Design and Optimization of Branched Piped Water Networks

Submitted in partial fulfillment of the requirements of the degree of

Doctor of Philosophy

submitted by

Nikhil Hooda

(Roll no: 114050009)

Under the Supervision of

Prof. Om P. Damani

Department of Computer Science and Engineering

Indian Institute of Technology Bombay, Mumbai 400076

2019

To my parents

Surender and Rachna Hooda

Approval Sheet

This thesis entitled "Design and Optimization of Branched Piped Water Networks" by Nikhil Hooda is approved for the degree of Doctor of Philosophy.

Examiners

5/3/2020

5/3/2020

Supervisors

O. P. Damani

Chairman

Date: 05th March, 2020

Place: Dept. of CSE, IIT Bombay

Declaration

I declare that this written submission represents my ideas in my own words and where others' ideas or words have been included, I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. I understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

Nikhil Surender Hooda

Roll No: 114050009

Date: 05/03/2020

Abstract

Government bodies responsible for drinking water distribution in developing countries face the challenging task of designing schemes that provide a quality of service that is adequate to meet the needs of the citizens at a cost below the strict government norms. Engineers at these government bodies must undertake the design process using tools that are not optimal and consider only pipe diameter selection, which is only one component of the entire scheme design. As such, much of the design process is undertaken in an ad-hoc and heuristic manner, relying on the experience and intuition of the engineers.

The problem of the capital cost optimization of branched piped networks consists of choosing pipe diameters for each pipe in the network from a discrete set of commercially available pipe diameters. Each pipe in the network can consist of multiple segments of differing diameters. Water networks also consist of intermediate tanks also known as Elevated Storage Reservoirs (ESRs). These act as buffers between incoming flow from the primary source and the outgoing flow to the demand nodes. The network from the primary source to the ESRs is called the primary network, and the network from the ESRs to the demand nodes is called the secondary network. During the design stage, the primary and secondary networks are optimized separately, with the ESRs acting as demand nodes for the primary network. Typically the choice of ESRs locations, their elevations, and the set of demand nodes to be served by different tanks is manually made in an ad-hoc fashion before any optimization is done. It is desirable therefore to include this ESRs configuration choice in the cost optimization process itself. Valves and pumps help control the water head in the network. The introduction of pumps also introduces operational cost to the objective function. Particular care must be taken since continuous high operational costs increases the chance of scheme failure.

Current tools used by the government consider only pipe diameter selection and provided non optimal designs. Other network components like ESRs, pumps and valves are designed manually and iteratively, resulting in a final design that is also non optimal. We present and implement a formulation that solves the problem of pipe diameter selection optimally. The solution provided is general i.e. each link can consist of any number of discrete pipe diameters. We extend the formulation to look at multiple components of a piped water network. In particular, the formulation captures the specific two stage design approach of rural networks in developing countries, with a partition of the network into primary and secondary networks. This allows the pipes, ESRs, pumps and valves

of the entire network to be optimized simultaneously instead of the manual iterative process of design and simulation employed today. Multiple refinements are done to the formulation to significantly improve performance. We prove that these improvements result in tighter models, i.e. the set of points of linear relaxation is strictly smaller than the linear relaxation for the initial model. The resulting model is guaranteed to be optimal and solves networks of real world importance in a matter of minutes.

The model has been implemented in JalTantra system which is free to use and publicly available at https://www.cse.iitb.ac.in/jaltantra/. Thus practitioners now have access to optimization and design system that is free and optimal and considers not just pipes but also tanks, pumps and valves. JalTantra includes usability features such as handling multiple file formats and has GIS functionality integrated for ease of providing network details. Maharashtra Jeevan Pradhikaran (MJP), the government body responsible for the planning, designing, and implementation of water supply schemes for the state of Maharashtra in India, has officially adopted it as one of the software packages to be used in the design of water supply schemes. Maharashtra Environmental Engineering Training and Research Academy (MEETRA), which is responsible for the training of MJP engineers, has integrated JalTantra into its curriculum. Details of the JalTantra system has been included in a compendium titled "Improving the performance of rural water supply and sanitation sector in Maharashtra".

Keywords: Water Distribution; Optimization; Integer Linear Program; Pipe Diameter Selection; Tank Configuration Selection; Pumps and Valves Selection

Contents

List of Figures

List of Tables

List of Abbreviations

CBC	COIN-OR Branch and Cut
CTARA	Centre for Technology Alternatives for Rural Areas
DMA	District Metered Area
ESR	Elevated Storage Reservoir
GA	Genetic Algorithms
GIS	Geographic Information System
GLPK	GNU Linear Programming Kit
GRG	Generalized Reduced Gradient
ILP	Integer Linear Programming
LP	Linear Programming
LPG	Linear Programming Gradient
MBR	Mass Balancing Reservoir
MJP	Maharashtra Jeevan Pradhikaran
MEETRA	Maharashtra Environmental Engineering Training and Research Academy
MLD	Million litres per day
MVS	Multi Village Scheme
NLP	Non Linear Programming
SDG	Sustainable Development Goals
SA	Simulated Annealing
WTP	Water Treatment Plant

Chapter 1

Introduction

Water is one of the most basic needs of human life. Water distribution networks are an integral infrastructure component for any society. Globally 750 million people still do not have access to an improved source of drinking water [5]. The problem is more acute in the rural areas of developing countries and is only going to exacerbate with growing populations, rising standards of living and increased awareness of the importance of clean water for health. Access to safe drinking water by 2030 was one of the 17 Sustainable Development Goals (SDGs) set by the General Assembly of the UN in 2015 [2].

Large scale projects are undertaken by governments to provide water to huge populations with a high initial capital cost as well as continuous operational and maintenance costs. These projects consist of several infrastructure components like pipes, tanks, pumps, valves etc. The designers of these projects need to choose the type, size, location and configurations for each of these components. These choices not only impact the quality of service but also impact the cost of the scheme. These systems are governed by complex nonlinear hydraulic equations and have to deal with uncertainty from various sources i.e. short term and long term demand changes, quantity and quality of water supply and component failures. The ability to pay the water tariff is often limited for people in rural areas. For large networks, invariably there are regions with worse coverage and a greater risk of failure. Any disruption of service quickly leads to people in that region unwilling to pay and reverting to previous unsafe local sources for their water needs. The economic stress added leads to further deterioration of the performance of the scheme effecting more and more people. This vicious cycle as seen in figure 1.1, leads to an eventual collapse of the entire scheme. As such given the costs and complexities involved and the crucial nature of the service being provided, these networks must be designed with great care.

The piped water networks for rural schemes are typically gravity fed, since reliable electricity supply is not a given. The most important aspect in the design of these systems is the choice of pipe diameters from a discrete set of commercially available pipe diameters.

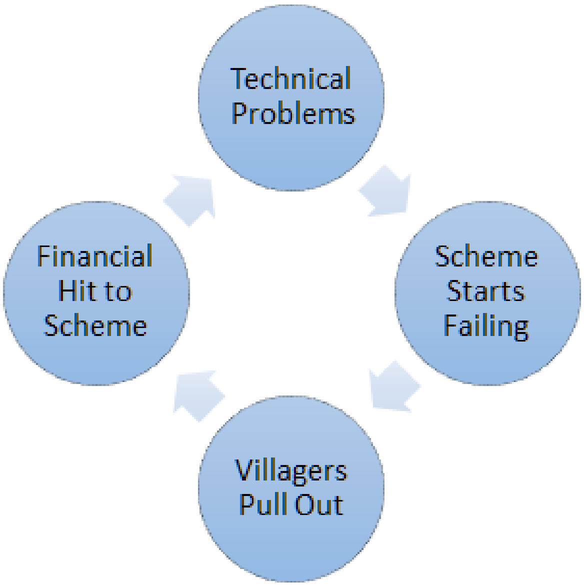


Figure 1.1: Vicious Cycle Leading to Scheme Failure.

In general, each link (connection between two nodes) can consist of several pipe segments of differing diameters. Larger the pipe diameters, better the service (pressure), but higher is the capital cost. The branched piped water network cost optimization problem is the selection of pipe diameters that minimize the system cost while providing the requisite service (pressure at demand points).

The water network design problem has been studied in various forms for over 50 years [38]. Different mathematical and algorithmic techniques ranging from deterministic ones like Linear Programming (LP), Non Linear Programming (NLP) etc. to modern metaheuristic ones like Genetic Algorithms (GA), Simulated Annealing (SA) etc. have been used over these past five decades. The networks under consideration can have different configurations. They can be branched or looped, gravity fed or pumped. Additionally, different subset of components of the network can be considered. Branched networks are common in rural areas since the redundancy provided by looped networks is an unaffordable luxury.

Maharashtra Jeevan Pradhikaran (MJP) is the government body responsible for the planning, designing and implementation of water supply schemes for the state of Maharashtra in India. It employs over 1500 engineers and over the past several decades has designed more than 11,000 rural water supply schemes. MJP when deciding to design and implement a scheme must adhere to strict government cost guidelines. In 2013, a study was undertaken by CTARA, IIT Bombay [33] to evaluate the feasibility of augmentation of the scope of the Karegaon scheme to include a cluster of 13 tanker fed villages in its neighbourhood. The primary objective of the study was the evaluation of the techno-economic feasibility of a multi village water supply scheme (MVS) to supply drinking water to the cluster of tanker fed villages in the neighbourhood of Karegaon scheme in Mokhada Taluka. A step by step process following guidelines and protocols used by MJP in their design process was used for this purpose. The secondary objective was to understand the process thoroughly and identify where and how the process could be improved. Details of the scheme design can be found in Appendix III.

Government bodies in India like the MJP use software like BRANCH, EPANET and WaterGEMS ([33], [69], [71], [79]) for the design of multi village schemes. BRANCH and LOOP [53] are optimization tools developed by the World Bank that attempt to minimize pipe cost for branched and looped pipe networks respectively. It is the software of choice for MJP engineers when designing a rural water scheme. Alternatively, some engineers use the commercial software WaterGEMS [4] to design and analyse water networks. Since it uses genetic algorithms, the cost optimization is heuristic and thus non-optimal. EPANET [68] is a water network modelling software that performs extended period simulation of the hydraulics of a water network. It is used to analyse and verify the network once its components have been designed.

Both BRANCH and WaterGEMS consider only the pipe diameter selection component of water network design. Beyond just pipe diameter selection, various other decisions need to be made regarding the components comprising the scheme design. Most choices like source selection, ESR distribution, valve location, etc. are made in an ad-hoc manner rather than optimizing overall. The designer's intuition and experience are relied on to make these choices. When using BRANCH for calculating pipe diameters in the design of Karegaon scheme, we found it has limited capabilities in terms of the number of pipes (at most 125) and does not guarantee an optimal solution. Despite BRANCH being free to use, due to difficulty in its usage, many designers even use spreadsheets to manually choose pipe diameters by a trial and error process.

Given the lack of free and optimal options we decided to develop an open source system, JalTantra that could not just optimally select pipe diameters but also aid the design of the other components of a typical water network. We ran several training sessions with government engineers on using JalTantra. Maharashtra Jeevan Pradhikaran (MJP) has officially adopted it as one of the software packages to be used in the design of water supply schemes. Maharashtra Environmental Engineering Training and Research Academy (MEETRA), which is responsible for the training of MJP engineers, has integrated JalTantra into its curriculum. Additionally, details of the JalTantra system has been included in a compendium titled "Improving the performance of rural water supply and sanitation sector in Maharashtra".

1.1 Objective

Our aim, in most general terms, is to help the real world practitioners in the design and optimization of piped water networks. In this work, we specifically focus on rural networks in developing countries. Government engineers in charge of designing these networks currently utilize software that only help design a part of the network and that too sub-optimally. Other components are designed by trial and error and by relying on the intuition and experience of the designer. Therefore our specific aim in the current work is to develop a general formulation that captures several network components and create a design and optimization system that implements this formulation. Given this context, here are some of the properties that the system must have:

- It must optimize the location and sizing of multiple network components like pipes, tanks, pumps and valves.
- The solutions it provides must be optimal, fast and consistent.
- It should be cross-platform, easy to use and be able to take inputs in legacy formats, so that users can transition to the new system easily. Modern technology like GIS should be leveraged to ease the design process. It must be extensible so that such new features can be easily implemented based on user feedback.

1.2 Contributions

Existing software used by government engineers in the design of water networks are non-optimal and restrict themselves to the optimization of pipe diameters only. Remaining components are designed by trial and error using the simulation software EPANET. In this work, we extend the problem and create a formulation that includes tanks, pumps and valves in addition to pipe diameters. This formulation is fast and optimal and is implemented in our free to use design and optimization system JalTantra.

