reservoir size has a bigger effect on the unit cost. Since the demand is constant, the unit cost represents the total cost and it naturally starts increasing after optimal size of upper reservoir. Figure 4.17c presents the results of dual analysis for upper reservoir size. For this analysis, an upper bound constraint is added to the model so that dual variables at different reservoir sizes can be calculated. Figure shows the marginal effect of increasing upper reservoir size on the objective function. The smaller the upper bound in the model is, the bigger is the improvement in the objective function when reservoir size is increased by one unit.

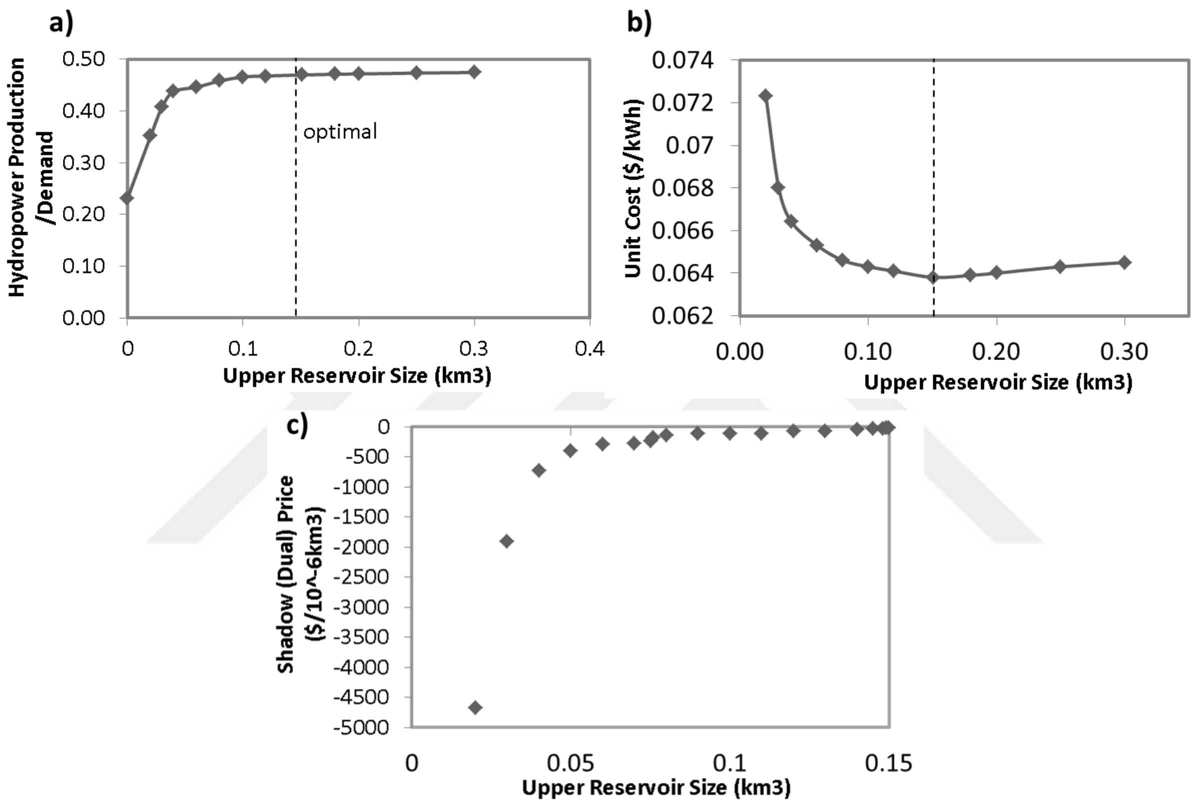


Figure 4.17 | Change in the hydropower production and unit cost of the system as upper reservoir sizes are fixed at different values. **a**, Hydropower production is normalized by demand. Figure shows that marginal value of increasing reservoir size decreases as the reservoir size is increased. **b**, Unit cost of the system is minimized when the upper reservoir is at optimal value. **c,** The dual price gives the improvement in the objective function if the constraint is relaxed by one unit. The smaller the upper bound in the model is, the bigger is the improvement in the objective function when reservoir size is increased by one unit.

4.6.2 Analysis on Time Resolution of the Model

The time resolution (increment) used in the model becomes quite important while working with variable and intermittent sources such as solar and wind. As seen in Figure 4.7, solar radiation has fluctuations on both short and long time scales. An approach that avoids capturing the variability of solar in short time scales could easily fail to accurately estimate the system sizes for new infrastructure. Moreover, as can be seen in Figure 3.1, we also observe diurnal variation in Himalayan stream flows occurring due to snow melts correlated with the sun light in the summer months. This also emphasizes the fact that systems including renewable sources may need to be modeled using short time increments as low as hours or minutes.

Furthermore, given that solar energy can be stored in the pumped hydro system in our model, and there is significant solar radiation variability in hourly level, using hourly time increments in addressing the problem becomes even more important. For our base case studies, our model is run with 3 hourly time increments. Clearly, the shorter the time increment is, the more accurate the results can be expected. However, increasing the resolution of the model will also increase the solution time significantly. Therefore, one needs to find the most suitable time increment that captures almost all the variability without making the analysis unnecessarily complicated. Here, we show that the critical time resolution due to the nature of our data is 3-hour increments.

To understand the effect of resolution for solar, output is first examined in terms of the frequency spectrum of solar radiation. As discussed in [32], spectral density of the output of solar systems provides insights for the diurnal, daily and seasonal cycles as well as the weather related, non-cyclic fluctuations. The method we followed to estimate the power spectrum has been described in [33]. Since the solar radiation data provided by NREL has hourly time intervals, to be able to increase the resolution even further, we used one-month real solar power output of a power station located in Gujarat (The output of the station in March 2012 is provided in Figure 4.18). The data is first rescaled to Delhi (for our case study) by latitude and linear interpolation method is used to expand the data for one year. Here, we applied Fourier transform to our input solar radiation data and plotted an estimate for the spectral density in logarithmic scale. The result for the data with 15-min resolution is provided in Figure 4.19.

The frequency spectrum shows that the highest peak is observed at a frequency of 1.157x10-5 Hz (24 hours) as expected because of the cyclic daily availability of solar. The resolution of the available data (15-min resolution) limits the observations in spectrum to 1.11x10-3 Hz. The most critical outcome obtained from the spectrum is that for time increments less than 3 hours, there is no significant peak observed in the data. Therefore, one can conclude that for our problem, using a time increment longer than 3 hours would fail to capture the variability of solar, whereas time increments less than 3 hours might not provide further improvements on the results.

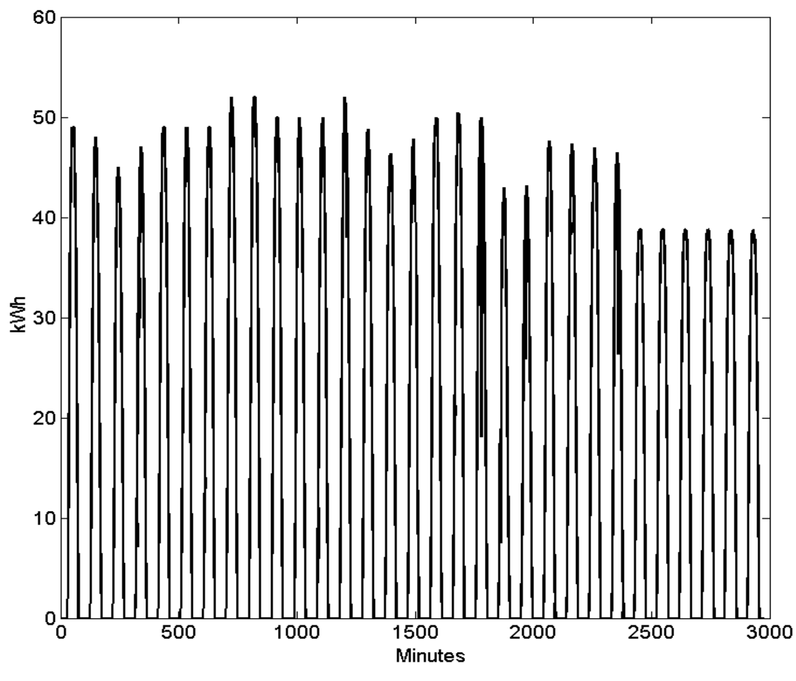


Figure 4.18 | Real energy output data from a solar power station in Gujarat over one month (March 2012) with 15-min sampling frequency.

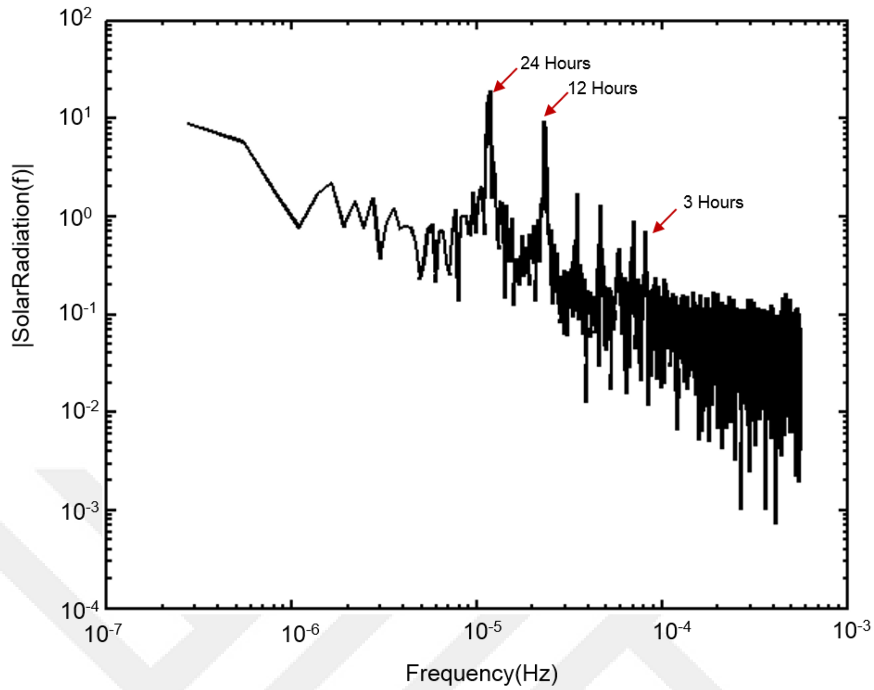


Figure 4.19 | An estimate for the spectrum of solar radiation data sampled at 15 min resolution. Highest peak is observed at 24hours and beyond 3hours, there is no significant peak is observed.

Next, we want to verify the conclusion above and analyze the effect of increased resolution. Decreasing the time increment of the model increases the number of constraints and reduces the solution space of the problem. Therefore, it is expected to observe higher cost for our minimization problem with smaller time increments. We present the results with 1-hour and 15-min time increments in comparison with 3-hour increments for Bhagirathi River and Delhi case study in Table 4.24. Here, we regenerated 1-hour and 3-hour resolution data using the 15-min resolution data to be able to have the sum of the radiation for one year equal for all data sets. The results for our analysis with different resolutions show that unit cost of the system indeed increases as time increment decreases; however, as the difference between the results of 3-hour resolution and 15-min resolution is not significant (2%), we can confidently assume that 3-hour resolution is sufficient to include the vast majority of the variations exist in the data.

Table 4.24| Sensitivity Analysis with Different Resolutions

	RESOLUTION					
	3-hour		1-hour		15-min	
Investments						
Upper Reservoir Size(km^3/~GWh)	0.1474 (35.4)		0.156 (37.4)		0.1565 (37.5)	
Lower Reservoir Size(km^3/~GWh)	0.0410 (9.84)		0.0405 (9.72)		0.0408 (9.79)	
Pump/Generator Size(GW)	1.478		1.543		1.560	
Solar Panel Size(km^2/~GW)	28.28 (0.68)		28.48 (0.68)		28.43 (0.68)	
Production	GWh	% Demand	GWh	% Demand	GWh	% Demand
Solar	3490	46.48	3276	43.63	3242	43.17
Hydro	3476	46.30	3681	49.02	3710	49.40
Diesel	542	7.22	552	7.35	557	7.42
Total	7508	100	7508	100	7508	100
Unit Cost						
Unit Cost of Hydro ($/kWh)	0.0189		0.0187		0.0187	
Unit Cost of Solar ($/kWh)	0.0536		0.0540		0.0540	
Unit Cost of Diesel ($/kWh)	0.1500		0.1500		0.1500	
Unit Cost of the System ($/kWh)	0.0632		0.0642		0.0644	

4.6.3 Analysis on Multiple Years Stream Flow Data

To be able to see the effect of different stream flows on infrastructure sizes in more detail, we have also completed an analysis with the complete time series of 53 years. We first run our model presented in Section 4.4.1 for every year individually. In Table 4.25, we observe that upper and lower reservoir sizes are more dependent on the inter-annual variability of the stream flow as expected and highest variability is observed in upper reservoir size as discussed in scenario analysis in Section 4.4.3.

Table 4.25 | Variability of Infrastructure Sizes for Different Years

	Mean	STDEV	CV
Upper Reservoir (km^3)	0.31	0.17	0.53

Lower Reservoir (km^3)	0.05	0.01	0.16
Generator/Pump (GW)	1.70	0.11	0.06
Solar Panel Area (km^2)	31.23	1.74	0.06
Transmission Line (GW)	1.70	0.11	0.06

Here, we use Bivariate Kernel Density Estimator to fit a joint probability distribution to lower and upper reservoir sizes and derive the bivariate density function shown in Figure 4.20. Joint density function appears to be bimodal and wide suggesting that the deterministic optimal solution calculated over the set of different years is not tight and wide range of options can give similar unit costs. First, we calculate the probability matrix given in Table 4.26 for different ranges of upper and lower reservoirs. We observe that according to deterministic results, for most of the years lower reservoir ranges between 0.040 km^3 and 0.055 km^3 and upper ranges between 0.1 km^3 and 0.4 km^3. For given ranges, we also present the average unit cost and percentage of demand met by the alternative fossil fuels (diesel) available for 53 years in Tables 4.27 and 4.28 respectively. As expected unit cost does not show significant variation for different ranges but diesel usage does. We observe that diesel distribution changes between 4% and 9%. The values around 4% is observed for larger upper reservoir size and the values around 9% is observed for smaller lower reservoir sizes.

Furthermore, we also estimate the expected unit cost of the system based on the probabilities of each year calculated using the density estimator. The expected unit cost of the system is calculated to be 0.069 $/kWh.

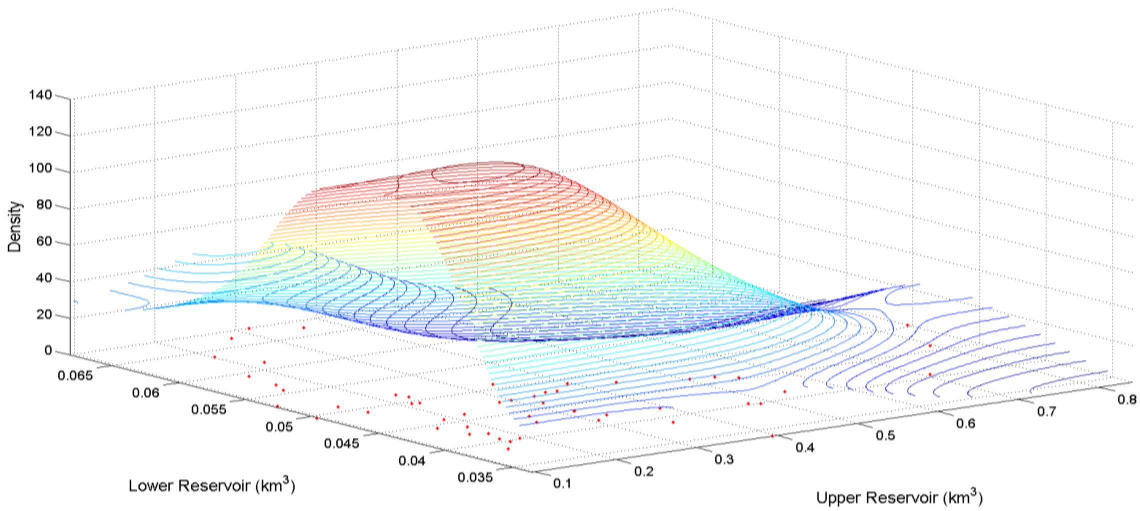


Figure 4.20 | Bivariate Kernel Density Estimate of Reservoir Sizes. Kernel density estimator based on [34] in MATLAB (kde2.m) is used with default parameters to estimate the densities. Red dots represent the observations for each year.

Table 4.26 | Probabilities for different ranges of upper and lower reservoir sizes

		Upper Reservoir (km3)							
		0.095-0.2	0.2-0.3	0.3-0.4	0.4-0.5	0.5-0.6	0.6-0.7	0.7-0.8	0.8-0.8406
Lower Reservoir Size (km3)	0.0335-0.035	0.004	0.004	0.004	0.003	0.002	0.001	0.000	0.000
	0.035-0.04	0.032	0.029	0.024	0.019	0.012	0.007	0.004	0.001
	0.04-0.045	0.065	0.061	0.050	0.037	0.024	0.014	0.010	0.004
	0.045-0.05	0.068	0.066	0.055	0.038	0.022	0.013	0.012	0.005
	0.05-0.055	0.050	0.040	0.027	0.015	0.007	0.004	0.006	0.003
	0.055-0.06	0.027	0.018	0.008	0.003	0.001	0.000	0.001	0.000
	0.06-0.065	0.019	0.016	0.011	0.005	0.001	0.000	0.000	0.000
	0.065-0.0681	0.011	0.014	0.013	0.007	0.002	0.000	0.000	0.000

Table 4.27 | Average unit cost for different ranges of upper and lower reservoir sizes

	Unit Cost ($/kWh)	Upper Reservoir (km3)							
		0.095-0.2	0.2-0.3	0.3-0.4	0.4-0.5	0.5-0.6	0.6-0.7	0.7-0.8	0.8-0.8406
Lower Reservoir Size (km3)	0.0335-0.035			0.065					
	0.035-0.04	0.065		0.069	0.068				
	0.04-0.045	0.069	0.069	0.068	0.066	0.067		0.068	
	0.045-0.05	0.070	0.071	0.070	0.075	0.071	0.074		
	0.05-0.055	0.069	0.070	0.070					0.078
	0.055-0.06	0.070							
	0.06-0.065		0.069						
	0.065-0.0681		0.073	0.072					

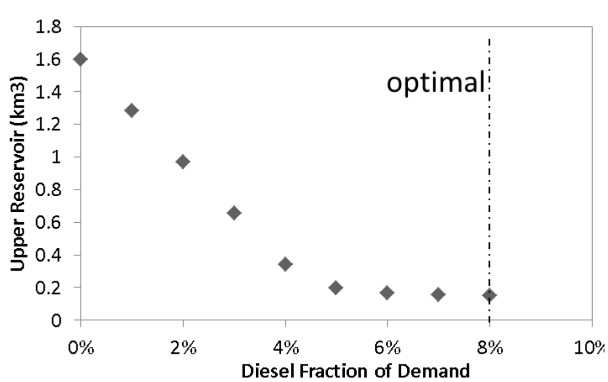
Table 4.28 | Percentage of demand met by diesel for different ranges of upper and lower reservoir sizes

	% Demand met by Diesel	Upper Reservoir (km3)							
		0.095-0.2	0.2-0.3	0.3-0.4	0.4-0.5	0.5-0.6	0.6-0.7	0.7-0.8	0.8-0.8406
Lower Reservoir Size (km3)	0.0335-0.035			9.05					
	0.035-0.04	7.76		9.20	7.22				
	0.04-0.045	8.36	7.55	6.09	5.70	5.21		4.63	
	0.045-0.05	8.49	6.70	6.44	6.06	4.90	4.33		
	0.05-0.055	7.17	6.24	5.59					4.35
	0.055-0.06	6.51							
	0.06-0.065		6.02						
	0.065-0.0681		7.63	6.35					

For the stream flow of a particular year (our base case, 1970), we also add a 'reliability' constraint to the model. In our model, the mismatched demand is penalized by an alternative fossil fuel cost (eg. diesel) and the goal is to meet the demand by renewable sources. Although system provides 100% reliabililty meeting the demand either by hyro,solar or diesel, we include a reliabiliy constraint into the model by limiting the percentage of demand met by diesel source. Figure 4.21 presents these results. At optimal solution, the diesel fraction to meet the demand is

8%. When diesel fraction is limited by a constraint, system increases the size of the upper reservoir. If the reliability constraint is above 4%, system increases the size of solar panels, upper and lower reservoirs; however between 0% and 4% system only increases the upper reservoir size. This can also be understood from the dual price as the dual price can be calculated from the change in upper reservoir when the reliability constraint is between 0 and 4%. As discussed in Chapter 3, the unit cost of reservoir capacity (i.e. constant incremental cost of additional capacity) that we used for our analysis is $1/m^3. If an additional 1 m^3 of a reservoir is used only once (in one time period only) during our planning horizon (1 year), we could generate 0.24 kWh of electricity (with 100 m head and 88% efficiency) with that 1m^3 water. In addition, annualized cost of an additional 1m^3 reservoir is $0.07 (with 25 year life time and 5% discount rate). Therefore, marginal cost of hydro can be roughly calculated as $0.29/kWh. We use unit cost for diesel as $0.15/kWh. For example, at 2% increasing the diesel usage by 1GWh will cause an increase in the objective function by $150000; however the gain from decreasing the upper reservoir size will be around $290000. Therefore, between 0% and 4% , the dual price is about $140000/ GWh and as the constraint approaches to optimal solution, the dual price goes to zero.

a)



b)

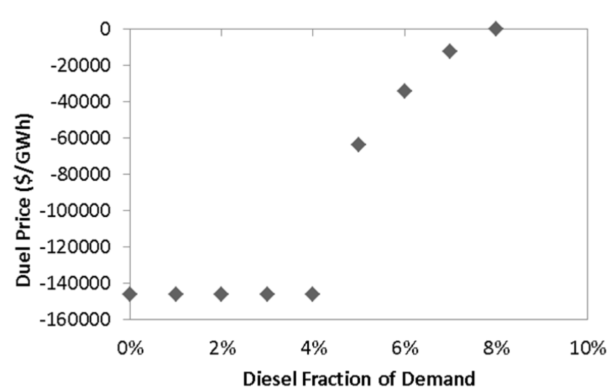


Figure 4.21 | Analysis on percentage of demand met by diesel source for a particular year. **a,** At optimal solution 8% of the demand is met by diesel. When diesel percentage is limited by a constraint, system increases the size of the upper reservoir. Up to 4%, system increases the size of solar panels, upper and lower reservoirs; however after 4% system only increases the upper reservoir size. **b**, Duel price of the constraint also presents that increasing diesel by 1 GWh reduces the objective function by \$140000 between 0% and 4%. The effect of relaxing the constraint decreases around optimal solution (8%).