In [55], an ILP formulation is proposed for the special case of one pipe diameter per link. While ILP in general is NP-hard, we found that for network sizes of real world interest, an optimal solution to the special formulation could be computed in reasonable time. This means that currently one can either get an optimal solution for the special case of one piped segment per link [55] or get a non-optimal solution for the general case of multiple pipe segments per link [53]. We propose a formulation that solves the general problem while still maintaining optimality. We implement this model into our network design system called JalTantra.

As part of extending JalTantra beyond just pipe diameter selection, we introduce other network components into the model. Water networks also consist of intermediate tanks also known as ESRs (Elevated Storage Reservoirs) that act as buffers between incoming flow from the primary source and the outgoing flow to the demand nodes. The network from the primary source to the ESRs is called the primary network, and the network from the ESRs to the demand nodes is called the secondary network. During the design stage, the primary and secondary networks are optimized separately with the ESRs acting as demand nodes for the primary network. Typically, the choice of ESR locations, their elevations, and the set of demand nodes to be served by different ESRs, is manually made in an ad-hoc fashion before any optimization is done. It is desirable therefore to include this ESR configuration choice in the cost optimization process itself. Therefore, we extend our model to an Integer Linear Program (ILP) model that integrates the same to the standard pipe diameter selection problem. Other components like valves and pumps are also incorporated into the network model. The inclusion of pumps in particular is significant since it means that apart from the capital cost, operational cost also needs to be considered.

The extension to an Integer Linear Program model has a significant impact on the performance of the model. Large networks whose pipe diameter was selected in a matter of seconds, need hours for the ESR configuration selection. In order to improve performance, we refine the model in several iterations. These involve exploring alternative formulations and tightening existing constraints i.e. reducing the set of points in the linear relaxation of the set described by the constraints. For each of the refinements, we prove that the resulting model is a strict improvement over the previous model. This brings down the time taken to solve even large networks of 200 nodes to a few minutes.

In summary, in this thesis we make the following contributions:

- Existing software, used by government agencies in developing countries, look at only the pipe diameter selection problem and do so non optimally. Other network components like ESRs, pumps and valves are designed manually and iteratively, resulting in a final design that is also non optimal. We first consider the pipe diameter selection problem and present a formulation that solves the pipe diameter selection problem optimally. The solution provided is general i.e. each link can consist of any number of discrete pipe diameters.

- We broaden the problem statement and extend the formulation to consider the other components of a piped water network i.e. ESRs, pumps and valves. In particular, the formulation captures the specific two stage design approach of rural networks in developing countries, with a partition of the network into primary and secondary networks. This allows the pipes, ESRs, pumps and valves of the entire network to be optimized simultaneously instead of the manual iterative process of design and simulation employed today.

- Multiple refinements are done to the model by exploring alternative formulations and tightening existing constraints i.e. reducing the set of points in the linear relaxation of the set described by the constraints. For each of the refinements, we prove that the resulting model is a strict improvement over the previous model. Networks of real world importance can be solved optimally in a matter of minutes.
- We implement these provably optimal and fast formulations in our web based network design system, JalTantra. Thus, we fill an important gap in the real world design of piped water networks for rural communities in the developing world. This has led to the adoption of JalTantra by government bodies in the design of such networks.

1.3 Structure of the Thesis

The thesis is structured as follows:

Chapter 1: Introduction. We introduce the problem of design and optimization of piped water networks. We motivate the importance of robust design and the difficulty in achieving it. We conclude the Chapter by stating the objective and the contributions of the thesis.

Chapter 2: Literature Review. We go over the existing literature on design of water networks. We discuss various approaches employed over the years for both branched and looped water networks. We further provide some background on closely related problems of network reliability, operation and generation.

Chapter 3: Components of a Rural Piped Water Scheme. We describe the components of a typical rural piped water scheme. We describe how each of the component contributes to both quality of service provided and the cost of the scheme.

Chapter 4: Pipe Diameter Optimization. We look at various approaches to the selection of pipe diameters for a water network. The first approach fixes each link in the network to at most one pipe diameter. We then extend this to at most two pipe diameters for each link using the result that in the optimal case at most two pipe diameters would be required. We show how a general LP model outperforms the two pipe approach in terms of time taken while providing the same final optimal solution. We also extend the model to include laying of parallel pipes in the case of augmentation of existing schemes.

Chapter 5: ESR Sizing and Allocation. We describe the ESR configuration problem for a piped water network and how the introduction of ESRs splits the water network into primary and secondary networks. We motivate the importance of including

ESR configuration into the optimization process due its impact on total capital cost. We then extend the general LP model for pipe diameter selection to an ILP model that considers ESR configuration and the resulting constraints that influence the hydraulics in the system.

Chapter 6: Pumps/Valves Integration. We describe the impact including pumps and valves has in a water network by increasing/decreasing the pressure head available in the system. The introduction of pumps in particular extend the cost to not just the initial capital cost but also an ongoing operational cost that must be considered. We extend our ILP model to include pumps and valves and modify the objective to the sum of capital and operational cost.

Chapter 7: Model Improvements. The iterative extension of the model to include more complicated constraints and objectives led to a deterioration in performance. In this chapter, we describe three major improvements that were made to tighten the model. For each, we prove that the linear relaxation of the new constraints is a strict subset of the linear relaxation of the previous points. We then show practical results of these improvements by comparing performance over eight water networks.

Chapter 8: Edge Based Model. The ILP model described so far uses node based variables to capture the partitioning of the network into primary and secondary networks. In this chapter we briefly look at an alternative edge based model. We show that the set of constraints that describe valid network configurations using edge based variables is tight. The performance overall remains worse than the node based approach.

Chapter 9: JalTantra System Description. We describe the JalTantra web system that implements the model to optimize network components.

Chapter 10: Conclusion and Future Work. We conclude the thesis by summarizing the work done and suggest some future research directions.

Appendix I: Complete ILP Model. We describe the variables, constraints and the objective of the complete ILP model.

Appendix II: JalTantra Usage Details. We describe details on the various tabs of the JalTantra system and how to use the system.

Appendix III: Karegaon Scheme Redesign. We provide details of the Karegaon scheme redesign done in 2013, which led to the research work presented in this thesis.

Chapter 2

Literature Review

The design of a multi village piped water network involves the choosing the sizing and location of the various networks components that make up the network. These choices determine the level of service provided as well as the cost of the network. As early as 1895 [67], networks were being manually designed using the economic velocity principle. Piped water network cost optimization using computer science techniques has been studied for more than 50 years now. Various approaches have been employed to different versions of the network optimization problem. Different versions of the problem involve either looking at different subsets of the components of the network or making different assumptions about the network configuration.

2.1 Branched Network Design

The earliest approaches looked at solely the pipe diameter optimization problem. In this case the demand and the network topology were fixed, and the pipe diameters were the variables to be determined such that sufficient heads were achieved at each of the nodes while minimizing the capital cost of the pipes. A piped water network can be branched or looped i.e. the network links are connected such that there is at least one loop in the network. Rural networks are typically branched. Such networks are easier to optimize since conservation of mass equations can be used to determine the flow required in each pipe of the network. In looped networks however, apart from the pipe diameter, flow is an additional parameter to be determined.

2.1.1 Linear Programming Techniques

In 1968, Karmeli et al. [38] used a Linear Programming (LP) model for determining pipe diameters in a branched network. The head loss in a pipe is computed using the Hazen-Williams equation [74] and the cost of the pipe is assumed to be a linear function of the length. The choice of diameters for the pipes is made from a discrete set of commercially available pipe diameters with the unit cost for each diameter known a priori. Since the network is branched, the flow through each link is known and the headloss per unit length can be pre computed for each of the commercially available pipe diameters. The above model is extend by Calhoun [11] to include a pump at the source of the network. Thus the source head becomes an additional parameter to be optimized. The model is further extended by Robinson and Austin [54] by considering pressure ratings for pipes and including pressure reducing valves. These inclusions therefore introduce maximum head constraints in the model. Additionally the cost of source head is now considered to be nonlinear and therefore the network is designed iteratively by solving a series of LP models. Although providing optimal solutions, these LP models were found to be too large for networks with large number of pipes. Bhave [6] presented a heuristic method which reduced the number of candidate pipe diameters being considered.

2.1.2 Non Linear Programming Techniques

An alternate approach to heuristically reducing the search space was to attempt to obtain an analytical solution. Mandry [48] considers the cost to be an explicit function of the diameter, rather than looking up unit cost of discrete diameter sizes. The resultant model therefore contains non linear terms and Non Linear Programming (NLP) techniques are used to solve the model. Swamee et al. [65] extend the model to include initial head as a variable, with both pumping and overhead tank costs considered. The solution provided by these models however contains continuous pipe diameters instead of the discretely sized commercially available pipe diameters. For any practical application the suggested pipe diameters would have to be rounded to discrete diameters, which renders the solution non optimal. Fujiwara and Dey [23] propose a two step approach using both LP and NLP formulations. In the first step an NLP formulation similar to the one proposed by Swamee et al. [65] is used to obtain continuous pipe diameters. These are then used to create a subset of candidate discrete diameters from all available ones. Then in the second stage a LP model similar to Karmeli et al. is employed to obtain the optimal solution. These NLP techniques however make the assumption that the minimum head required at all end nodes are equal. Additionally, no minimum head requirements are considered for any internal nodes in the network. Young [78] proposes a NLP approach which can handle both non equal head requirements and consider internal nodes.

2.2 Looped Network Design

Branched networks are typically used in the case of rural areas. Urban areas however contain looped networks since they provide additional reliability. In branched networks, if demands are known, the flow through each pipe in the network is fixed. The complexity introduced by allowing networks to be looped is that the flows in the pipes are now variables. This causes the optimization of looped networks problem to be necessarily nonlinear and nonconvex [20]. Techniques used to solve looped networks, be they deterministic or metaheuristic, have a common iterative approach. An initial solution is proposed which is then tested and verified using a hydraulic solver. Then depending on the output from the hydraulic solver variables are modified to generate a new solution. These process is iteratively repeated till a predetermined stopping criteria is met. The difference in the various approaches lies in what components of the network are being considered and how candidate solutions are modified at each iteration.

2.2.1 Deterministic Techniques

In 1968, Jacoby [34] proposed a non linear integer program to design a looped network. The diameters considered are continuous which are later rounded off to integers. Watanatada [73] extends the model to include flow rate and minimum head requirement constraints. Shamir [62] further extends the Watanatada model and considers multiple demand patterns. These approaches use gradient methods. An initial distribution of flows and diameters is assumed for the network which satisfy all the constraints. The gradient is then computed for the objective function. Diameters and flows are moved on the gradient to give the next set of variables. This process is then repeated till the stopping criterion is met. Lansey and Mays [40] use a model similar to Shamir [62]. Additionally, they utilize a network solver for the optimization. The variables are separated into two sets, dependant(node heads) and independent(diameters and flows). For a given set of independent variables, the network solver computes the dependant variables using hydraulic equations. The optimizer then computes a new set of independent variables using a Generalized Reduced Gradient (GRG) method and the process is repeated. Alperovtis and Shamir [1] introduce a Linear Programming Gradient (LPG) method to design looped networks. Initially a set of fixed values is assumed for the flows in the links. Then a LP model similar to Karmeli [38] is run to compute the pipe diameters. The flows are then modified using a gradient method and the process is repeated till the stopping criterion is met. Fujiwara et al. [25] use a method similar to [1] and improve on the modifying flows step which results in a faster convergence. Eiger et al. [20] show that any formulation of the looped network optimization must be nonconvex and nonlinear. They also show that previous gradient approaches ignore the fact that the gradient of the objective function need not always exist. Eiger et al. explicitly consider this and propose a nonsmooth optimization algorithm which provided local optima. The problem is optimized globally by using branch and bound techniques. Samani and Mottaghi [55] use binary variables to represent if a particular pipe diameter is used for a link and therefore use an Integer Linear Program (ILP) to model the problem. The use of binary variables results in a single diameter being suggested for each link.

2.2.2 Metaheuristic Techniques

Most recent work in the design of piped water networks has been in utilizing various meta-heuristic techniques. These techniques encompass a range of "meta" level algorithms used to explore large search spaces in order to find optimal solutions. They are "meta" level since they are independent of the specific problem being considered. This is unlike most heuristic approaches which attempt to exploit some structure unique to the problem under consideration. Constraints are modelled using penalty functions that are added to the objective cost function. Fixing the values for the set of variables determines the objective cost function. The variables are then iteratively modified in several generations till a stopping criterion is met. The difference in the various techniques lies in how this iterative modification of variables is done. Meta-heuristic techniques are inspired from various phenomena observed in the natural world.