4.7 Conclusion

We have introduced a model to determine optimal sizing of infrastructure needed to match demand and supply in a most reliable and cost effective way. We have combined for the first time three important concepts which help reduce the intermittency of renewable sources: hybrid systems, energy storage (pumped hydro) and long distance transmission in regional or national level. This model is obtained by enhancing the model in Chapter 3 by introducing the pump hydro ability with a secondary lower reservoir. We show that pumped hydro systems increases the cost efficiency and utilization of renewable sources increasing the flexibility to match supply and demand. We also show examples of how transmission lines provide more geographic aggregation which smoothes the variability of intermittent sources over large distances. This helps design more efficient hybrid systems reducing the storage costs and generate more dispatchable (controllable) power.

Bibliography

[1] Jacobson, M. Z., & Delucchi, M. A. (2011). Providing all global energy with wind, water, and solar power, Part I: Technologies, energy resources, quantities and areas of infrastructure, and materials. Energy Policy, 39(3), 1154-1169.

[2] Delucchi, M. A., & Jacobson, M. Z. (2011). Providing all global energy with wind, water, and solar power, Part II: Reliability, system and transmission costs, and policies. Energy Policy, 39(3), 1170-1190.

[3] Ashok, S. (2007). Optimised model for community-based hybrid energy system.*Renewable energy*, *32*(7), 1155-1164.

[4] Ekren, O., & Ekren, B. Y. (2010). Size optimization of a PV/wind hybrid energy conversion system with battery storage using simulated annealing. *Applied Energy*, *87*(2), 592-598.

[5] Ekren, O., Ekren, B. Y., & Ozerdem, B. (2009). Break-even analysis and size optimization of a PV/wind hybrid energy conversion system with battery storage–a case study. *Applied Energy*, *86*(7), 1043-1054.

[6] Powell, W. B., George, A., Simão, H., Scott, W., Lamont, A., & Stewart, J. (2012). SMART: A Stochastic Multiscale Model for the Analysis of Energy Resources, Technology, and Policy. *INFORMS Journal on Computing*, *24*(4), 665-682.

[7] Mallah, S., & Bansal, N. K. (2010). Allocation of energy resources for power generation in India: business as usual and energy efficiency. *Energy Policy*,*38*(2), 1059-1066.

[8] Cormio, C., Dicorato, M., Minoia, A., & Trovato, M. (2003). A regional energy planning methodology including renewable energy sources and environmental constraints. *Renewable and Sustainable Energy Reviews*, *7*(2), 99-130.

[9] NEMS (2003), The National Energy Modeling System: An Overview 2003, Technical report, Energy Information Administration, Office of Integrated Analysis and Forecasting, U.S. Department of Energy, Washington, DC 20585.

[10] Loulou, R., Goldstein, G. & Noble, K. (2004), Documentation for the MARKAL Family of Models, Technical report, Energy Technology Systems Analysis Programme

[11] Fishbone, L. G., & Abilock, H. (1981). Markal, a linear-programming model for energy systems analysis: Technical description of the bnl version. *International journal of Energy research*, *5*(4), 353-375.

[12] Stoyan, S. J., & Dessouky, M. M. (2011). A Stochastic Mixed-Integer Programming approach to the energy-technology management problem. *Computers & Industrial Engineering.*

[13] Lall, U., & Miller, C. W. (1988). An optimization model for screening multipurpose reservoir systems. Water Resources Research, 24(7), 953-968.

[14] Lall, U., & Mays, L. W. (1981). Model for planning water-energy systems. *Water Resources Research*, 17(4), 853-865.

[15] Sinha, A. K., Rao, B. V., & Lall, U. (1999). Yield model for screening multipurpose reservoir systems. Journal of water resources planning and management, 125(6), 325-332.

[16] Archibald, T. W., McKinnon, K. I. M., & Thomas, L. C. (1997). An aggregate stochastic dynamic programming model of multireservoir systems. *Water Resources Research*, 33(2), 333-340.

[17] Lamond, B. F., & Sobel, M. J. (1995). Exact and approximate solutions of affine reservoir models. *Operations Research*, 43(5), 771-780.

[18] Nandalal, K. D. W., & Bogardi, J. J. (2007). Dynamic programming based operation of reservoirs: applicability and limits. Cambridge university press.

[19] Pereira, M. V. F., & Pinto, L. M. (1991). Multi-stage stochastic optimization applied to energy planning. *Mathematical Programming*, 52(1-3), 359-375.

[20] Jacobs, J., Freeman, G., Grygier, J., Morton, D., Schultz, G., Staschus, K., & Stedinger, J. (1995). SOCRATES: A system for scheduling hydroelectric generation under uncertainty. *Annals of Operations Research*, 59(1), 99-133.

[21] De Ladurantaye, D., Gendreau, M., & Potvin, J. Y. (2009). Optimizing profits from hydroelectricity production. *Computers & Operations Research*, 36(2), 499-529.

[22] Johannesen, A., Gjelsvik, A., Fosso, O. B., & Flatabo, N. (1991). Optimal short term hydro scheduling including security constraints. Power Systems, IEEE Transactions on, 6(2), 576-583.

[23] Castronuovo, E. D., & Lopes, J. P. (2004). On the optimization of the daily operation of a wind-hydro power plant. *Power Systems, IEEE Transactions on*,19(3), 1599-1606.

[24] Brown, P. D., Peas Lopes, J. A., & Matos, M. A. (2008). Optimization of pumped storage capacity in an isolated power system with large renewable penetration. *Power Systems, IEEE Transactions on*, 23(2), 523-531.

[25] Kuznia, L., Zeng, B., Centeno, G., & Miao, Z. (2011). Stochastic optimization for power system configuration with renewable energy in remote areas. *Annals of Operations Research*, 1-22.

[26] Garcia-Gonzalez, J., de la Muela, R. R., Santos, L. M., & González, A. M. (2008). Stochastic joint optimization of wind generation and pumped-storage units in an electricity market. *Power Systems*, *IEEE Transactions on*, *23*(2), 460-468.

[27] Liang, X., Lettenmaier, D., Wood, E., & Burges, S. (1994). A simple hydrologically based model of land surface water and energy fluxes for general circulation models. *Journal of Geophysical Research*, 99(14), 415–14.

[28] Liang, X., Wood, E. F., & Lettenmaier, D. P. (1996). Surface soil moisture parameterization of the VIC-2L model: Evaluation and modification. *Global and Planetary Change*, *13*(1), 195-206.

[29] www.ilog.com

[30] Lamont, A. D. (2008). Assessing the long-term system value of intermittent electric generation technologies. *Energy Economics*, *30*(3), 1208-1231.

[31] Rueda, F., Moreno-Ostos, E., & Armengol, J. (2006). The residence time of river water in reservoirs. Ecological Modelling, 191(2), 260-274.

[32] Curtright, Aimee E., and Jay Apt. "The character of power output from utility-scale photovoltaic systems." *Progress in Photovoltaics: Research and Applications* 16.3 (2008): 241-247.

[33] Apt, Jay. "The spectrum of power from wind turbines." *Journal of Power Sources* 169.2 (2007): 369-374.

[34] Botev, Z. I., Grotowski, J. F., & Kroese, D. P. (2010). Kernel density estimation via diffusion. *The Annals of Statistics*, 38(5), 2916-2957.

Chapter 5

Main Conclusions and Future Work

5.1 Main Conclusions

Main results from this thesis through its methodologies and algorithms are tools for energy generation, transmission and distribution system design and help policy makers to make cost assessments in energy infrastructure planning rapidly and accurately. In the first part, we focus on power distribution systems planning for rural electrification using techniques from combinatorial optimization and developed new methods to estimate the cost of local-level distribution systems. In the second part, we focus on power generation and transmission using clean energy sources and demonstrate how we can get more reliable systems by combining various renewables and build a system where we can save one source of energy (solar power) in the form of other (hydro) using an infrastructure optimization (lower reservoir for a pumped hydro).

We note that, all of the results presented in this thesis are based on actual data and realistic assumptions. For optimizing distribution systems, we have analyzed some villages in Sub-Saharan Africa as there is either no network coverage or a partial coverage in these villages, a perfect setting for testing our algorithm. We have worked on Northern India for the advanced hybrid systems with clean energy sources, where there is rich solar energy in addition to high hydro power potential in Himalaya Mountains, to see if renewables could be utilized to meet growing electricity demand without fossil fuels in electricity generation.

5.1.1 Power Distribution Systems Design for Rural Electrification

The investment costs of off-grid approaches are easier to estimate, but the investment costs of networked approaches are more difficult to estimate, taking into account both the spatial distribution of demand and the optimal placement of infrastructure to meet that demand. We have used digitized QuickBird satellite imagery with remote sensing techniques to identify household-level demand points in the Sub-Saharan African villages. Figure 5.1 shows the household distribution and the existing grid for Mbola, Tanzania as an example.

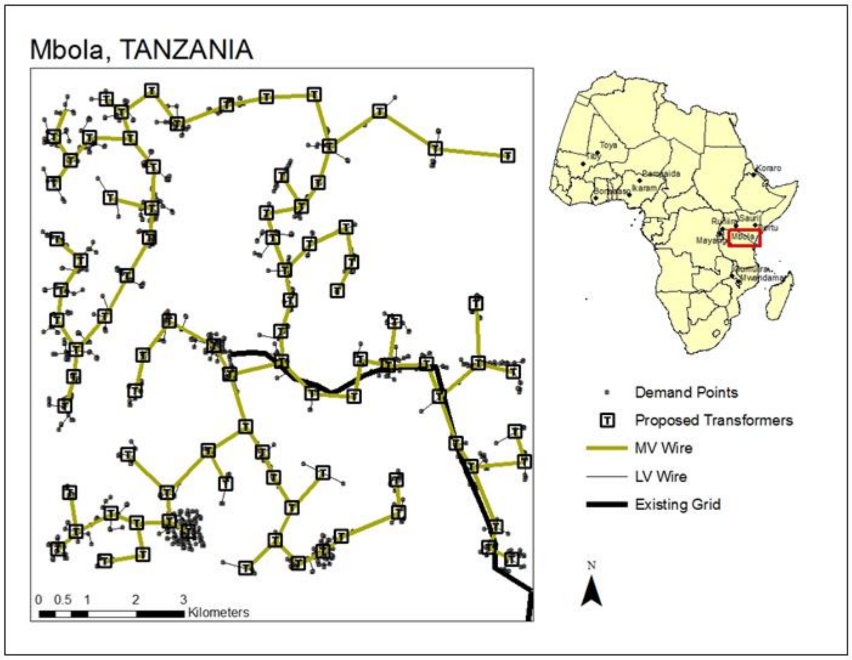


Figure 5.1 | Match between proposed network and existing grid. Proposed transformers, low voltage (LV) and medium voltage (MV) networks compared to partial existing grid. Algorithm outputs 90 transformers for 1175 demand points.

We present a heuristic algorithm for multi-level distribution system layouts while minimizing overall cost of infrastructure costs; specifically the combined costs of transformers and the two-tiered network together. To our knowledge, this algorithm is the first heuristic algorithm that selects the locations and service areas of transformers without requiring candidate solutions and simultaneously builds two-level grid network in a green-field setting. It allows one to specify different costs for the higher throughput lines upstream of the transformer as compared to downstream of the transformer. This point is particularly important as that is the case in the real world implementations and through our analysis we have showed that the ratio between the unit cost for low voltage (LV) lines and medium voltage(MV) lines could be quite critical in designing the overall network and give policy planners a good starting point for more detailed smart grid projects. Figure 5.2 summarizes the network sensitivity results.

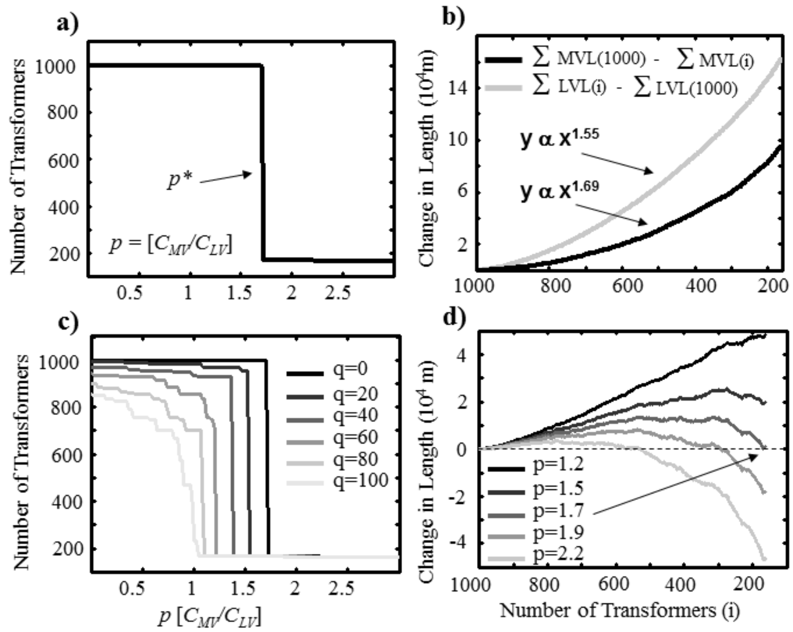


Figure 5.2 | Network Sensitivity Analysis. a, The sudden change in the number of transformers as the ratio between C_{MV} (unit cost for medium voltage line) and C_{LV} (unit cost for low voltage line) cost parameters increases. **b,** Change in total MV and LV lengths from the 1000 transformer case. **c,** The change in the number of transformers for different *q* ratios as *p* ratio increases.(See Section 5.1.2 for q). **d**, The difference between the curves in b with MV weighted with different *p* ratios.

5.1.2 Power Generation and Transmission Using Clean Energy Sources

In Chapter 3 and 4, we focus on power generation and transmission using clean energy sources. To reduce the role of fossil fuels and ease the concerns on the electricity generation, energy models which involve clean and renewable energy sources are necessitated. However, renewable sources are generally are intermittent and heavily dependent on the spatial location. One of the ways to mitigate the intermittency of renewable sources is to design hybrid systems which operate as a combination of alternative resources. We discuss the problem of modeling hydro and solar energy production and allocation, including long-term investments and storage, capturing the stochastic nature of hourly supply and demand data.

We imagine a country in the future where hydro and solar are the dominant sources and fossil fuels are only available in minimal form, in the shape of diesel generators as an example. In this country, we first identify candidate basins for hydro power stations and aggregated demand point locations (the cities or the states of the country). Within demand points solar energy production is possible. Then, we determine the possible transmission network between supply and demand points. Northern India with widespread solar power potential and high hydro power potential in Himalaya Mountains seems to be a very good candidate for the imagined country in the future. Figure 5.3 shows the identified basins and demand points.

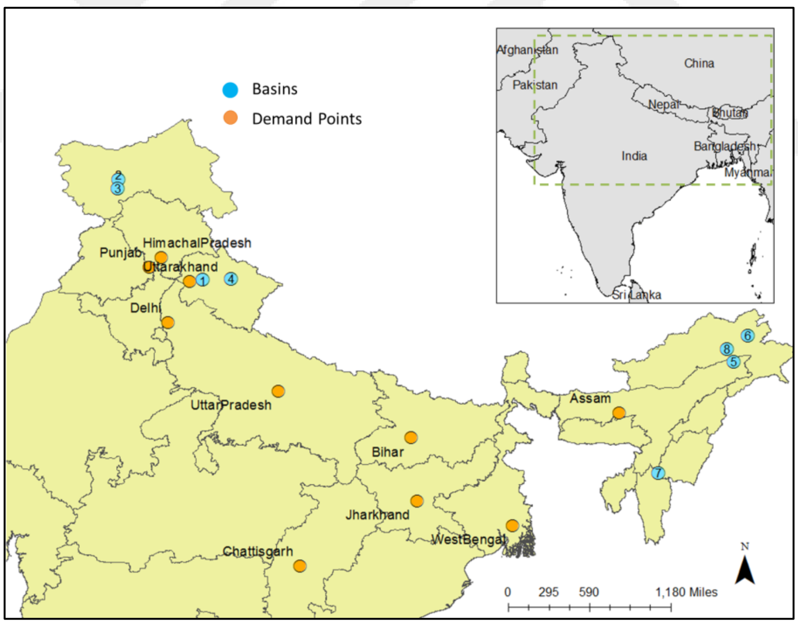
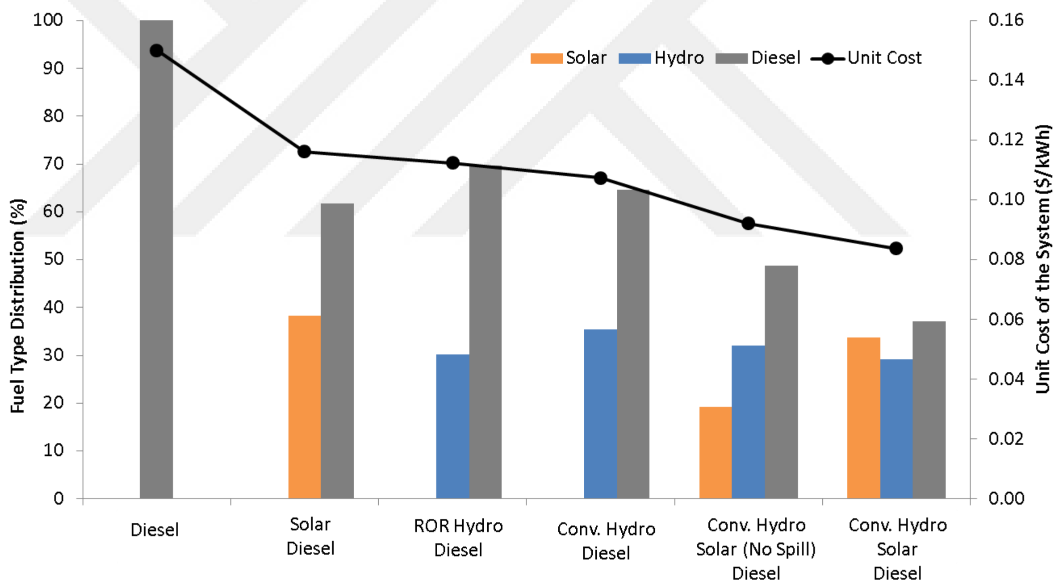


Figure 5.3 | Basins and demand points determined in India for analysis. Data is collected from CEA (Central Electricity Authority, Power Ministry of India) and other official websites to accurately estimate the 3-hourly demand load profile of each state for one year. If there is missing data for some days or hours within a day, interpolation/extrapolation methods are performed for projection.

An innovative contribution of the work in chapter 3 and chapter 4 is the establishment of a new perspective to energy modeling by including fine-grained sources of uncertainty such as stream flow, solar radiations and demand in hourly level as well as spatial location of supply and demand in national level.

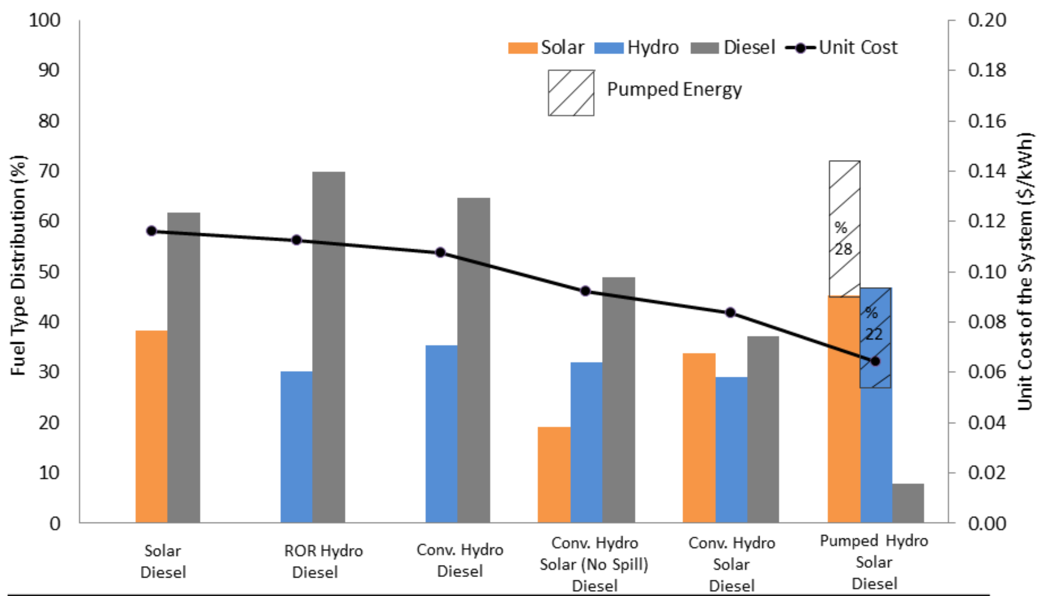
In Chapter 3, a scenario based linear programming approach is described for modeling the hybrid solar and hydropower system with conventional hydro storage. In this system, incoming stream flows are stored in large reservoirs in dams and water release is deferred until it is needed. Figure 5.4 clearly shows that as the use of the renewable increase and gets more advanced, unit cost of the system decreases and the most cost efficient alternative is the one that we have used in our model for Chapter 3.



	Solar Diesel	ROR Hydro Diesel	Conv. Hydro Diesel	Conv. Hydro Solar (No Spill) Diesel	Conv. Hydro Solar Diesel
Reservoir (km3)	-	-	0.24	0.24	0.29
Generator (GW)	-	1.09	1.13	1.05	1.06
SolarPanel (km2)	16.61	-	-	6.42	12.50

Figure 5.4 | Comparison of alternative technologies. Solely diesel is the most expensive. As the use and the variety of clean energy increase, the unit cost goes down substantially. It is more cost efficient to design solar panel area based on high demand and spill some of the solar energy in low demand periods.