The largest subset of these meta heuristic techniques that have been used in the water network design problem are genetic algorithms (GA). Simpson et al. [63] used GAs to optimally determine pipe diameters for piped water networks. GA is inspired from evolution theory. In GA several instances of the decision variables are simultaneously considered and maintained. Each instance represents an individual in a population. With each iteration, individuals are modified by reproduction (elements from two individuals are combined to create a new individual) and mutation (some elements of an individual are randomly changed). "Fitness" (objective cost) for the new set of individuals is then computed. More fit (lower cost) individuals have a higher probability of surviving and being selected for future reproduction and mutation. Dandy et al. [16] suggest an improved GA model which uses a modified fitness function and mutation is done only to adjacent pipe diameters. Montesinos et al. [50] also use a modified GA where in each generation a fixed number of the population is discarded. Additionally each individual can undergo at most one mutation. Wu and Simpson [75] use GA to design networks which include pumps, tanks and valves. They use a separate network solver like in Lansey and Mays [40]. Babayan et al. [3] extend the GA model to include demand uncertainty.

Various meta heuristic techniques, other than GA, have also been employed to design piped water networks. These include: Simulated Annealing (Cunha and Sousa [15]) which is inspired by the annealing process where the cost is allowed to increase with a certain probability allowing one to escape local optimas; Tabu Search (Cunha and Riberio [14]) which is inspired by human memory processes and tracks the list of already explored solutions; Ant colony optimization (Maier et al. [44]) which is inspired by group behavior of large ant colonies where each "ant" looks at and optimizes one component of the overall structure; Shuffled Frog Leaping Algorithm (Eusuff and Lansey [21]) which is inspired by evolution of memes (cultural evolution) among a population of frogs where individuals can be modified not just by parents (in the case of GAs) but also peers.

With the increase in usage of meta-heuristic techniques, multi objective optimization approaches began to be considered. Halhal et al. [32] first applied a multi objective approach to network design, attempting to maximize network benefit while minimizing network cost. Mala Jetmarova et al. [47] provide a detailed review of various approaches to the optimization of water distribution systems.

Over the decades of research on the design of looped networks, several small benchmark networks such as Hanoi network [25], New York City tunnels [60] and Two Loop network [1] have been used to compare performance. The two loop network is a simple network consisting of 8 links(each of length 1000m), 6 demand nodes and 14 commercially available pipe diameters. We compare the performance of some of the approaches described above using this network in Table 2.1.

Table 2.1: Comparison of performance of various approaches on the two loop network

	Alpervotis and Shamir (1977)		**Eiger et al. (1994)**		**Savic and Walters (1997)**		**Samani and Mottaghi (2006)**	
Pipe	**Length (m)**	**Dia (in.)**	**Length (m)**	**Dia (in.)**	**Length (m)**	**Dia (in.)**	**Length (m)**	**Dia (in.)**
1	256 744	20 18	1000 0	18	1000 0	18	1000 0	18
2	996.38 3.62	8 6	238.02 761.98	12 10	1000 0	10	1000 0	10
3	1000 0	18	1000 0	16	1000 0	16	1000 0	16
4	319.38 680.62	8 6	1000 0	1	1000 0	4	1000 0	4
5	1000 0	16	628.86 371.14	16 14	1000 0	16	1000 0	16
6	784.94 215.06	12 10	989.05 10.95	10 8	1000 0	10	1000 0	10
7	1000 0	6	921.86 78.14	10 8	1000 0	10	1000 0	10
8	990.93 9.07	6 4	1000 0	1	1000 0	1	1000 0	1
Cost	497.32		402.35		419		419	

2.3 Related Problems

Apart from the primary objective of determining network components, there are a range of problems related to water distribution networks. These range from additional considerations during the design stage like reliability analysis to network operation and maintenance once the scheme has been implemented.

2.3.1 Reliability Considerations

So far the only objective being optimized is the cost of the scheme. Goulter and Coals [30] explicitly include reliability in the design of looped networks. Each available pipe diameter is assumed to have a rate of failure per unit length per year. It is assumed than any connecting link can satisfy the entire demand of a node. The probability of node isolation is thus computed by computing the probability of failure of each of its connecting links. A constraint is added to the model for each node restricting this probability within acceptable limits. Goulter and Bouchart [29] extend the above model to also model demand failure i.e. the probability that demand at a node exceeds the design values. This is combined to the link failure probability and a single reliability metric is considered. Bouchart and Goulter [10] consider demand not only at nodes but also along the entire length of the links. Su et al. [64] utilize the approach of using a network solver and the GRG method proposed in [40] and include pumps, tanks and reliability constraints. Lansey et al. [41] propose a chance constrained NLP model that considers uncertainties in demands, required pressure heads and pipe roughness. This model is also solved using GRG. Xu and Goulter [76] extend the Lansey et al. [41] model to include uncertainty of pipe failure. Duan et al. [19] further extend the model to consider pump failure.

The above reliability analysis is modelled as a failure of components or under conditions of demand uncertainty. However, in areas with limited water availability a common point of failure is the uncertainty in supply. Under normal supply conditions the hydraulics of the system are demand driven i.e. assuming some demand for various nodes and assuming that those nodes are being fulfilled. Software like EPANET [68] which are used to analyse these networks assume demand driven hydraulics. But under low supply and pressure conditions the flows are limited not by the demands but by the pressure available in the system. Therefore the network needs to be analysed using pressure driven hydraulics. Bhave [7] proposed a network analysis method which represents the flow as function of available pressure. Gupta and Bhave [31] extend this approach to consider reliability under pressure deficient considerations where reliability is based on a node-reliability factor, volume-reliability factor, and network reliability factor. Sayyed et al. [59] use the demand driven analysis of EPANET to simulate a pressure driven regime by adding artificial valves and emitters to each node. Mahmoud et al. [43] extended this approach and improved the scalability by adding the artificial components to a selected number of nodes. This selection was determined by first doing a normal demand driven analysis and only those nodes with negative pressures were modified.

2.3.2 Network Generation

In all the above approaches the network layout was assumed to be fixed and known before the design stage. In 1983, Bhave and Lam [8] consider the problem of determining the optimal layout of a network given only the locations of the demand nodes and the supply points. The problem is modelled as a Steiner tree problem where the objective is to minimize the total length of the links required to connect all the nodes. Lejano [42] develops a Integer Linear Programming (ILP) formulation for optimizing network configuration. The objective to be optimized is empirical and includes the capital cost of pipes and pumps as well as the cost of serving each customer. Therefore it allows for selecting a subset of customers to be served instead of serving everybody as is typical.

Despite the numerous approaches to the design of networks, their testing has been inadequate due to a lack of benchmark networks. They are limited to a few small sized networks such as Hanoi network [25], New York City tunnels [60] and Two Loop network [1]. This leads to most metaheuristic techniques overfitting the few instances available. De Corte and Sorenson [18] have developed a tool, HydroGen that can algorithmically generate large scale water networks with characteristics resembling those of real life networks.

2.3.3 Network Operation

Even after the network has been designed the operation and maintenance of large water networks remains a complex problem. The operation of these networks broadly encompasses two areas. Firstly, pump and valve operation which is used to ensure adequate supply of water under varying and uncertain demand patterns. In particular optimization of pump usage is crucial since it constitutes the largest expenditure in the operation of water networks [70]. The second aspect is the optimization of water quality in the network.

As with network design, work on network operation has followed a similar path, starting with deterministic techniques in the 1970s ([1] , [61], [13]) to metaheuristics ([45], [51]) after the 1990s. Sankar et al. [56] provide an algorithm that looks at the problem of optimal pump operation in distressed situations where water supply is insufficient. The optimization thus has to consider cost but also equitable distribution of water. Mala

Jetmarova et al. [46] provide a detailed review of the work done in this field over the past three decades.

Related to network operation is the problem of network maintenance. Leaks are common due to the pressured environments and result in massive losses of water [66]. In order to combat leaks and better manage existing networks, urban networks are increasingly getting partitioned into independent sections called District Metered Areas (DMAs). This partitioning is achieved by cutting off parts of the network from each other using valves. Savic and Ferrari [57] propose an automated method to create DMAs for an existing network and compare the characteristics of candidate solutions.

Chapter 3

Components of a Rural Piped Water Scheme

The basic premise of a multi village piped water scheme is that water needs to be transported from a set of common sources of water to a group of villages. This is achieved by several components pipes, pumps, reservoirs, treatment plants, valves etc. The choices made in regard to these components contribute to the cost of the scheme as well as the quality of the service provided. The aim is to minimize cost while maintaining desired service quality. We briefly go over the components that comprise a typical scheme and detail how each impact the scheme cost and quality. The layout of a typical multi village scheme is depicted in Figure 3.1 below (courtesy: CTARA, IITB).

Water is pumped from the source to the water treatment plant (WTP) by the raw water rising main. From the WTP it is then pumped to the mass balancing reservoir (MBR) by the pure water rising main. Then we have the primary network where the water is distributed from the MBR to several Elevated Storage Reservoirs (ESRs). Then comes the secondary network where the water goes from the ESRs to the individual hamlets. Finally, we have the tertiary network where water goes from the hamlets to individual stand posts/homes.

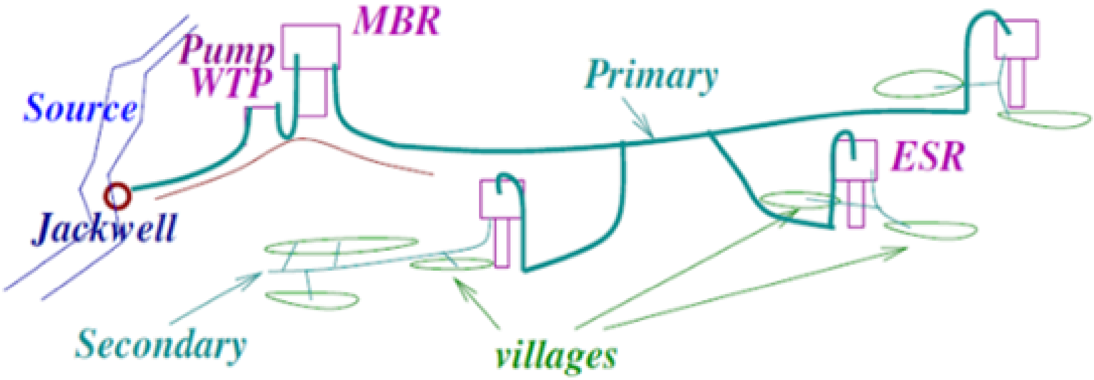


Figure 3.1: Components of a typical Rural Piped Water Scheme.

3.1 Water Source

The water source can be surface water (like lakes, reservoirs, rivers etc.) or ground water. Several different sources might be available from which one or more sources must be selected to service the scheme. The choice of source depends on several factors like head of the source, location in relation to the rest of the network, quality of water available, amount of water that can be sustainably drawn and reliability in times of stress (summer months). In the present work we assume there is a single known water source which provides a constant water head to the network irrespective of the amount of water drawn.

3.2 Water Treatment Plant (WTP)

The water that is drawn from the source needs to be treated at a WTP. WTPs vary in the kind of treatment that they can provide (i.e. which chemicals are used for treatment) and their capacity (i.e. amount of water that the WTP can process in a day). The factors influencing the choice of WTP are the supply of water that the scheme needs to provide, the quality of source water and the task for which water will be consumed (drinking, irrigation, etc.). In the present work the design of the WTP component is ignored since we do not consider water quality. The network is designed post the water treatment with the MBR effectively acting as the source of water. This is because the choice of type and size of WTP can be made independent of the network design. The type of WTP depends on the quality of water available. The size of the WTP depends on the total demand which is known beforehand.

3.3 Mass Balancing Reservoir (MBR)

Water from the WTP is stored in the MBR and then released into the rest of the network. Thus, it serves as a buffer in the supply of water to the rest of the network. It acts as an "effective" source of the network and the water head it provides is a key component in shaping the rest of the network. The choices to be made in regards to the MBR are its location, elevation and sizing.

3.4 Elevated Storage Reservoirs (ESRs)

These are the reservoirs that are placed at various points of the network from which water is delivered to the villages. Water is supplied to these ESRs from the MBR. Each ESR can serve one or more villages. This depends on the relative location of the villages and the ESR, as well as the amount of water demanded by the villages and the capacity of the ESR. An important consideration is the elevation of the ESR as this impacts the choice of pipes in the primary and secondary networks. If the location of the ESR is at a relatively high elevation, the ESR might be placed on the ground and is termed as a Ground Storage Reservoir (GSR). In the present work, we assume that if ESRs are included in the optimization, then each ESR must be served by the source and each demand point must be served by exactly one ESR.

3.5 Pipes

Pipes are the backbone of the water distribution network. As water flows in the pipe, there is a loss in the water head along the length of the pipe. This is caused by friction losses due to the movement of water in the pipe. This headloss depends on factors like the diameter, length and material of the pipe as well as the amount of water flowing in the pipe. The choice of pipe therefore depends on all these factors. The headloss in a pipe is calculated by the Hazen-Williams equation [74]. Also, pipes need to withstand the water pressure that is applied at all times. The amount of pressure that a pipe can withstand is represented by its pressure rating. There are several types of pipes available with varying reliability and pressure ratings [49]. In particular for distribution networks the following three types are commonly used:

- Reinforced Cement Concrete (RCC): Cheap and available in large diameters (upto 2000mm). However can be subject to corrosion if water is acidic and extremely heavy, thus difficult to handle.

- Ductile Iron (DI): Costlier than RCC but better corrosion resultant and thus longer lasting. Also available in large diameters (upto 1000m).

- High Density Polyethylene (HDPE): Costliest of the three options. Light and corrosion resistant. However, requires specialized welding, is easier to break and is not available in high diameters (upto 630mm).

In the present work we assume, for each pipe diameter there is a single type of pipe available. This is because the choice of pipe is not restricted to hydraulic considerations and cost. The choice is also determined by operational concerns like the quality of water available and the reliability of pipe desired which is outside the scope of the current work.

3.6 Pumps/Valves

Pumps and valves help manipulate the water head in the network. Pumps are required to supply water where water cannot be supplied naturally i.e. via gravity. The power of the pump required depends on the amount of water to be pumped as well as the water head that needs to be provided. Pumps are one of the primary sources of operational cost and as such their usage should be strictly on an as needed basis. In the present work, the efficiency of a pump is assumed constant irrespective of the flow through the pump. This makes the pump power a linear function of the head provided and the flow through the pipe. In practice, the efficiency varies over different water flow and head values, which is represented by a characteristic pump curve. While determining operational cost only the cost due to pumps is considered. Operational costs arising due to operation of WTP and maintenance costs are ignored. This is because these costs are independent of the design of the rest of the network. At times part of the network has excess water pressure due to huge elevation differences. Pipes therefore must be chosen which can withstand such pressures leading to a rise in capital cost. If a lower head can serve for the downstream network, pressure reducing valves can be employed to reduce the pressure and thus allow the use of pipes with lower pressure rating.

Additionally, in the present work we make the following assumptions about the network:

- Typically the demand at a location would be distributed across many households. We aggregate the demand at a village into a single point. If required, the tertiary network which distributes the water within a village can be separately designed.

- The demand at each point is known and constant throughout the day. This is known as steady state analysis. In practice however, once the network has been designed extended period simulation is required to verify the functioning of the network. Our system JalTantra outputs the result of the optimization as an EPANET file which can be used to perform the extended period simulation.

- The network configuration i.e. the links connecting the demand points is fixed, known and branched i.e. there are no loops in the network. This fixes the flow through each pipe since the demands are known.

Under these assumptions, we develop our design and optimization model by considering pipes, tanks and pumps/valves over the next three chapters. We then discuss the modifications made to the model to improve performance in Chapter 7.

Chapter 4

Pipe Diameter Optimization

In this chapter we describe the primary problem of network design, selection of pipe diameters. We first look at an approach from literature that restricts each link in the network to at most one pipe diameter. We then extend this model to at most two pipes per link which is guaranteed to be optimal. We show how a more general model which allows any number of diameters per link results in a much faster formulation. We conclude by extending the model to include parallel pipes in the case of augmentation of existing schemes.

4.1 OnePipe Model

Although in general each link in the network can consist of multiple pipe diameters, the first formulation that we looked at for our problem was to restrict each link in the network to a single pipe diameter. [55] has an Integer Linear Programming (ILP) formulation for the problem where only one pipe diameter is used per link. It uses binary variables to represent the choice of commercial pipe to be made for each link. Henceforth we use the term OnePipe model while referring to this formulation.

4.1.1 The Objective Function

The objective function to be minimized is the total cost of the pipes chosen for the links in the network:

$$O(.) = \sum_{i=1}^{NL} L_i C_i(D_i). \tag{1}$$

Where,

> $O(.)$: The total pipe cost which is a function of the pipe diameters chosen for each link
>
> NL: The number of links in the network

D_i: Pipe diameter for the link i

L_i: Length of link i

C_i: Cost/length of link i, a function of the pipe diameter D_i

The diameters D_i can only be chosen from the set of available commercial pipe diameters. This restriction is represented via binary variables x_{ij}. The modified objective function is:

$$O(.) = \sum_{i=1}^{NL} \sum_{j=1}^{NP} L_i C_{ij}(D_{ij}) x_{ij}. \tag{2}$$

Where,

NP: Number of commercially available pipe diameters

x_{ij}: 1 if link i uses pipe diameter j, else 0

4.1.2 Pipe Constraint

For each link the objective function $O(.)$, has terms for each of the available commercial pipe diameters j. Since each link is to be installed with only one diameter, it means that exactly one of the binary variables corresponding to each link must be one. Therefore we get the following pipe constraint for each link i:

$$\sum_{j=1}^{NP} x_{ij} = 1 \ , \qquad i = 1, \ \ldots, \ NL. \tag{3}$$

4.1.3 Node Constraint

At each node a minimum amount of pressure needs to be maintained. The pressure at any node is calculated from the headloss in the pipes connecting the node to the reference node i.e. the source for the network. Therefore the pressure constraint for each node n is:

$$P_n \leq H_R - E_n - \sum_{i \in S_n} hl_i \ , \qquad n = 1, \ \ldots, \ NN. \tag{4}$$

Where,

P_n: The minimum pressure that must be maintained at node n

H_R: The head supplied by reference node

E_n: The elevation of node n

NN: The number of nodes in the network

S_n: Set of pipes that connect node n to the reference node R

hl_i: Headloss in link i. It is modeled as summation over HL_{ij} i.e. headloss in link i if the pipe diameter j is chosen. We use the Hazen-Williams formula [74] for headloss which is given by:

$$HL_{ij} = \frac{10.68 * L_i * \left(\frac{FL_i}{R_j}\right)^{1.852}}{D_j^{4.87}}, \qquad i = 1, \ldots, NL, \quad j = 1, \ldots, NP. \tag{5}$$

Where,

FL_i: Flow in link i

R_j: Pipe roughness of commercial pipe j

D_j: Diameter of commercial pipe j

As with the objective cost function, we introduce x_{ij} to linearize the node constraint which gives us:

$$P_n \leq H_R - E_n - \sum_{i \in S_n} \sum_{j=1}^{NP} HL_{ij} x_{ij}, \qquad n = 1, \ldots, NN. \tag{6}$$

Now for a given pair of (i, j) values, HL_{ij} is just a constant, so we get a linear constraint with x_{ij} as our variables.

4.2 TwoPipe Model

The previous formulation assumes a single pipe diameter for each link in the network. But this might not be (and usually is not) the optimal solution. In the most general case, we can have several pipes of varying diameters for a single link in the network. In [24] it is shown that in the optimal solution, each link will consist of at most two pipe segments of adjacent diameters. The above will hold if the commercial pipe cost is a convex function of a power of its diameter, which is the case in practice. We modify the objective function to incorporate this knowledge about the optimal solution structure as described below.

4.2.1 The Objective Function

To generalize the formulation in order to account for multiple pipes per link we introduce two new continuous variables for each link: l_{i1} and l_{i2}. The previous binary variable representing the choice of commercial pipe for each link is now broken into two: x_{ij1} and

x_{ij2}. Here l_{i1}/ l_{i2} represent the length of the two pipes for link i and x_{ij1}/x_{ij2} represent the choice of diameters for the two pipes. The modified objective function is:

$$O(.) = \sum_{i=1}^{NL} \sum_{j=1}^{NP} l_{i1} C_{ij1}(D_{ij1}) x_{ij1} + l_{i2} C_{ij2}(D_{ij2}) x_{ij2}. \tag{7}$$

This formulation is not linear since we have terms like $l_{i1}x_{ij1}$ which are a product of two variables. We linearize this equation by introducing variables z_{ij1}:

$$z_{ij1} = l_{i1} x_{ij1} \ , \qquad i = 1, \ \ldots, \ NL, \quad j = 1, \ \ldots, \ NP. \tag{8}$$

This nonlinear equality is represented by the following set of linear inequalities:

$$z_{ij1} \leq L_i x_{ij1} \ , \tag{9}$$

$$z_{ij1} \leq l_{i1} \ , \tag{10}$$

$$z_{ij1} \geq l_{i1} - L_i(1 - x_{ij1}) \ . \tag{11}$$

This linearization is possible since z_{ij1} is a product of a continuous variable and a binary variable. A similar set of inequalities will be introduced for z_{ij2} as well. Our new objective function:

$$O(.) = \sum_{i=1}^{NL} \sum_{j=1}^{NP} C_{ij1}(D_{ij1}) z_{ij1} + C_{ij2}(D_{ij2}) z_{ij2} \ . \tag{12}$$

4.2.2 Pipe Constraint

As before, the choice of pipe diameter must be made from the available commercial pipe diameters.

$$\sum_{j=1}^{NP} x_{ij1} = 1 \ , \qquad i = 1, \ \ldots, \ NL \ . \tag{13}$$

Similar equalities hold for x_{ij2}. We also have an additional constraint for each link i.

$$l_{i1} + l_{i2} = L_i \ , \qquad i = 1, \ \ldots, \ NL \ . \tag{14}$$

4.2.3 Node Constraint

The node constraint for minimum pressure requirement now changes since length of each pipe segment is no longer a constant. Since headloss is linear in the length of the pipe (5), the length term can be taken out. The headloss HL_i in each link i now becomes:

$$HL_i = \sum_{j=1}^{NP} HL_{ij}(x_{ij1} l_{ij1} + x_{ij2} l_{ij2}) \ , \qquad i = 1, \ \ldots, \ NL \ , \tag{15}$$

$$HL_{ij} = \frac{10.68 * \left(\frac{FL_i}{R_j}\right)^{1.852}}{D_j^{4.87}} \;, \qquad i = 1, \; \ldots, \; NL, \quad j = 1, \; \ldots, \; NP \; . \tag{16}$$

Replacing the $x_{ij}l_{ij}$ terms with z_{ij} gives the new node constraint:

$$P_n \leq H_R - E_n - \sum_{i \in S_n} \sum_{j=1}^{NP} HL_{ij}(z_{ij1} + z_{ij2}) \;, \qquad n = 1, \; \ldots, \; NN \; . \tag{17}$$

4.2.4 Computational Results

On testing the new formulation, the run-time performance is found to be very poor. Whereas for the OnePipe model a 100 node network can be solved in 1.5 seconds, the new formulation cannot solve optimally even a 10 node network before getting timed out at 100 seconds.

The reason for this is the large number of new constraints that are added to the system. Previously all constraints were linear in the number of nodes. We had only one constraint for each link and one for each node (for an acyclic network, number of links is one less than the number of nodes). But now, in order to introduce z_{ij1} and z_{ij2}, we have introduced six new constraints for each (link, pipe diameter) combination. Thus the number of constraints goes from $O(n)$ for the OnePipe model to $O(n^2)$ in the TwoPipe model.

4.3 General Model

If multiple pipes per link are permitted, at most two pipes of adjacent diameters will be chosen [24]. In the TwoPipe model this knowledge is explicitly used by modelling each link to be made up of two segments. We then independently determine the commercial pipe diameter x_{ij} and the length l_{ij} for each segment. This introduces $O(n^2)$ product terms, each of which need to be linearized causing a blowup in the number of constraints. To overcome this constraint blowup, we ignore the two pipe segment structure of the original solution. Each link can consist of multiple pipe segments for each possible commercial pipe diameter. The length of each segment needs to be determined. No explicit choice is made regarding which commercial pipe diameter is chosen or not. The commercial pipe diameter is considered chosen if its segment has non-zero length.

We introduce continuous variables l_{ij} and do away with the binary choice variables x_{ij1} and x_{ij2}. Each link is made up of NP components corresponding to the NP pipe diameters. l_{ij} then represents the length of each of these components.