Furthermore in chapter 4, to the same problem in chapter 3, we apply pumped hydro power stations in which water is pumped from lower reservoir to upper reservoir during periods of low demand to be released for generation when demand is high. We show that introduction of the pumped hydro storage increases the utilization of both hydro and solar sources and decreases the unit cost of the system significantly. The losses of the pumping process makes the pumped hydro stations a net consumer of energy overall (6% in Figure 5.5 is lost due to losses). However, as the solar energy can be also stored in the pumped systems, it provides more flexibility to the systems utilizing more clean energy and helps reducing the variability of sources and decreases the unit cost of the system (25% in the case study in Figure 5.5). Especially in the case study for which we provided this figure, addition of a small lower reservoir to the system (0.04 km3) causes to have a smaller "big" (upper) reservoir compared to conventional hydro system (0.15 km^3 compared to 0.29 km^3).



	Solar Diesel	ROR Hydro Diesel	Conv. Hydro Diesel	Conv. Hydro Solar (No Spill) Diesel	Conv. Hydro Solar Diesel	Pumped Hydro Solar Diesel
Upper Reservoir (km3)	-	-	0.24	0.24	0.29	0.15
Lower Reservoir (km3)	-	-	-	-	-	0.04
Generator (GW)	-	1.09	1.13	1.05	1.06	1.50
SolarPanel (km2)	16.61	-	-	6.42	12.50	28.18

Figure 5.5 | Alternative Technologies - Compared to other hybrid systems, the system with pumped hydro storage is the most cost efficient design with much lower unit cost. Having two-level reservoir with a pump system reduces the intermittency effect of renewable sources and the system can utilize more

clean energy. In particular, total hydro production increases with even smaller reservoir size compared to conventional system.

The role of solar energy in pumped system is twofold: internal solar and pumped solar. Therefore, it is expected that the solar panel area in the hybrid system with pumped storage is larger compared to a system with conventional storage. Here, what is interesting is that internally used solar energy in the demand points is also higher in the systems with pumped storage. Figure 5.6 compares conventional and pumped hydro systems in terms of the solar energy produced in one day. It can be seen that total production is scaled up by increased the solar panel area as the solar radiation curve is same. Solar energy used internally for the 4th and 5th time periods are the same as extra solar energy is spent for pumping, however shaded area in 3rd and 6th time periods in Figure 5.6 b represents the parts which explain the increase in the amount of solar energy internally used in demand points.

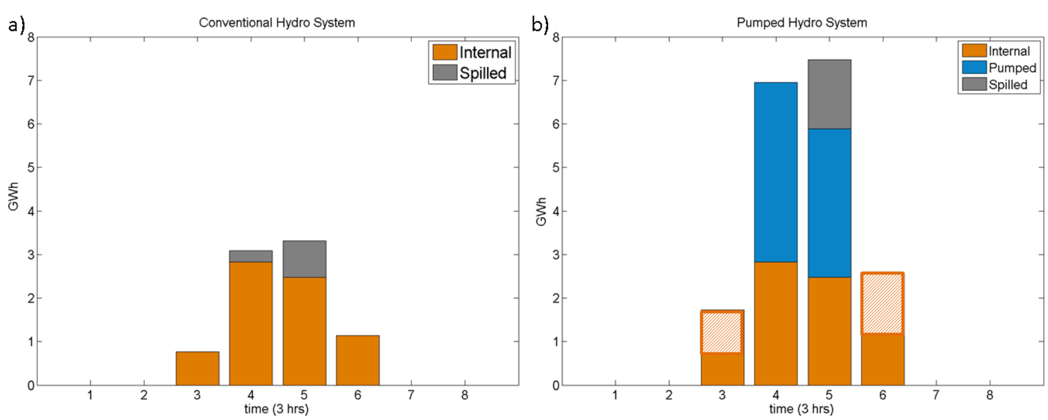


Figure 5.6 | Comparisons of solar production profile of one day in November for conventional. **a,** and pumped hydro systems. **b, -** Total production is scaled up by increased the solar panel area as the solar radiation curve is same. The role of solar energy in pumped system is twofold: internal solar and pumped solar. Solar energy used internally for the 4th and 5th time periods are the same between two systems as extra solar energy is spent for pumping, however shaded area in 3rd and 6th time periods in **b** represents extra solar internal.

Moreover, as individual systems are connected with each other through transmission lines and work as a single system, the variability and intermittency of renewable sources is expected to be smoothed out. To demonstrate effect of geographic aggregation, we sequentially add more basins and demand points to the system that we discussed in previous sections. We observe that unit cost of combined system is reduced when sources are used together. Total generator size used in the system with two basins is lower than the one with one basin. This indicates that combining two sources with different variability increases the flexibility of obtaining cheaper systems.

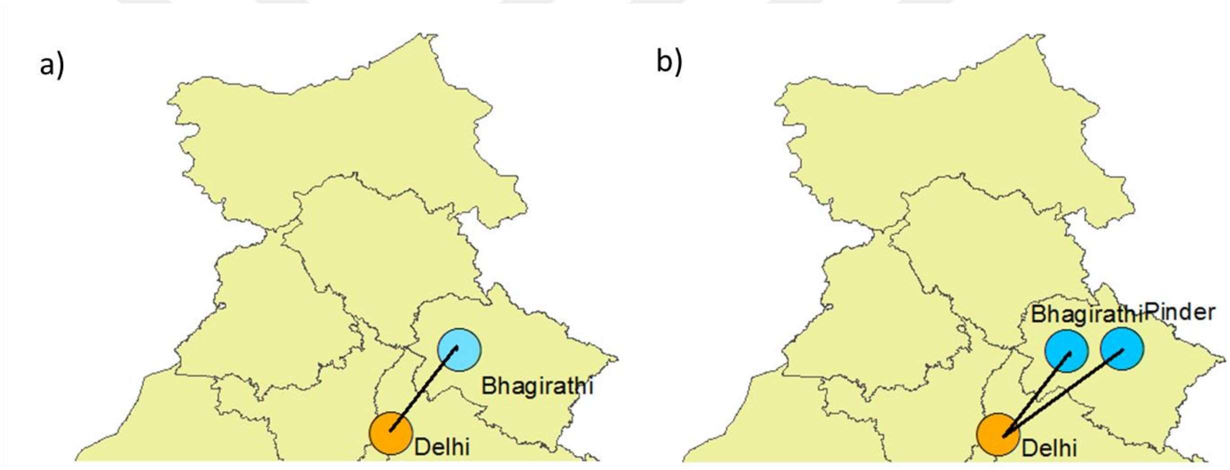


Figure 5.7 | Pinder basin is added to the system to show that geographic aggregation of alternative sources can reduce the variability of renewable sources.

Table 5.1 | Results for Bhagirathi and Pinder Rivers combined system to meet Delhi's demand

	Pinder + Bhagirathi		Bhagirathi	
Investment Size	Pinder	Bhagirathi		
Upper Reservoir Size(km^3)	0.032	0.093	0.148	
Lower Reservoir Size(km^3)	0.004	0.028	0.039	
Pump/Generator Size(GW)	0.216	1.139	1.499	
Transmission Line to Delhi (GW)	0.216	1.139	1.500	
Solar Panel Size(km^2)	24.98		28.18	
	(1 km2 ~24MW)			
Production	GWh	% Demand	GWh	% Demand
Solar	3291.95	43.87	3224.17	45.35
Hydro	3621.35	48.26	3695.38	46.87
Diesel	590.03	7.86	583.76	7.78
Total	7503.30	100.00	7503.30	100.00
Unit Cost				
Unit Cost of Hydro ($/kWh)	0.016		0.019	
Unit Cost of Solar ($/kWh)	0.054		0.054	
Unit Cost of Diesel ($/kWh)	0.150		0.150	
Unit Cost of the System ($/kWh)	0.059		0.064	

5.2 Future Work

There are many interesting optimization problems in the sustainable development field, especially the ones with a focus on rural electrification are particularly important. In chapter 2, we discuss about optimization of single and multi-level power distribution systems. Our discussion includes the initial decision of extending the network into an unserved area given the limited funding for a single level grid network. If it is only possible to connect some portion of the households to the electricity, how to choose that portion is a very critical and interesting question in terms multi-level networks as well and there is a need for analyzing this problem in the literature. To this end, we want to extent our algorithms on partial electrification problems to multi-levelnetworks.

Furthermore, renewable energy based decentralized systems are getting extremely favorable as they constitute an alternative to power generation in centralized power plants with fossil fuels due to the fact that they are expected to play a significant role in emission reduction and climate change mitigation. A decentralized power system can function either in the presence of a grid or as a stand-alone isolated system to meet the local demand. In literature, there a lot of studies about isolated cases and applications of decentralized systems [5]. However, we believe, more generalized studies are needed for assessing the feasibility of grid-connected and standalone decentralized energy systems.

We can also expand our contribution to the Smart Grid literature. Substantial research efforts continue to focus on Smart Grid technologies and many of the power system components in these emerging technologies such as energy storage devices and hybrid generation systems which include renewable resources are discussed in this thesis. However, there are still interesting concepts such as demand response programs which help integration of renewables into the grid and lead to more energy-efficient, environmentally-friendly, sustainable and more reliable electricity supply chains. Effective integration of these components poses important challenges and also good opportunities in planning of future power systems operations. Main point here is that demand response includes all electricity consumption pattern modifications by customers that are intended to change the timing and level of instantaneous demand and this is expected to make the grid more cost efficient. However, systems where consumers can directly participate in demand management require new efforts for forecasting the electric loads of individual consumers.

Our work on incorporating renewable sources into the grid can be expanded to analyze the frequency spectrum of intermittent sources in more detail. Frequency spectrum of the fluctuating data can help understand the coherence between the demand, solar production and the stream flow. For example, in our dataset we observe yearly, seasonal cycles (Figure 5.8a) and, for some sites, diurnal cycles (Figure 5.8b) in the stream flow data. Diurnal cycle in Bhagirathi River, for example, is observed for a short period of time especially in March and this cycle is supposed to be correlated with sun light and solar production (we also discussed about the spectrum of solar in Chapter 4). Frequency spectrum of 3-hourly demand data for one year in Figure 5.8c also shows the cycles observed in the data. Further analysis can be done with the individual power stations to understand how the combining alternative sources which have same or different cyclic behaviors can affect the use of back up source (diesel in our model). This analysis can give us idea about the alternative back up sources and their ramp rate characteristics to fill-in when there is not enough renewable power. Moreover, we can have more information about the demand response programs used to handle intermittency of renewables. Last but not least, understanding spectral distribution of resources and demand can also arise the possibility of additional decomposition methodologies across different time scale.