4.3.1 The Objective Function

We replace the terms $l_i * x_{ij}$ with l_{ij} in our objective cost:

$$O(.) = \sum_{i=1}^{NL} \sum_{j=1}^{NP} C_{ij}(D_{ij}) l_{ij} \ . \tag{18}$$

4.3.2 Pipe Constraint

Now the pipe constraint simply reduces to:

$$\sum_{j=1}^{NP} l_{ij} = L_i \ , \qquad i = 1, \ \ldots, \ NL \ . \tag{19}$$

4.3.3 Node Constraint

In the node constraint we replace $x_{ij} * L_i$ with l_{ij}:

$$P_n \leq H_R - E_n - \sum_{i \in S_n} \sum_{j=1}^{NP} HL_{ij} l_{ij} \ , \qquad n = 1, \ \ldots, \ NN \ . \tag{20}$$

4.3.4 Computational Results

As in the case of OnePipe the number of constraints is linear in the number of nodes and the performance turns out to be much better. Table 4.1 shows the size of model in terms of variables and constraints for the two approaches.

Table 4.1: Size comparison of TwoPipe and General models.

	TwoPipe Model		**General Model**	
Nodes	**Constraints**	**Variables**	**Constraints**	**Variables**
n	6n + 8nm*	2n + 4nm	2n	nm
10	860	420	20	200
100	8600	4200	200	2000

*m is the number of commercial pipe diameters which is taken as 10.

For a 100 node network we have around 200 inequalities for the general formulation. But for the two-pipe model, a network with just 10 nodes the number of constraints is around 800! Also by eliminating binary variables and replacing them with continuous variables we have converted what is an Integer Linear Program (ILP) to a Linear Program (LP). While ILPs are NP hard [52], LPs can be solved in Polynomial time [39]. This is very significant since we are able to solve the LP formulation of a 1000 node network in two seconds even though it contains 2000 constraints.

We now have a more general formulation for our pipe diameter optimization problem. We have the added benefit of not restricting the solution to two pipes per link. If the necessary convexity conditions [24] are met, our solution should naturally contain only two pipes per link with adjacent diameters.

We evaluate the performance of the general approach in comparison to the OnePipe model and the popular design software BRANCH. Table 4.2 presents the performance in terms of objective cost as well as running time of the three methods over six test networks. For the Gen100 network, BRANCH terminated with a memory overflow message. Since the stated maximum number of nodes is 125, the Gen1000 network is not run on BRANCH.

The time taken by both the OnePipe model and the general model is less than half a second for the first four networks. But the ILP vs. LP nature is borne out for the two larger networks. In fact, for the Gen1000 the OnePipe model also gets timed out.

Being optimal, the general model indeed outperforms both BRANCH and the OnePipe model for all six networks in terms of objective cost. Both BRANCH and our proposed model use a more general formulation and hence better cost results are to be expected. The general model performs better than BRANCH since we use a LP formulation that is solved optimally whereas BRANCH uses a heuristic approach. In fact, for two of the first four networks even the OnePipe model performs better than BRANCH.

Table 4.2: Comparison of the General model to the OnePipe model and BRANCH.

		BRANCH		**OnePipe Model**		**General Model**	
Network	**Number of Nodes**	**Cost (1000 Rs)**	**Time (msec)**	**Cost (1000 Rs)**	**Time (msec)**	**Cost (1000 Rs)**	**Time (msec)**
Mokhada	37	24,652	119380	24,317	429	24,181	387
Shahpur	21	29,082	34878	29,523	356	28,895	318
Khardi	11	21,281	4207	21,267	256	21,184	247
Gen5	5	1958	1105	2020	243	1949	228
Gen100	100	-	-	93,718	1531	90,512	628
Gen1000	1000	-	-	5,80,578	Timeout	5,62,564	2023

4.4 Parallel Pipes

A common design challenge is the augmentation of existing networks. The previous infrastructure is frequently of not sufficient capacity since it was designed for a smaller population. As part of the previous model, a user could provide diameters for existing pipes in the network. Therefore, some links in the model could have fixed diameters. It is possible for this network configuration, that all constraints cannot be met, especially if the fixed diameters provided are very small.

It is therefore desirable to allow pipes to be placed in parallel to existing pipes to augment capacity if and when required. To capture this requirement we introduce binary variables p_{ij} and p_i for each link i and commercial pipe diameter j. p_{ij} represents the choice of diameter for the parallel pipe for link i. No parallel pipe being chosen is represented by the variable p_i. Note that since we are introducing binary variables our program now once again reverts to an Integer Linear Program. But in this particular case the performance is not affected significantly since these variables are introduced only in the case of links with existing pipes, and then too only if user permits the addition of parallel pipes to those links. These are typically going to be a very small number as compared to the total links in the network. The modified objective function and constraints are given below.

4.4.1 The Objective Function

Let S_E be the set of links which have existing pipes and $S_{\sim E}$ the set of remaining links. Note that existing links don't contribute to the cost. Then the modified objective function is:

$$O(.) = \sum_{i \in S_{\sim E}}^{NL} \sum_{j=1}^{NP} C_{ij}(D_{ij}) l_{ij} + \sum_{i \in S_E}^{NL} \sum_{j=1}^{NP} L_i C_{ij}(D_{ij}) p_{ij} \; . \tag{21}$$

4.4.2 Pipe Constraint

For all links that have existing pipes we add an additional pipe constraint:

$$p_i + \sum_{j=1}^{NP} p_{ij} = 1 \; , \qquad i = 1, \; \ldots, \; NL \; . \tag{22}$$

4.4.3 Node Constraint

For links with existing pipes, headloss changes since the flow in those pipes now depends on the parallel pipe chosen:

$$P_n \leq H_R - E_n - \sum_{i \in S_n} \sum_{j=1}^{NP} HL_{ij} l'_{ij} \; , \qquad n = 1, \; \ldots, \; NN \; , \tag{23}$$

$$HL_{ij} = \frac{10.68 * \left(\frac{FL_{ij}}{R_j}\right)^{1.852}}{D_j^{4.87}} \; , \qquad i = 1, \; \ldots, \; NL, \quad j = 1, \; \ldots, \; NP \; , \tag{24}$$

$$l'_{ij} = l_{ij} \; , \qquad i \in S_{\sim E} \; , \tag{25}$$

$$l'_{ij} = L_i * p_{ij} \; , \qquad i \in S_E \; , \tag{26}$$

$$FL_{ij} = FL_i \ , \qquad i \in S_{\sim E} \ , \quad (27)$$

$$FL_{ij} = FL_i \left(1 + \frac{R_j}{R_i} * \left(\frac{D_j}{D_i} \right)^{\frac{4.87}{1.852}} \right)^{-1} \ , \qquad i \in S_E \ . \quad (28)$$

Where R_i and D_i are the roughness and pipe diameter values of the existing pipe.

In this chapter, we discussed the core problem of network design, the selection of pipe diameters from a discrete set of commercially available pipe diameters and looked at several approaches to solve the problem. Software in use by government engineers today, restrict themselves to this problem. But networks consist of several other components as well which impact the cost of the network and level of service provided. In the next chapter we discuss tanks/ESRs, why they are important to the cost of the network and how we integrated them into our general model.

Chapter 5

ESR Sizing and Allocation

5.1 The ESR configuration problem

The ESR/Tank configuration problem is to determine the ESR locations, heights, capacities, and the downstream demand nodes that each ESR will service. The purpose of using ESRs is to divide the network into a primary network and secondary networks. Primary network distributes the water from the source to the ESRs. Each ESR then distributes water to the demand nodes it is responsible for. This division of responsibility helps in providing a more equitable distribution of water in the entire network. Typically this choice of ESR configuration is made in an ad-hoc manner, relying on the intuition and experience of the designer. This choice is integrated in our capital cost optimization formulation and the same is implemented in JalTantra.

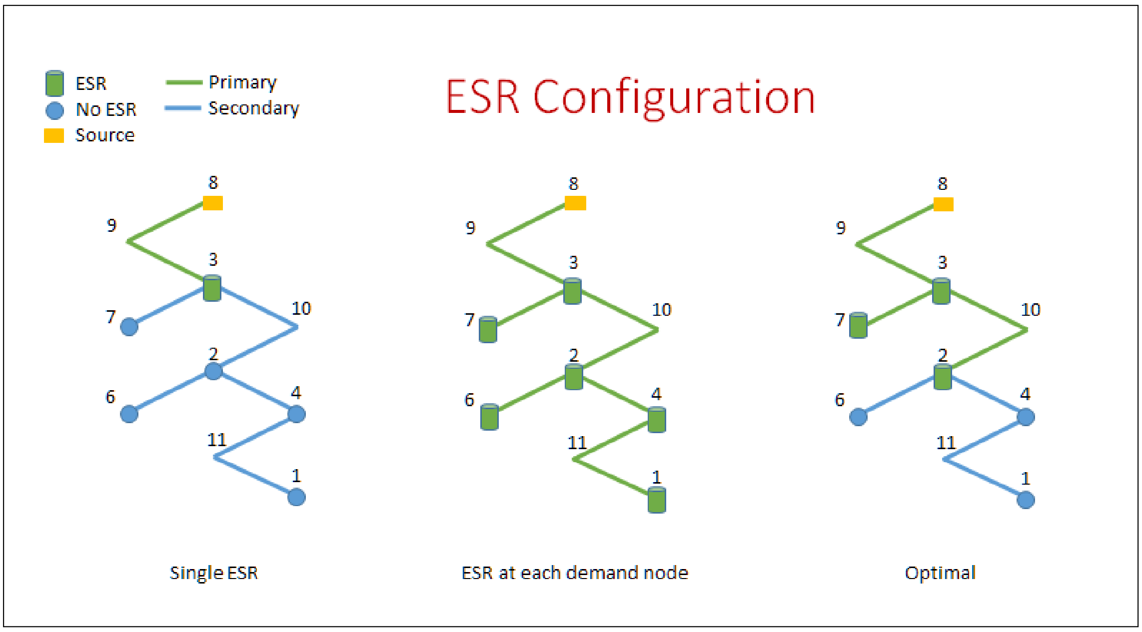


Figure 5.1: Alternate ESR configurations for a sample network.

ESR allocation in a network can be done in several ways. The choice can be an ESR for each demand node or a single ESR for the entire network or any other configuration in between these two extremes. This allocation then determines the ESR capacity. Figure 5.1 depicts the two extreme configurations as well as the "optimal" one for a sample network with six demand nodes. Note how the choice affects the primary/secondary networks.

5.2 ESR Capital Cost

In Chapter 4 we discussed the capital cost of pipes which is the first source of network capital cost. The second component of network cost is the cost of the ESRs.The capital cost for ESRs depends on the size of the ESRs to be built. However note that the cost of a ESR rises sub-linearly. That is, doubling the ESR capacity changes the cost to less than double the original cost. The non-linear ESR cost is a piece-wise linear function, represented by a table which divides the ESR capacity into several brackets. A typical ESR cost table from [49] is shown in table 5.1. An ESR with a capacity of 55000 litres will therefore cost 935800 + 9.64 * (55000 - 50000) = 984000 Rs.

Table 5.1: Typical ESR Cost Table.

Minimum Capacity (l)	Maximum Capacity (l)	Base Cost (Rs)	Unit Cost (Rs)
0	25000	0	24.47
25000	50000	611800	12.96
50000	75000	935800	9.64
75000	100000	1178800	8.64
100000	150000	1394800	7.23
150000	200000	1772800	6.03
200000	250000	2096800	5.4
250000	300000	2366800	5.4
300000	400000	2636800	5.12
400000	500000	3176800	4.32
500000	750000	3608800	4.32
750000	1000000	4688000	4.32
1000000	1500000	5768800	4.32
1500000	2000000	7928800	3.92
2000000	-	9548800	3.24

5.3 The Push and Pull of Pipes and ESRs on the Total Capital Cost

The distribution of headloss in the network dictates if the node pressure requirements are being satisfied or not. Headloss in a pipe depends on the length and diameter of the pipe used, as well as the flow through the pipe. For the branched networks, the flow in a pipe depends on whether the pipe is part of a primary network, or a secondary network, which in turn depends on the choice of ESR configurations. Typically the primary network runs for the entire day whereas secondary networks are scheduled to run for a few hours every day in order to manage the distribution of water. Thus flow rate in a secondary network is higher than that in a primary network. Therefore, for the same headloss across a pipe, higher diameters are required in case of a secondary network. This means that the total pipe cost is minimized when the entire network is a primary network, that is, there is a ESR installed at each demand node, and there are no secondary networks (for e.g. in the second configuration shown in Figure 5.1).

The total ESR capacity required for the network is same regardless of the ESR configuration, that is, the number, locations, and the allocation of demand nodes to the ESRs. The cost for various configurations, however, would be different, since as mentioned earlier, individual ESR cost rises sublinearly with its capacity. Therefore the total ESR cost is minimized when a single ESR serves the entire network (for e.g. in the first configuration shown in Figure 5.1).

For the "ESR at each demand node" configuration, the pipe cost is minimum but the ESR cost is maximum compared to any alternative configuration. In the case of a single ESR, the ESR cost is minimum but the pipe cost is maximum. The cost optimum ESR configuration therefore depends on the network topology and can lie anywhere at or between these two extremes. For example in the sample network shown in Figure 5.1, the capital cost is minimized if 3 ESRs are built.

In summary, the choice of ESR configuration, that is, the location, height and capacity of the ESRs, and the set of demand nodes that each ESR serves, is a non-trivial decision that has a direct impact on the capital cost optimization of piped water networks.

5.4 Model Details

Several variables are added to the model to implement ESR allocation. We briefly describe the added variables and constraints below. Note that these variables and constraints are in addition to those described in our previous model.

5.4.1 Variables

Continuous Variables:

t_n: Height of ESR at node n

d_n: Total demand served by ESR at node n

h_n: Water head at node n

hl_i: Total headloss across link i

h'_{ni}: Effective head provided to link i by its starting node n

Binary Variables:

f_i: 1 if flow in link i is primary, 0 if secondary

es_{ni}: 1 if source of water for link i is its immediate upstream node n, 0 otherwise

s_{nm}: 1 if ESR at node n is source for node m, 0 otherwise

e_{nk}: 1 if ESR at n^{th} node is costed by the k^{th} row of ESR cost table, 0 otherwise

Parameters:

NN: Number of nodes in the network

NE: Number of rows in the ESR cost table

B_k: Base cost of the k^{th} row of the ESR cost table

UN_k: Unit cost of the k^{th} row of the ESR cost table

UP_k: Upper limit capacity for the k^{th} row of the ESR cost table

LO_k: Lower limit capacity for the k^{th} row of the ESR cost table

DE: The total water demand of the network

T_{min}: Minimum ESR height allowed

T_{max}: Maximum ESR height allowed

A_n: Set of nodes that are ancestors of node n

D_n: Set of nodes that are descendants of node n

C_n: Set of child nodes of node n

P_n: Parent node of node n

I_n: Incoming link for node n

O_n: Set of outgoing links from node n

FL_i^p: Flow in i^{th} link if i is part of the primary network

FL_i^s: Flow in i^{th} link if i is part of the secondary network

PH: Number of hours of water supply in the primary network

SH: Number of hours of water supply in the secondary network

5.4.2 Objective Cost

The additional objective cost term is simply the ESR cost at each node.

$$O(.) = \sum_{i \in S_{\sim E}}^{NL} \sum_{j=1}^{NP} C_{ij}(D_{ij}) l_{ij} + \sum_{i \in S_E}^{NL} \sum_{j=1}^{NP} L_i C_{ij}(D_{ij}) p_{ij} + \sum_{n=1}^{NN} \sum_{k=1}^{NE} e_{nk} * (B_k + UN_k * (d_n - LO_k)) \ . \tag{29}$$

5.4.3 Constraints

- Note that the ESR cost term is non-linear since it contains a product of two variables $e_{nk} * d_i$. But this term is linearizable since e_{nk} is a binary variable. We introduce z_{nk} to represent $e_{nk} * d_i$ and add the following constraints:

$$z_{nk} \leq DE * e_{nk} \ , \qquad n = 1, \ \ldots, \ NN, \quad k = 1, \ \ldots, \ NE \ , \tag{30}$$

$$z_{nk} \leq d_n \ , \qquad n = 1, \ \ldots, \ NN, \quad k = 1, \ \ldots, \ NE \ , \tag{31}$$

$$z_{nk} \geq d_n - DE * (1 - e_{nk}) \ , \qquad n = 1, \ \ldots, \ NN, \quad k = 1, \ \ldots, \ NE \ . \tag{32}$$

  Note that this will be a common occurrence in the constraints to come i.e. the decomposition of a product of a continuous and a binary variable into linear constraints. For the sake of clarity and brevity the non-linear constraints will be mentioned as is.

- The first ESR constraint is to ensure that every ESR height t_n is between constants T_{min} and T_{max}.

$$T_{min} \leq t_n \leq T_{max} \ , \qquad n = 1, \ \ldots, \ NN \ . \tag{33}$$

- Next the head constraint at each node is modified to include the ESR height.

$$P_n \leq h_n - t_n - E_n \ , \qquad n = 1, \ \ldots, \ NN \ . \tag{34}$$

- We then look at the constraints that deal with allocation of demand nodes to ESRs.If a node n does not serve its own demand i.e. it is part of a secondary network, then all its downstream nodes will also be part of a secondary network.

$$s_{mm} \leq s_{nn}, \quad n = 1, \ \ldots, \ NN, \quad m \in D_n \ . \tag{35}$$

- If a node n does not serve its own demand, then it cannot serve the demand of its downstream nodes.

$$s_{nm} \leq s_{nn}, \quad n = 1, \ \ldots, \ NN, \quad m \in D_n \ . \tag{36}$$

- For every node n, only one upstream node m can serve its demand.

$$\sum_{m \in A_n \cup \{n\}} s_{mn} = 1, \quad n = 1, \ \ldots, \ NN \ . \tag{37}$$

- The total demand d_n served by node n is the sum of the demands of the downstream nodes that it serves i.e. all m such that $s_{nm} = 1$.

$$d_n = \sum_{m \in D_n \cup \{n\}} s_{nm} * DE_m, \quad n = 1, \ \ldots, \ NN \ . \tag{38}$$

- For a node n, its incoming pipe i will have primary flow only if the node serves itself.

$$f_i = s_{nn}, \quad n = 1, \ \ldots, \ NN, \quad i = I_n \ . \tag{39}$$

- If a node n serves node m i.e. $s_{nm} = 1$, each node o in the path from n to m belongs to a secondary network and therefore cannot serve itself.

$$s_{nm} \leq 1 - s_{oo}, n = 1, \ \ldots, \ NN, m \in D_n, o \in D_n \cap A_m \ . \tag{40}$$

- Next, we have the constraints that relate the demand that a ESR serves to its cost variables e_{nk}. Note that we require z_{nk} in our objective function to replace the non linear term $e_{nk} \times d_n$:

$$z_{nk} = e_{nk} \times d_n, \quad n = 1, \ \ldots, \ NN, \quad k = 1, \ \ldots, \ NE \ . \tag{41}$$

- Since every ESR can be costed using exactly one row of the table, the sum of e_{nk} for a given n must be 1:

$$\sum_{k=1}^{NE} e_{nk} = 1, \quad n = 1, \ \ldots, \ NN \ . \tag{42}$$

- Next we have constraints that make sure that the ESR capacity d_n lies between the minimum and maximum capacity of the selected row of the cost table:

$$For \quad n = 1,\ \ldots,\ NN, \quad k = 1,\ \ldots,\ NE:$$

$$LO_k e_{nk} \leq d_n\ , \tag{43}$$

$$DE * e_{nk} + d_n \leq UP_k + DE\ . \tag{44}$$

- Across every link i there is headloss hl_i. This headloss depends on the flow, length and diameter of the pipe chosen. As before, we use the Hazen-Williams equation [74] to calculate the headloss. The flow through the link depends on whether the link is part of the primary or secondary network:

$$hl_i = \sum_{j=1}^{NP}(HL_{ij}^p f_i + HL_{ij}^s(1-f_i))l_{ij}, \qquad i = 1,\ \ldots,\ NL\ , \tag{45}$$

$$HL_{ij}^p = \frac{10.68 * \left(\frac{FL_i^p}{R_j}\right)^{1.852}}{D_j^{4.87}}, \qquad i = 1,\ \ldots,\ NL, \quad j = 1,\ \ldots,\ NP\ , \tag{46}$$

$$HL_{ij}^s = \frac{10.68 * \left(\frac{FL_i^s}{R_j}\right)^{1.852}}{D_j^{4.87}}, \qquad i = 1,\ \ldots,\ NL, \quad j = 1,\ \ldots,\ NP\ , \tag{47}$$

$$FL_i^s = FL_i^p * \frac{PH}{SH}, \qquad i = 1,\ \ldots,\ NL\ . \tag{48}$$

Before the introduction of ESRs, the source node provided head to the entire network. Therefore the head at each node was computed as the head provided by the source minus the sum of all headlosses along the path from the source to the node. But now each ESR serves the role of the source to the secondary network it is responsible for. The impact of the introduction of ESR is illustrated in Figure 5.2. The source remains responsible for the primary network. Therefore for each link i with a start node n:

$$h_n = h'_{mi} - hl_i, \qquad n = 1,\ \ldots,\ NN, \quad m = P_n, \quad i = I_n, \tag{49}$$

$$h'_{mi} = (t_m + E_m) * es_{mi} + h_m * (1 - es_{mi}), \qquad m = 1,\ \ldots,\ NN, \quad i \in O_m, \tag{50}$$

$$es_{mi} = s_{mm} * (1 - f_i), \qquad m = 1,\ \ldots,\ NN, \quad i \in O_m. \tag{51}$$

es_{mi} here represents if the secondary source of pipe i is node m. It is 1 only if node m serves itself and if the flow in pipe i is secondary. If es_{mi} is 1 then the effective head served by node m is the sum of its elevation and the esr height. Else it is simply the head provided by the upstream source.

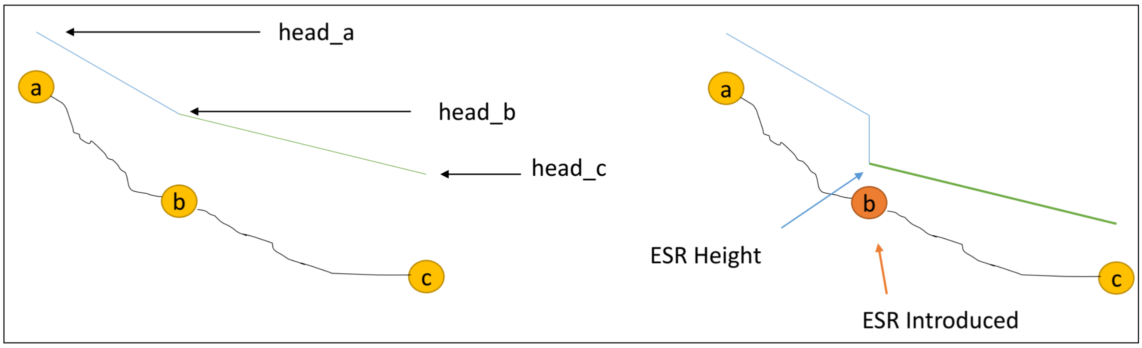


Figure 5.2: Change in downstream head due to introduction of ESR.

5.4.4 Computational Results

The new model was tested on the sample network shown in 5.1. Apart from the optimal configuration we looked at the cost breakup of the two extreme configurations, namely a single ESR and ESRs at each demand point.

Table 5.2: Cost breakup of alternative ESR configurations.

Configurations	**Number of ESRs**	**ESR Cost (1000 Rs)**	**Piping Cost (1000 Rs)**	**Total Cost (1000 Rs)**
Single ESR	1	**3703**	20214	23917
ESR at every demand node	6	7644	**14642**	22286
Optimal	3	5694	16041	**21735**

The results shown in table 5.2 are in line with expectations. The single ESR configuration has the minimum ESR cost and the ESR at every node configuration has the minimum piping cost. The overall optimal configuration however has both ESR and piping cost in the middle but an overall lower cost.

We have looked at the two primary contributors to capital cost in a network, pipes and ESRs. We next look at pumps which not only contribute to the capital cost of the scheme but also the operational cost. Additionally with the introduction of pumps and valves, the designer gains the ability to increase/decrease the head in the network.

Chapter 6

Pumps/Valves Integration

So far we have only looked at networks that have solely relied on gravity for the transmission of water. But in most real life scenarios it is often the case that part of the distribution network is at a higher elevation than the source. It might be possible to increase the source head by constructing the MBR at a height but then this might put a heavy load on the capital cost of the scheme. To ensure the high elevation locations receive the water, one might be forced to employ pipes of large diameters that are very expensive.