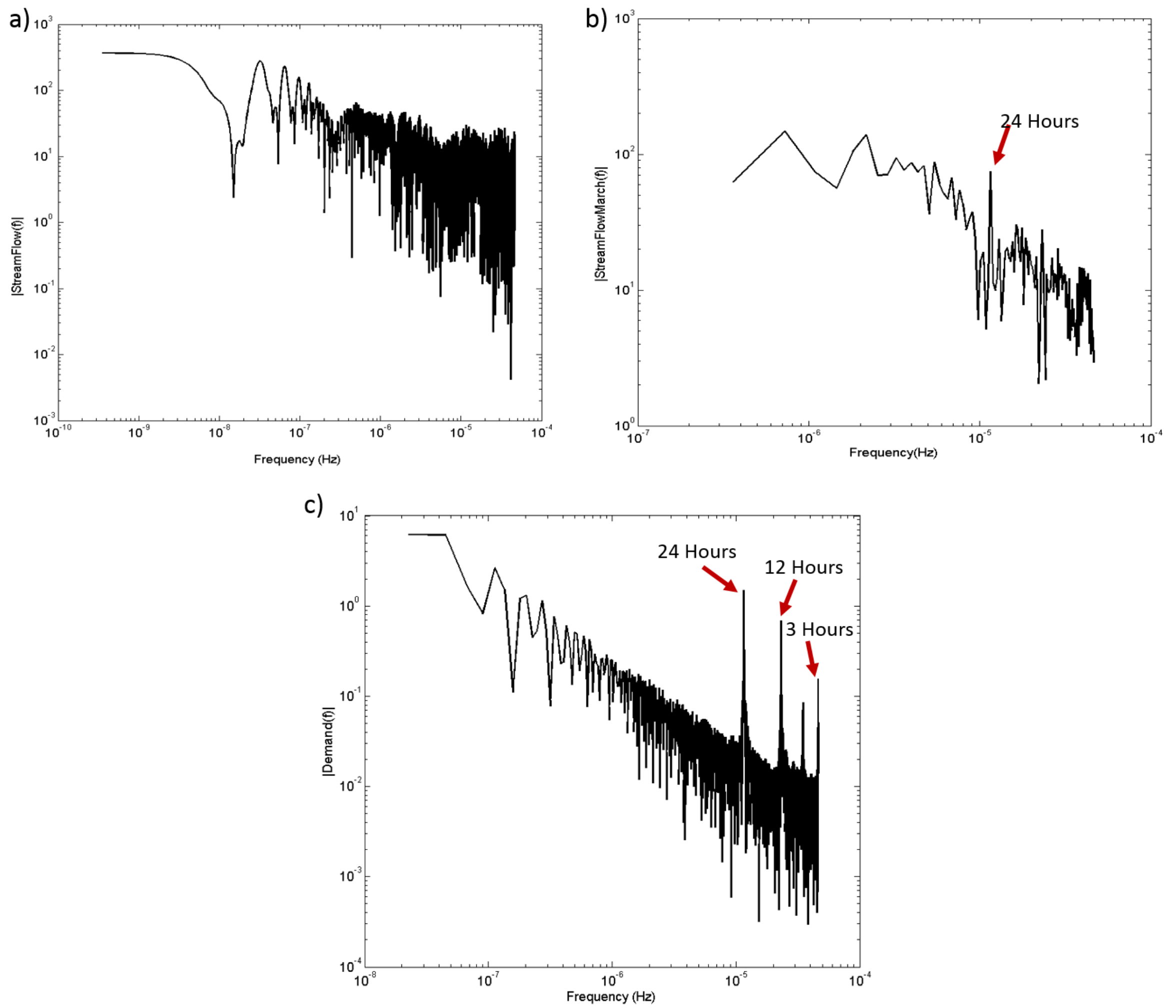


Figure 5.8 | Spectral distribution of stream flow and demand for Bhagirathi River and Delhi case study. **a,** The spectrum for 51 years of 3-hour resolution stream flow data shows the yearly fluctuations. **b,** The diurnal fluctuations observed in March is showed on the spectrum for 1 month data with 3-hour resolution. **c,** Spectral distribution of demand is presented for 1 year data of 3-hour resolution and therefore seasonal fluctuations can not be observed in the figure.

Another valuable extension to Chapter 4 would be to examine the variability of different intermittent sources and their effect on the pumped hydro system. The intermittent renewables have certain time scales that they vary. The energy spectrum of wind suggests that the diurnal variability of wind is not as strong as solar; however, as opposed to solar, wind also shows significant variation for the periods longer than one day (Figure 5.9). To answer the question of how the results would differ if we used another intermittent renewable source along with pumped hydro, we have also completed an preliminary analysis with wind data. A representative wind speed data with hourly time intervals is collected from NREL's website. We assume that we have 3MW turbines that we can place one in every $40D^2$ of area where D is the diameter of the wind turbine. To calculate the wind output, the characteristics given in a technical report for a 3 MW turbine is used [1]. The diameter is assumed to be 80 meters. To be able to use the current models for wind without making any significant changes, the wind input is also provided for unit area. Then, the model estimated the potential of wind generation by optimizing the area (which can also be converted to number of turbines needed using our $40D^2$ coverage assumption). In order to have a meaningful comparison between wind and solar, the cost of turbines is selected in a way that average cost of wind and solar per unit area would be same (provided that the average amount of the input data both for wind and solar could be generated throughout the year).

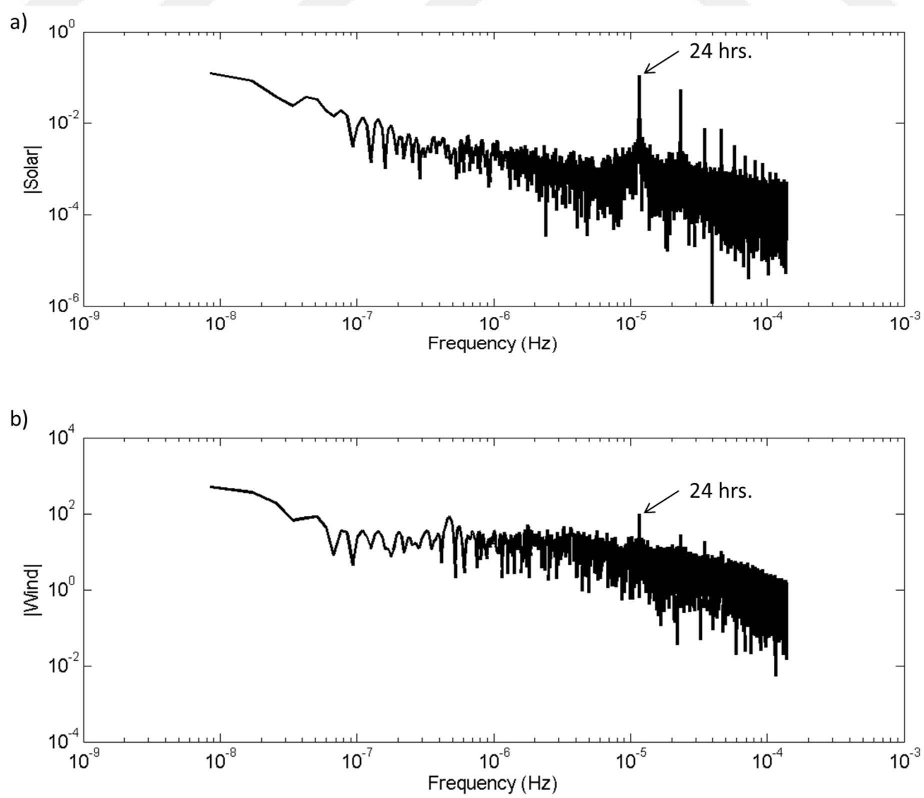


Figure 5.9 | Spectral distribution of intermittent sources. **a,** The spectrum of solar data used in base case studies (Delhi) with 1-hour resolution. **b,** The spectrum of wind data used as an example and collected from NREL website.

Figure 5.10 summarizes the initial results. Figure 5.10a shows that the wind energy generation is lower than the solar energy generation; however, the proportion of wind energy which is directly being used in demand points is higher than solar energy. This is possibly because the wind component with less diurnal variability and possibly being available during the night can be more suitable to be used directly. Moreover, the Figure 5.10 b shows the difference in hydropower generation with pumped water and the incoming stream flow when pumped hydro system is combined with an intermittent source (solar or wind). In both systems, the same amount of stream flow enters to the system but hydro power generation for the system combined with solar is higher than the one combined with wind. This work can be expanded to better understand the main driver which sizes the pumped hydro system and changes the proportions of sources to meet the demand.

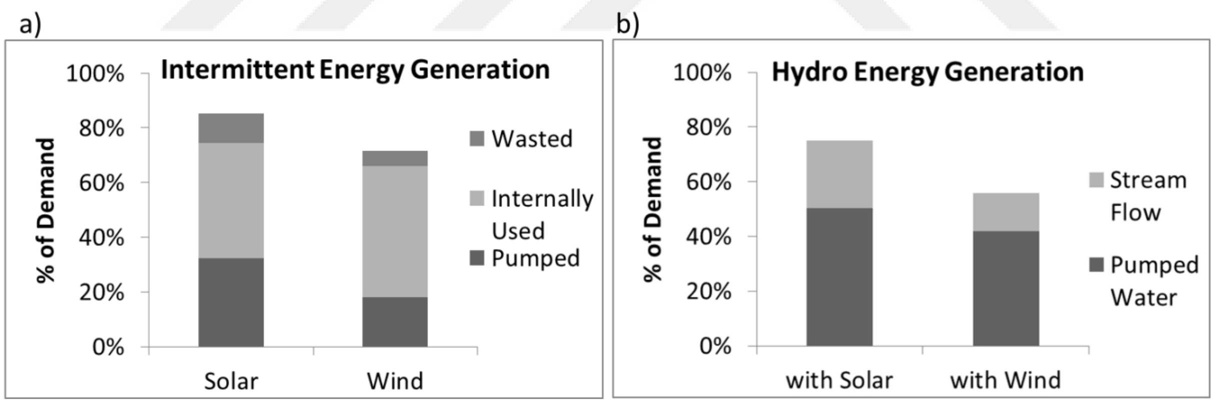


Figure 5.10 | Energy generation distribution comparisons for solar and wind. **a,** Proportion of wasted, internally used and pumped energy as a percentage of demand. **b,** Proportion of hydro energy generated by natural stream flow or pumped water.

Moreover, in chapter 3 and 4, we discuss about designing hydropower stations with large dams in order to be able to store energy in the reservoirs in the form of potential energy. Although, large hydro stations are sometimes considered environmentally unacceptable (building of major dams may involve displacement of land and populations and some geographical problems), commercial interest in conventional and pumped hydro systems has been increasing. The problems we discuss here consider the feasibility of resources in terms of their potentials and examine whether they are enough if fossil fuels can be eliminated from the electricity generation without looking into other environmental, social and political constraints related to them. In addition to the concerns related to building large dams such as including disrupting local ecology, and displacement of nearby people and animals, there might be other political constraints such as water sharing agreements between countries. Our work can be expanded to include more political and geographic aspects, social and environmental considerations.

[1] ENERCON Wind energy converters. Product overview. Version: July 2010.

APPENDIX A

Transmission Line Cost

Here, we will provide more details about the investment cost analysis for transmission lines used in Chapter 3 and Chapter 4. In our implementation, we consider the transmission cost is dependent on the distance and the capacity of the lines and we do not include the losses in our analysis.