On the flip side there might be networks where the source is at a very high elevation and therefore there is excessive pressure in the entire network despite using the smallest pipe diameters. The problem with excess pressure is that it might lead to bursting of pipes and as such higher resilience pipes may need to be employed which again increases the capital cost.

To address the problems of too less or too much head, network components like pumps and valves can be employed. Pumps help provide additional head to the network. Up to now we have only considered the capital cost of the scheme. But with pumps an important component of its cost, if not the most, is its operational cost. The energy required to run the pump is a continuous cost that the scheme must bear. So with the introduction of pumps, we must now consider the operational cost in addition to the capital cost of the scheme.

6.1 Model Extension to Include Pumps

To include pumps in our model, we consider a pump being allowed at each pipe in the network (unless user specifically says otherwise). What remains to be determined is the power of each pump and the head provided by the pump. If the pump power selected is 0, it means that a pump is not selected for that pipe. Also the pumps are assumed to be operational for the entire duration of the supply. Therefore a pump in the primary/secondary network will run for the primary/secondary supply hours respectively. As before the following objective cost terms, constraints and variables are in addition to the ones already mentioned earlier.

6.1.1 Variables

Continuous Variables:

p_i: Power of pump installed at link i

p_i^p: Power of pump installed at link i, if link i is part of primary network

p_i^s: Power of pump installed at link i, if link i is part of secondary network

ph_i: Head provided by pump at link i

Binary Variables:

pe_i: 1 if a pump is installed at link i, 0 otherwise

Parameters:

ρ: Density of water

g: Acceleration due to gravity

η: Efficiency of pump

PP_{min}: Minimum pump power allowed

PP_{max}: Maximum pump power allowed

$INFR$: Inflation rate

$INTR$: Interest rate

CP: Capital cost of pumps per unit kW

EP: Energy cost of pumps per unit kWh

DF: Discount factor for the energy cost over the entire scheme lifetime

6.1.2 Objective Function

Now that operational cost is also being considered, there are several ways to incorporate it along with the capital cost. One possibility is optimizing for the operational cost while considering capital cost as a constraint. This would be in line with schemes having a fixed upper limit on their budget depending on the size of the population that they serve. Another possibility is to consider both capital cost and operational cost in the objective function. Here since the operational cost is borne across several years, the effective operational cost used is after considering both the interest rate and the inflation rate.

$$
\begin{aligned}
O(.) = \quad & \sum_{i \in S_{\sim E}}^{NL} \sum_{j=1}^{NP} C_{ij}(D_{ij}) l_{ij} \quad + \quad \sum_{i \in S_E}^{NL} \sum_{j=1}^{NP} L_i C_{ij}(D_{ij}) p_{ij} \\
+ \quad & \sum_{n=1}^{NN} \sum_{k=1}^{NE} e_{nk} * (B_k + UN_k * (d_n - LO_k)) \\
+ \quad & \sum_{i=1}^{NL} CP * p_i \quad + \quad EP * DF * \left(\sum_{i=1}^{NL} PH * p_i^p + \sum_{i=1}^{NL} SH * p_i^s \right) \ ,
\end{aligned}
\tag{52}
$$

$$
where \quad DF = \quad \sum_{n=1}^{Y} \left(\frac{1 + INFR}{1 + INTR} \right)^{n-1} \ .
$$

6.1.3 Constraints

- The original headloss constraint is modified to include the artificial head provided by the pump:

$$
hl_i = \sum_{j=1}^{NP} (HL_{ij}^p f_i + HL_{ij}^s (1 - f_i)) l_{ij} \quad - \quad p_i, \quad i = 1, \ \ldots, \ NL \ . \tag{53}
$$

- Next we have to relate the head provided by the pump to its power. Head ph_i provided by the pump is a continuous variable and the flow can take on two values depending on whether the pipe is in the primary network or the secondary network. Therefore we compute power as the sum of primary and secondary power and compute each component separately.

$$
\begin{aligned}
p_i &= p_i^p + p_i^s, && i = 1, \ \ldots, \ NL, && (54) \\
p_i^p &= \frac{(\rho * g * FL_i^p * ph_i)}{\eta} * f_i, && i = 1, \ \ldots, \ NL, && (55) \\
p_i^s &= \frac{(\rho * g * FL_i^s * ph_i)}{\eta} * (1 - f_i), && i = 1, \ \ldots, \ NL \ . && (56)
\end{aligned}
$$

- Finally, the pump power for each pump must lie between minimum and maximum allowed pump power. This is implemented using the binary variable pe_i.

$$PP_{min} * pe_i \leq p_i \leq PP_{max} * pe_i, \quad i = 1,\ \ldots,\ NL\ . \tag{57}$$

Table 6.1: Cost breakup of alternative ESR configurations with pumps.

Configurations	**Number of ESRs**	**ESR Cost (1000 Rs)**	**Piping Cost (1000 Rs)**	**Pump Cost (1000 Rs)**	**Total Cost (1000 Rs)**
Single ESR	1	**3703**	20214	0	23917
ESR at every demand node	6	7644	14642	0	22286
Optimal	3	5694	16041	0	21735
Optimal with Pumps	2	4856	**13628**	2132	**20615**

6.1.4 Computational Results

As in section 5.4.4, the model including pumps was tested on the sample network shown in figure 5.1. Along with the different ESR configurations, we now compare the cost once pumps are allowed as well. As can be seen in table 6.1, the inclusion of pumps has resulted in a further decrease in overall cost. This is due to a significant decrease in piping cost since pipes of lesser diameters could be utilized.

6.2 Model Extension to Include Valves

The introduction of pressure reducing valves reduces the downstream head along the pipe they are installed, hence reducing excess pressure in the network. But from the optimization perspective if no penalties or constraints are included to "punish" excess pressure, the valves will never be chosen by the model. Currently the notion of excess pressure being a negative is purely an external to the model notion. Based on discussions with government engineers, we have decided to not strictly enforce any maximum pressure constraints since in some cases (like hilly areas) it might be unavoidable. Thus, to incorporate pressure reducing valves we have included a manual option where the user can introduce for any link i a valve with pressure reducing setting VH_i. The headloss equation therefore becomes:

$$hl_i = \sum_{j=1}^{NP} (HL_{ij}^{p} f_i + HL_{ij}^{s}(1 - f_i)) l_{ij} \quad - \quad p_i \quad + \quad VH_i, \quad i = 1,\ \ldots,\ NL\ . \tag{58}$$

With the inclusion of ESRs, pumps and valves the complexity of the model increased significantly leading to poor performance on larger networks. In order to address this, the a series of improvements to the model were made which are described in the next chapter.

Chapter 7

Model Improvements

The first version of the model included just the pipe diameter optimization for branched networks (typical in the case of rural areas). It was a LP model and thus solved the problem quickly and optimally. This allowed even networks of a thousand nodes to be solved in a couple of seconds. We extended the model to include ESRs. The added complexity of considering both primary and secondary networks simultaneously, required an ILP model. Although still optimal in terms of cost, the time taken was significantly worse. In this chapter we describe three significant improvements that were made to the model. These improvements reduced the time taken to optimize the larger networks by orders of magnitude. The time taken to optimize a 150 node network has gone from over 40 minutes to 5 seconds, and a 200 node network which could not be solved within 24 hours now takes just 70 seconds.

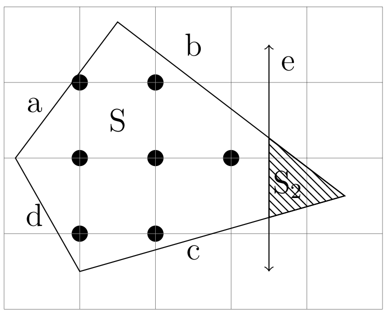


Figure 7.1: Tightening linear relaxation by introducing constraint.

The improvements consist of tightening the set of constraints used to describe the ILP model. Consider the example shown in figure 7.1. The points represent the integer points over which we are trying to optimize. The lines a, b, c, d and e represent the constraints that encompass those integer points. When solving the linear point (LP) relaxation, the entire set S is considered. By introducing the constraint e, we can still capture the same integer points while cutting off a part (S_2) of the linear relaxation. Since a smaller solution space is now considered while solving the LP relaxation, this speeds up the optimization. For each of the three improvements presented, we prove that the newer set of constraints have a linear relaxation that is a strict subset of the linear relaxation of the older set, while maintaining the same set of integer points. In particular, for the ESR configuration improvement we show that the newer subset of constraints is as tight as possible, i.e. the linear relaxation has no fractional points. Since the overall model is complex, while discussing each improvement, we only consider a small subset of relevant constraints at a time.

7.1 Pipe Headloss Improvement

7.1.1 Initial Model

We first focus on a part of the model whose purpose is to determine the pipe diameters chosen for each link in the network. Each link can consist of multiple pipe diameters. Also, each link can be part of the primary network or the secondary network. The headloss across the link depends on these choice of pipe diameters and whether it belongs to the primary or secondary network. The set of variables and parameters used for this purpose are defined as follows. Consider a network of NL links. Let NP be the number of pipe diameters available.

Variables:

l_{ij} = length of the j^{th} pipe diameter component of link i, $\quad i = 1, \ldots, NL, \quad j = 1, \ldots, NP$,

l^p_{ij} = length of the j^{th} pipe diameter component of link i, if link i is part of the primary network, $i = 1, \ldots, NL$, $j = 1, \ldots, NP$,

hl_i = headloss across link i, $\quad i = 1, \ldots, NL$,

f_i = 1 of link i is part of the primary network, 0 if it is part of the secondary network, $\quad i = 1, \ldots, NL$.

Parameters:

L_i = Length of link i, $\quad i = 1, \ldots, NL$,

HL^p_{ij} = Unit headloss for the j^{th} pipe diameter component of link i, if i is part of primary network, $\quad i = 1, \ldots, NL, \quad j = 1, \ldots, NP$,

HL^s_{ij} = Unit headloss for the j^{th} pipe diameter component of link i, if i is part of secondary network, $\quad i = 1, \ldots, NL, \quad j = 1, \ldots, NP$.

Constraints:

The first constraint captures l_{ij}^p as a product of l_{ij} and f_i:

$$l_{ij}^p = l_{ij} \times f_i, \quad i = 1,\ \ldots,\ NL, \quad j = 1,\ \ldots,\ NP. \tag{59}$$

Equation (59) consists of a product of two variables, and is therefore a non linear equation. Fortunately since f_i is a binary variable, we can linearize the equation using the following inequalities:

$$\begin{aligned}
0 \leq l_{ij}^p,&\ i = 1,\ \ldots,\ NL,\ j = 1,\ \ldots,\ NP, && (60)\\
l_{ij}^p \leq L_i f_i,&\ i = 1,\ \ldots,\ NL,\ j = 1,\ \ldots,\ NP, && (61)\\
l_{ij} - L_i(1 - f_i) \leq l_{ij}^p,&\ i = 1,\ \ldots,\ NL,\ j = 1,\ \ldots,\ NP, && (62)\\
l_{ij}^p \leq l_{ij},&\ i = 1,\ \ldots,\ NL,\ j = 1,\ \ldots,\ NP. && (63)
\end{aligned}$$

The sum of all pipe diameter components must equal the link length:

$$\sum_{j=1}^{NP} l_{ij} = L_i, \quad i = 1,\ \ldots,\ NL. \tag{64}$$

Next we have the equation for hl_i, which is the sum all headloss components contributed by the different pipe diameter components of link i :

$$hl_i = \sum_{j=1}^{NP} P_{ij} l_{ij}^p + \sum_{j=1}^{NP} S_{ij}(l_{ij} - l_{ij}^p), \quad i = 1,\ \ldots,\ NL. \tag{65}$$

Finally we have constraints that relate to the bounds for the variables:

$$\begin{aligned}
l_{ij} \geq 0,& \quad i = 1,\ \ldots,\ NL, \quad j = 1,\ \ldots,\ NP, && (66)\\
f_i \in \{0, 1\},& \quad i = 1,\ \ldots,\ NL. && (67)
\end{aligned}$$

Since there exists a l_{ij}^p for each link and pipe diameter combination in the network, a large number of linear decompositions of equation (59) need to be done. In the next section we show an improved model that has the same feasible 0-1 set of values but with a tighter LP relaxation, resulting in better performance.

7.1.2 Improved Model (Model-2)

In order to decompose the product of variables in (59), a large number of constraints need to be added. This is avoided in the new model by not explicitly defining l_{ij}^p. Instead its relation to l_{ij} and f_i is implicit. In the next section we show that new model is better.