We first calculate how the unit cost of unit power transmission varies based on the distance: We initially determine what type of transmission lines can be used to transmit a certain amount of power by a certain length of distance using surge impedance loading factor. Then, the investment costs of alternative transmission lines are calculated. For example, in order to transmit 300 MW by 250 miles, one can choose to use 500 kV DC Bipole, 500kV DC Bipole, 230 kV AC Double, 345 kV AC Single Circuit or 500 kV AC Single Circuit transmission lines where investment costs are estimated to be \$442 million, \$275.8 million, \$308 million and \$537.5 million respectively. We then repeat this procedure and calculate the cost for 1MW power transmission by 1 km for multiple scenarios including various distances and for various power levels. Table A1, A2, and A3 presents these results.

Table A4 : Transmission line type and associated investment costs for 100 miles

100 Miles		
	Circuit	**Transmission Line Cost ($M)**
1MW	11 kV AC Single	0.708
	33 KV AC Single	2.71
2MW	66 kV AC Single	6.14
	115 KV AC Single	30.23
100MW	161 kV AC Single	55.3
	230 KV AC Single	65.3
200MW	161 kV AC Double	78.6
	230 KV AC Single	79.2
300MW	345 KV AC Single	163
	230 KV AC Double	125.8
500MW	500 KV DC Bipole	250
	230 KV AC Double	243
	345 KV AC Single	185
	500 KV AC Single	295
1000MW	500 KV DC Bipole	300
	600 KV DC Bipole	335
	345 KV AC Double	280
	500 KV AC Single	380
1500MW	500 KV DC Bipole	370
	600 KV DC Bipole	412.5
	345 KV AC Double	335
	500 KV AC Single	475
2000MW	500 KV DC Bipole	440
	600 KV DC Bipole	490
	765 KV AC Single	681

Table A5: Transmission line type and associated investment costs for 500 miles

500 miles		
	Circuit	**Transmission Line Cost ($M)**
1MW	66 kV AC Single	20.0
	33 KV AC Single	3.5
10MW	66 kV AC Single	20.1
	138 KV AC Single	115.2
	115 KV AC Single	85.1
100MW	161 KV AC Single	199.3
	234 KV AC Single	312.3
	500 KV DC Bipole	803.3
200MW	500 KV DC Bipole	828.0
	345 KV AC Single	347.0
	500 KV AC Single	711.0
300MW	500 KV DC Bipole	842.0
	345 KV AC Double	358.0
500MW	500 KV DC Bipole	870.0
	345 KV AC Double	555.0
	500 KV AC Single	765.0
1000MW	500 KV DC Bipole	940.0
	500 KV AC Double	985.0
	765 KV AC Single	1030.0
1500MW	500 KV DC Bipole	1010.0
	600 KV DC Bipole	1132.5
	765 KV AC Double	1140.0
2000MW	500 KV DC Bipole	1080.0
	600 KV DC Bipole	1210.0
	765 KV AC Double	1245.0
	800 KV DC Bipole	1315.0

Table A6: Transmission line type and associated investment costs for 750 miles

750 miles		
	Circuit	**Transmission Line Cost ($M)**
1MW	66 kV AC Single	24.764
	115 KV AC Single	127.523
10MW	66 kV AC Single	24.89
	138 KV AC Single	172.73
	115 KV AC Single	37.83
100MW	345 KV AC Single	498.5
	230 KV AC Double	286.1
200MW	500 KV DC Bipole	1228
	345 KV AC Single	509.5
	500 KV AC Single	1048.5
300MW	500 KV DC Bipole	1242
	345 KV AC Double	783
	500 KV AC Single	1066.5
500MW	500 KV DC Bipole	1270
	500 KV AC Single	1102.5
	765 KV AC Single	1342.5
1000MW	500 KV DC Bipole	1340
	600 KV DC Bipole	1505
	765 KV AC Double	1447.5
1500MW	500 KV DC Bipole	1410
	600 KV DC Bipole	1582.5
	765 KV AC Double	2790
2000MW	500 KV DC Bipole	1480
	600 KV DC Bipole	1660
	800 KV DC Bipole	1802.5

After we have the unit cost values for all the scenarios, we then fit a linear curve to estimate the incremental cost per unit power per distance. The fitted curves are shown in Figure A1.

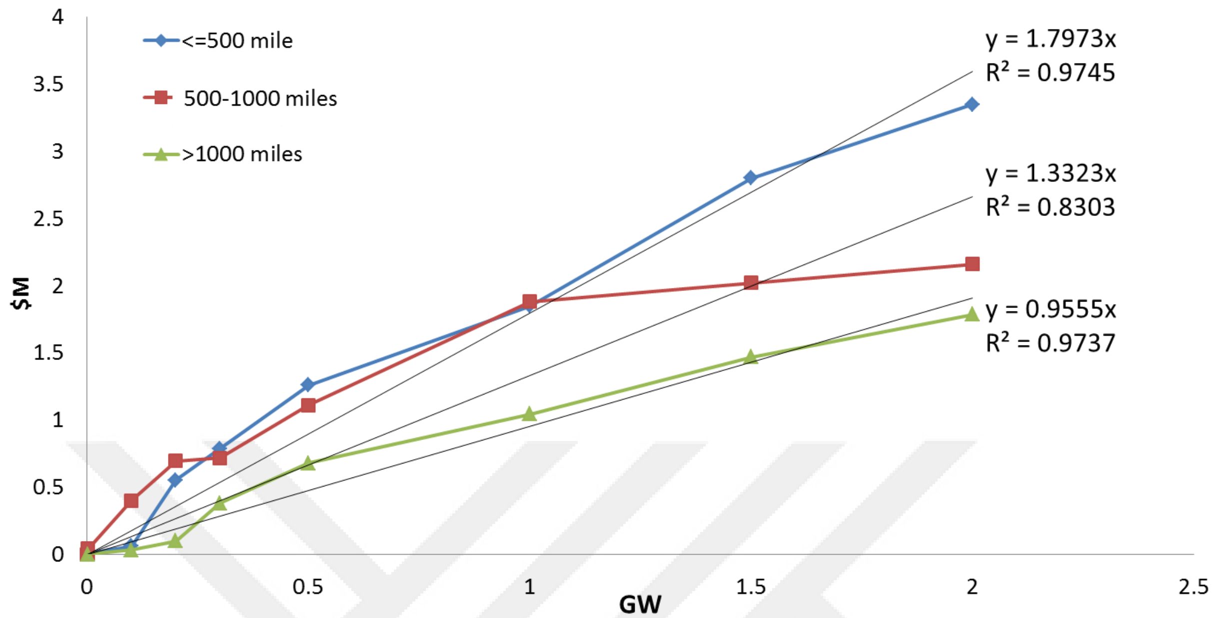


Figure A1| Curves fitted linearly to estimate the incremental cost per unit power per distance. We have calculated the cost for 1MW power transmission by 1 km for multiple scenarios including various distances and for various power levels (Table A1, A2 and A3).

In our analysis, we have determined a linear relationship between the amount of power and the cost. The slope of this relationship decreases with the distance transmitted. As a result, we have computed three cost parameters for three distance ranges (0-500 km, 500-1000 km and >1000 km) as $1.1 million, $0.8 million and $0.6 million per megawatt per kilometer, respectively. These findings are summarized in Table A4.

Table A7: Final transmission cost multipliers

Distance (km)	Cost multiplier per GW per mile ($M)	Cost multiplier per GW per km ($M)
<=500	1.79	1.1
500-1000	1.33	0.8
>=1000	0.96	0.6

Next, in our model, we calculate the spherical distance between two points and multiply it with the corresponding cost multiplier from above to estimate the unit investment cost of the transmission line between the two points per unit power [1]. Then, we use this unit cost in the objective function of our optimization model to minimize the total investment cost. The data used in the model is compiled by Selcuk Korel, MS'2010, Columbia University, using the sources [2-4].

Bibliography

[1] Sanoh, A., Kocaman, A. S., Kocal, S., Sherpa, S., & Modi, V. (2014). The economics of clean energy resource development and grid interconnection in Africa. *Renewable Energy*, *62*, 598-609.

[2] Gutman, R., Marchenko, P. P., & Dunlop, R. D. (1979). Analytical development of loadability characteristics for EHV and UHV transmission lines. Power Apparatus and Systems, IEEE Transactions on, (2), 606-617.

[3] Hao, J., & Xu, W. (2008, October). Extended transmission line loadability curve by including voltage stability constrains. In Electric Power Conference, 2008. EPEC 2008. IEEE Canada (pp. 1-5). IEEE.

[4] Bergen, A. R. (2000). Vijay Vittal Power System Analysis.

APPENDIX B

Demand Data

Here, we will show the demand data that we have used for the states discussed in Chapter 3 and Chapter 4. As it is mentioned in Chapter 3.3.2.2, total monthly power availabilities and requirements for each state for the year 2012 are provided on Central Electricity Authority, Power Ministry of India (CEA) website.

On the website of Load Dispatch Center of Delhi, daily load profiles for the days where peak demand occurred in each month are provided.

Table B1 shows the results for Delhi.

Delhi (MWh)	12am-3am	3am-6am	6am-9am	9am-12pm	12pm-3pm	3pm-6pm	6pm-9pm	9pm-12am
Jan	4,944	5,322	10,061	11,581	9,824	9,626	10,538	7,998
Feb	4,327	5,287	9,323	9,336	7,639	7,859	8,904	6,718
Mar	6,333	6,375	7,887	9,220	8,972	9,161	9,581	7,858
Apr	10,155	8,976	8,586	10,718	11,619	11,229	10,852	11,485
May	12,177	11,078	10,419	12,678	13,635	13,688	12,598	12,268
Jun	12,484	11,231	10,964	13,443	13,987	14,313	13,319	13,986
Jul	11,596	10,506	10,218	12,712	13,484	13,537	12,604	12,997
Aug	12,200	11,116	10,589	12,913	14,159	14,416	13,865	14,121
Sep	12,415	11,480	10,184	12,113	13,221	13,487	13,239	13,969
Oct	8,931	8,096	8,564	10,602	10,867	10,484	11,499	10,462
Nov	4,966	5,500	7,763	8,585	8,110	8,747	8,904	6,573
Dec	4,492	5,104	9,340	10,606	8,970	8,237	9,758	7,838

Chhattisgarh, Assam, Punjab, Himalach Pradesh and Uttarakhand also publish on their website daily, weekly or monthly reports for daily load profiles for a number of days or for the times when the minimum and maximum demands are observed. We have used the collected data to accurately estimate the 3-hourly demand load profile of each state for one year. If there is missing data for some days or hours within a day, interpolation/extrapolation methods are performed to project the data. When there is only limited number of days that we can use as a representative of all days in a month, we generated data from a normal distribution with mean equal to observed demand of given days and standard deviation equal to 5 percent of the observed data. The results can be seen in Table B2 – B6.

Table B2: Monthly average demand for Chhattisgarhi

Chattisgarhi(MWh)	12am-3am	3am-6am	6am-9am	9am-12pm	12pm-3pm	3pm-6pm	6pm-9pm	9pm-12am
Jan	5,496	6,161	7,500	6,161	5,600	6,161	7,612	6,161
Feb	6,648	6,942	7,554	7,315	6,588	7,302	7,710	6,994
Mar	5,496	6,161	7,500	6,161	5,600	6,161	7,612	6,161
Apr	5,900	6,973	7,500	6,973	5,738	6,973	7,753	6,973
May	6,400	6,814	7,500	6,814	6,292	6,814	7,733	6,814
Jun	5,700	6,202	6,300	6,202	5,440	6,202	6,810	6,202
Jul	6,000	6,769	7,400	6,769	5,874	6,769	7,602	6,769
Aug	6,000	6,480	7,100	6,480	5,795	6,480	7,303	6,480
Sep	5,200	5,968	6,400	5,968	5,049	5,968	6,642	5,968
Oct	6,700	7,351	7,800	7,351	6,612	7,351	8,083	7,351
Nov	5,600	6,142	7,400	6,142	5,524	6,142	7,586	6,142
Dec	5,850	6,336	7,500	6,336	5,729	6,336	7,640	6,336

Table B3: Monthly average demand for Punjab

Punjab (MWh)	12am-3am	3am-6am	6am-9am	9am-12pm	12pm-3pm	3pm-6pm	6pm-9pm	9pm-12am
Jan	8,954	11,403	13,852	11,069	11,069	11,069	12,407	11,403
Feb	10,500	10,404	16,445	13,099	14,398	14,398	15,050	13,099
Mar	12,583	10,905	14,901	12,583	12,583	12,583	14,504	12,583
Apr	12,732	10,139	14,914	12,732	10,344	12,732	14,709	12,732
May	15,089	14,349	18,216	17,073	17,073	13,580	17,806	17,073
Jun	22,210	18,842	25,617	22,210	22,210	21,469	26,467	22,210
Jul	27,077	23,752	26,870	26,870	26,870	21,662	26,870	28,320
Aug	25,914	21,382	23,916	23,916	23,916	21,768	23,916	26,400
Sep	18,715	17,971	20,087	20,087	21,571	20,087	21,252	18,397
Oct	17,826	17,274	19,027	19,027	20,228	19,280	21,376	19,027
Nov	10,807	12,921	12,921	14,527	12,921	12,921	12,921	11,396
Dec	8,017	10,198	12,282	10,198	10,198	8,422	11,682	10,198

Table B4: Monthly average demand for Assam

Assam (MWh)	12am-3am	3am-6am	6am-9am	9am-12pm	12pm-3pm	3pm-6pm	6pm-9pm	9pm-12am
Jan	1,560	1,919	1,920	1,919	1,471	1,919	2,726	1,919
Feb	1,660	2,054	1,975	2,054	1,677	2,054	2,901	2,054
Mar	1,546	1,907	1,732	1,907	1,581	1,907	2,770	1,907
Apr	1,697	2,017	1,821	2,017	1,764	2,017	2,785	2,017
May	1,857	2,113	1,827	2,113	1,893	2,113	2,875	2,113
Jun	2,193	2,400	2,123	2,400	2,215	2,400	3,069	2,400
Jul	2,235	2,476	2,240	2,476	2,212	2,476	3,217	2,476
Aug	2,543	2,609	2,294	2,609	2,284	2,609	3,315	2,609
Sep	2,204	2,513	2,114	2,513	2,320	2,513	3,412	2,513
Oct	2,059	2,407	1,941	2,407	2,116	2,407	3,512	2,407
Nov	1,755	2,117	2,055	2,117	1,709	2,117	2,947	2,117
Dec	1,807	2,226	2,134	2,226	1,832	2,226	3,130	2,226

Table B5: Monthly average demand for Uttarakhand

Uttarakhand (MWh)	12am-3am	3am-6am	6am-9am	9am-12pm	12pm-3pm	3pm-6pm	6pm-9pm	9pm-12am
Jan	3,230	3,445	4,153	3,961	3,639	3,862	4,522	3,736
Feb	3,135	3,496	4,475	4,168	3,606	4,130	4,561	3,821
Mar	3,099	1,962	3,417	3,743	3,564	3,739	4,746	3,827
Apr	3,909	3,810	4,039	3,777	2,593	3,299	4,040	3,967
May	3,721	3,834	3,948	4,057	4,203	4,213	4,555	4,470
Jun	3,930	4,617	4,749	4,519	3,951	4,084	4,143	4,173
Jul	4,305	4,270	4,074	3,923	3,875	3,719	4,045	4,144
Aug	3,746	3,807	3,973	3,522	3,768	3,815	4,279	3,961
Sep	3,774	3,829	4,258	3,804	3,779	4,035	4,357	3,830
Oct	3,113	3,675	3,687	4,124	4,016	4,084	4,644	3,721
Nov	3,302	3,758	4,044	3,638	3,115	3,746	4,584	3,514
Dec	3,173	3,516	4,486	4,165	3,767	4,067	4,447	3,509

Table B6: Monthly average demand for Himalach Pradesh

HimalachPradesh(MWh)	12am-3am	3am-6am	6am-9am	9am-12pm	12pm-3pm	3pm-6pm	6pm-9pm	9pm-12am
Jan	2,309	2,366	3,261	3,388	2,896	2,991	3,253	2,731
Feb	2,492	2,533	3,615	3,510	3,095	3,189	3,149	2,810
Mar	2,240	2,305	3,312	3,248	2,865	2,823	2,839	2,563
Apr	2,453	2,489	3,483	3,498	3,259	3,251	3,278	2,922
May	2,775	2,824	3,382	3,426	3,311	3,264	3,042	2,945
Jun	2,703	2,756	3,324	3,447	3,362	3,406	3,111	3,123
Jul	2,782	2,777	3,268	3,253	3,057	3,060	2,992	3,037
Aug	2,884	2,882	3,452	3,505	3,380	3,326	3,163	2,957
Sep	3,477	3,503	4,346	4,245	4,122	4,059	4,189	3,725
Oct	2,626	2,704	3,425	3,342	3,106	3,038	3,297	2,785
Nov	2,341	2,464	3,548	3,342	2,973	3,130	3,322	2,645
Dec	2,387	2,388	3,530	3,512	3,048	3,203	3,497	2,821

For the states which we do not have access to daily load profiles such as Uttar Pradesh, Bihar, Jharkhand and West Bengal, we used Chhattisgarh as a reference state. Daily load profiles of Chhattisgarh are rescaled by the ratio between total monthly demands collected from CEA's website. Tables B7 to B10 show the results.

Table B7: Monthly average demand for Uttar Pradesh

UttarPradesh (MWh)	12am-3am	3am-6am	6am-9am	9am-12pm	12pm-3pm	3pm-6pm	6pm-9pm	9pm-12am
Jan	24,292	27,233	33,150	27,233	24,752	27,233	33,647	27,233
Feb	28,303	29,554	32,160	31,142	28,047	31,087	32,824	29,776
Mar	24,149	27,073	32,955	27,073	24,606	27,073	33,449	27,073
Apr	23,759	28,079	30,202	28,079	23,106	28,079	31,219	28,079
May	29,523	31,432	34,598	31,432	29,027	31,432	35,671	31,432
Jun	32,387	35,236	35,796	35,236	30,911	35,236	38,694	35,236
Jul	29,136	32,872	35,935	32,872	28,524	32,872	36,918	32,872
Aug	28,927	31,240	34,230	31,240	27,936	31,240	35,206	31,240
Sep	27,495	31,553	33,840	31,553	26,696	31,553	35,122	31,553
Oct	27,917	30,629	32,500	30,629	27,552	30,629	33,678	30,629
Nov	26,060	28,582	34,436	28,582	25,707	28,582	35,301	28,582
Dec	27,174	29,432	34,838	29,432	26,613	29,432	35,487	29,432

Table B8: Monthly average demand for Bihar

Bihar (MWh)	12am-3am	3am-6am	6am-9am	9am-12pm	12pm-3pm	3pm-6pm	6pm-9pm	9pm-12am
Jan	4,326	4,850	5,904	4,850	4,408	4,850	5,992	4,850
Feb	4,994	5,215	5,674	5,495	4,949	5,485	5,792	5,254
Mar	4,689	5,257	6,399	5,257	4,778	5,257	6,495	5,257
Apr	4,592	5,426	5,837	5,426	4,466	5,426	6,033	5,426
May	4,445	4,732	5,209	4,732	4,370	4,732	5,370	4,732
Jun	4,969	5,406	5,492	5,406	4,743	5,406	5,937	5,406
Jul	4,742	5,350	5,849	5,350	4,643	5,350	6,009	5,350
Aug	4,843	5,230	5,731	5,230	4,677	5,230	5,894	5,230
Sep	4,171	4,787	5,134	4,787	4,050	4,787	5,329	4,787
Oct	4,647	5,099	5,410	5,099	4,586	5,099	5,606	5,099
Nov	5,046	5,535	6,668	5,535	4,978	5,535	6,836	5,535
Dec	4,843	5,245	6,208	5,245	4,742	5,245	6,324	5,245

Table B9: Monthly average demand for West Bengal

WestBengal (MWh)	12am-3am	3am-6am	6am-9am	9am-12pm	12pm-3pm	3pm-6pm	6pm-9pm	9pm-12am
Jan	10,215	11,451	13,939	11,451	10,408	11,451	14,148	11,451
Feb	12,168	12,706	13,827	13,389	12,059	13,365	14,112	12,802
Mar	12,265	13,750	16,737	13,750	12,497	13,750	16,988	13,750
Apr	13,075	15,452	16,621	15,452	12,716	15,452	17,180	15,452
May	14,221	15,141	16,665	15,141	13,982	15,141	17,182	15,141
Jun	15,326	16,674	16,939	16,674	14,628	16,674	18,310	16,674
Jul	13,955	15,744	17,211	15,744	13,662	15,744	17,681	15,744
Aug	14,551	15,714	17,218	15,714	14,052	15,714	17,709	15,714
Sep	14,135	16,222	17,397	16,222	13,725	16,222	18,056	16,222
Oct	13,838	15,183	16,110	15,183	13,658	15,183	16,695	15,183
Nov	11,371	12,471	15,025	12,471	11,217	12,471	15,403	12,471
Dec	10,443	11,310	13,388	11,310	10,227	11,310	13,637	11,310

Table B10: Monthly average demand for Jharkhan

Jharkhand (MWh)	12am-3am	3am-6am	6am-9am	9am-12pm	12pm-3pm	3pm-6pm	6pm-9pm	9pm-12am
Jan	1,977	2,216	2,697	2,216	2,014	2,216	2,738	2,216
Feb	1,985	2,073	2,256	2,184	1,967	2,180	2,302	2,088
Mar	1,963	2,200	2,678	2,200	2,000	2,200	2,719	2,200
Apr	2,039	2,410	2,592	2,410	1,983	2,410	2,679	2,410
May	2,192	2,334	2,569	2,334	2,156	2,334	2,649	2,334
Jun	2,177	2,368	2,406	2,368	2,078	2,368	2,601	2,368
Jul	2,206	2,489	2,721	2,489	2,160	2,489	2,795	2,489
Aug	2,243	2,422	2,654	2,422	2,166	2,422	2,730	2,422
Sep	1,999	2,294	2,461	2,294	1,941	2,294	2,554	2,294
Oct	2,036	2,234	2,370	2,234	2,009	2,234	2,456	2,234
Nov	2,221	2,436	2,935	2,436	2,191	2,436	3,009	2,436
Dec	2,131	2,308	2,732	2,308	2,087	2,308	2,783	2,308