Variables:

We introduce one new variable, which is similar to l^p_{ij} but for the secondary network:
l^s_{ij} = length of the j^{th} pipe diameter component of link i, if link i is part of the secondary network, and 0 if link i is part of the primary network $\quad i = 1, \ldots, NL, \quad j = 1, \ldots, NP.$

Constraints:

The first constraint simply states that l_{ij} is the sum of the primary and secondary components, i.e. l^p_{ij} and l^s_{ij} respectively:

$$l_{ij} = l^p_{ij} + l^s_{ij}, \quad i = 1, \ldots, NL, \quad j = 1, \ldots, NP. \tag{68}$$

For a given link i, either all l^p_{ij} are 0 or all l^s_{ij} are 0, depending on the value of f_i. And the sum of the non-zero components must equal the length of the link L_i. The first two inequalities of the new model capture this:

$$\sum_{j=1}^{NP} l^p_{ij} = L_i f_i, \quad i = 1, \ldots, NL, \tag{69}$$

$$\sum_{j=1}^{NP} l^s_{ij} = L_i(1 - f_i), \quad i = 1, \ldots, NL. \tag{70}$$

Next we have the equation for hl_i, which is the sum all headloss components contributed by the different pipe diameter components of link i. For the new model we equivalently use l^s_{ij} instead of $l_{ij} - l^p_{ij}$ due to equation (68):

$$hl_i = \sum_{j=1}^{NP} P_{ij} l^p_{ij} + \sum_{j=1}^{NP} S_{ij} l^s_{ij}, \quad i = 1, \ldots, NL. \tag{71}$$

Finally as before we have the bounds for the variables:

$$l_{ij} \geq 0, \quad i = 1, \ldots, NL, \quad j = 1, \ldots, NP, \tag{66}$$

$$l^p_{ij} \geq 0, \quad i = 1, \ldots, NL, \quad j = 1, \ldots, NP, \tag{60}$$

$$l^s_{ij} \geq 0, \quad i = 1, \ldots, NL, \quad j = 1, \ldots, NP, \tag{72}$$

$$f_i \in \{0, 1\}, \quad i = 1, \ldots, NL. \tag{67}$$

We now prove that the improved model is tighter than the initial model, that is the linear relaxation of the improved model is a strict subset of the linear relaxation of the initial model. Let S_1 be the set of points belonging to the initial model and S_2 be the set of points belonging to the improved model. Let R_1 and R_2 be the set of points corresponding to the LP relaxations of S_1 and S_2 respectively. Both R_1 and R_2 are defined by the same set of constraints that describe the initial sets S_1 and S_2, except for the constraint (67) which refers to the binary nature of f_i. Instead, the continuous bounds for f_i is defined as follows:

$$0 \leq f_i \leq 1, \quad i = 1, \ \ldots, \ NL. \tag{73}$$

Proposition 1. *R_2 is a strict subset of R_1 i.e. $R_2 \subset R_1$.*

We prove R_2 is a strict subset of R_1 in two steps. First we show that R_2 is a subset of R_1 and then we find a point in R_1 that is not in R_2.

Consider a point $P \in R_2$. It satisfies the constraints (60), (66) and (68) - (73). We prove that it also lies in R_1 by showing that it satisfies the constraints (60)-(66) and (73). Constraints (60), (66) and (73) are trivially satisfied since they are common for both sets.

$$For \quad i = 1, \ \ldots, \ NL, \quad j = 1, \ \ldots, \ NP$$

$$\text{Proving (61)}: \quad l^p_{ij} \leq L_i f_i$$

$$\sum_{j=1}^{NP} l^p_{ij} = L_i f_i \tag{69}$$

$$\equiv \quad \{\text{using } l^p_{ij} \geq 0 \text{ (60)}\}$$

$$l^p_{ij} \leq L_i f_i$$

Hence satisfied.

$$\text{Proving (62)}: \quad l_{ij} - L_i(1 - f_i) \leq l^p_{ij}$$

$$\sum_{j=1}^{NP} l^s_{ij} = L_i(1 - f_i) \tag{70}$$

$$\Rightarrow \quad \{\text{using } l^s_{ij} \geq 0 \text{ (72)}\}$$

$$l^s_{ij} \leq L_i(1 - f_i)$$

$$\equiv \quad \{\text{using } l_{ij} = l^p_{ij} + l^s_{ij} \text{ (68)}\}$$

$$l_{ij} - l^p_{ij} \leq L_i(1 - f_i)$$

$$\equiv \quad \{\text{rearranging}\}$$

$$l_{ij} - L_i(1 - f_i) \leq l^p_{ij}$$

Hence satisfied.

$$\text{Proving (63)}: \quad l^p_{ij} \leq l_{ij}$$

$$0 \leq l^s_{ij} \tag{72}$$

$$\equiv \quad \{\text{using } l_{ij} = l^p_{ij} + l^s_{ij} \text{ (68)}\}$$

$0 \leq l_{ij} - l^p_{ij}$

$\equiv$ {rearranging}

$l^p_{ij} \leq l_{ij}$

Hence satisfied.

Proving (64) : $\sum_{j=1}^{NP} l_{ij} = L_i$

$$\sum_{j=1}^{NP} l^p_{ij} = L_i f_i \tag{69}$$

$$\sum_{j=1}^{NP} l^s_{ij} = L_i(1 - f_i) \tag{70}$$

$\equiv$ {adding equations}

$$\sum_{j=1}^{NP} (l^p_{ij} + l^s_{ij}) = L_i f_i + L_i(1 - f_i)$$

$\equiv$ {using $l_{ij} = l^p_{ij} + l^s_{ij}$ (68) and simplifying}

$$\sum_{j=1}^{NP} l_{ij} = L_i$$

Hence satisfied.

Proving (65) : $hl_i = \sum_{j=1}^{NP} P_{ij} l^p_{ij} + \sum_{j=1}^{NE} S_{ij}(l_{ij} - l^s_{ij})$

$$hl_i = \sum_{j=1}^{NP} P_{ij} l^p_{ij} + \sum_{j=1}^{NP} S_{ij} l^s_{ij} \tag{71}$$

$\equiv$ {using $l_{ij} = l^p_{ij} + l^s_{ij}$ (68)}

$$hl_i = \sum_{j=1}^{NP} P_{ij} l^p_{ij} + \sum_{j=1}^{NE} S_{ij}(l_{ij} - l^s_{ij})$$

Hence satisfied.

Therefore point $P \in R_1$, since it satisfies the constraints (60)-(66) and (73). Therefore $R_2 \subseteq R_1$.

Next we find a point Q such that $Q \in R_1$ and $Q \notin R_2$. Take point $Q(l, l^p, l^s, hl, f) = ([L/2, L/2], [L/2, L/2], [0, 0], L, 1/2)$. Here $(n, m) = (1, 2)$ and $(L, P, S) = (L, [1, 1], [1, 1])$

where $L \geq 0$. We show $Q \in R_1$ since it satisfies all the constraints.

60 :

$0 \leq l^p_{11}$, replacing values we get $0 \leq L/2$

$0 \leq l^p_{12}$, replacing values we get $0 \leq L/2$

61 :

$l^p_{11} \leq L_1 f_1$, replacing values we get $L/2 \leq L/2$

$l^p_{12} \leq L_1 f_1$, replacing values we get $L/2 \leq L/2$

62 :

$l_{11} - L_1(1 - f_1) \leq l^p_{11}$, replacing values we get

$L/2 - L(1 - 1/2) \leq L/2$

$l_{12} - L_1(1 - f_1) \leq l^p_{12}$, replacing values we get

$L/2 - L(1 - 1/2) \leq L/2$

63 :

$l^p_{11} \leq l_{11}$, replacing values we get $L/2 \leq L/2$

$l^p_{12} \leq l_{12}$, replacing values we get $L/2 \leq L/2$

64 :

$l_{11} + l_{12} = L_1$, replacing values we get $L/2 + L/2 = L$

65 :

$hl_1 = P_{11}l^p_{11} + P_{12}l^p_{12} + S_{11}(l_{11} - l^p_{11}) + S_{12}(l_{12} - l^p_{12})$,

replacing values we get $L = L/2 + L/2 + 0 + 0$

66 :

$l_{11} \geq 0$, replacing values we get $L/2 \geq 0$

$l_{12} \geq 0$, replacing values we get $L/2 \geq 0$

73 :

$0 \leq f_1 \leq 1$, replacing values we get $0 \leq 1/2 \leq 1/2$

Therefore point $Q \in R_1$. To show that $Q \notin R_2$ consider equation 70:

$$l^s_{11} + l^s_{12} = L_1(1 - f_1), \text{ replacing values we get } 0 + 0 = L/2$$

Therefore $Q \notin R_2$. We showed that $R_2 \subseteq R_1$ and that $R_2 \neq R_1$. These two together imply $R_2 \subset R_1$. Proposition 1 therefore shows that the LP relaxation of the new model (R_2) has a tighter bound than the LP relaxation of the old model (R_1).

7.2 ESR Cost Improvement

7.2.1 Initial Model

We next focus on a part of the model whose purpose is to determine the capital cost of each ESR in the network. The ESR cost is a piece-wise linear function. An example is shown in figure 7.2 below. We need to determine which row in the ESR cost table the ESR capacity falls in. Each row in the ESR cost table has minimum and maximum capacity values. If the ESR capacity is within these values, then that row is used to compute the ESR's cost. In the example below if a ESR capacity of 5 litres is to be built, then the first row of the cost table will be used, since its capacity range is 0-10. Binary variables are used to capture for each ESR, the row in the cost table that is chosen to compute the cost. The set of variables and parameters used for this purpose are defined as follows. Consider a network of n locations. Let m be the number of linear components of the piecewise linear cost of construction of a ESR.

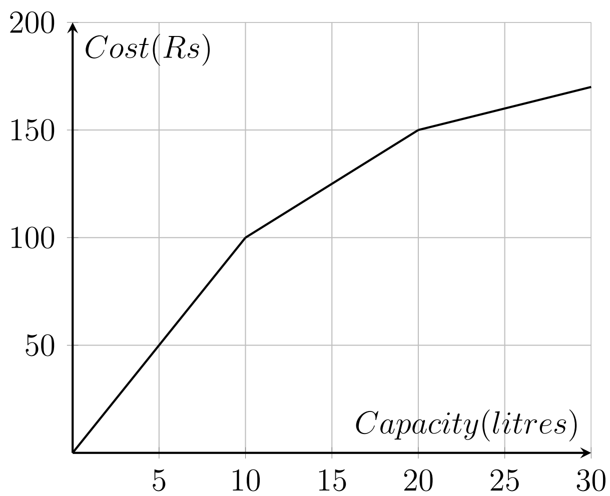


Figure 7.2: Graph of the cost of a ESR vs its capacity.

Variables:

e_{nk} = 1 if the ESR at location n is costed using the k^{th} row of the ESR cost table, $n = 1, \ldots, NN, \quad k = 1, \ldots, NE,$

z_{nk} = capacity of the ESR at location n if it is costed using the k^{th} row of the ESR cost table, 0 otherwise, $\quad n = 1, \ldots, NN, \quad k = 1, \ldots, NE,$

d_n = capacity of ESR at location n, $\quad n = 1, \ldots, NN.$

Parameters:

LO_k = minimum capacity that the k^{th} row of the ESR cost table can satisfy, $\quad k = 1, \ldots, NE,$

UP_k = maximum capacity that the k^{th} row of the ESR cost table can satisfy, $\quad k = 1, \ldots, NE,$

DE = value of the total water demand in the network, where $DE \geq UP_k, \quad k = 1, \ldots, NE.$

Constraints:

The first constraint relates the ESR capacity corresponding to the k^{th} row as a product of the ESR capacity and the binary choice variable e_{nk} :

$$z_{nk} = e_{nk} \times d_n, \quad n = 1, \ldots, NN, \quad k = 1, \ldots, NE. \tag{74}$$

Since equation (74) consists of a product of two variables, it is a non linear equation. We linearize the equation using the following inequalities:

$$0 \leq z_{nk}, \; n = 1, \ldots, NN, \quad k = 1, \ldots, NE, \tag{74.a}$$

$$z_{nk} \leq DEe_{nk}, \; n = 1, \ldots, NN, \quad k = 1, \ldots, NE, \tag{74.b}$$

$$d_n - DE(1 - e_{nk}) \leq z_{nk}, \; n = 1, \ldots, NN, \quad k = 1, \ldots, NE, \tag{74.c}$$

$$z_{nk} \leq d_n, \; n = 1, \ldots, NN, \quad k = 1, \ldots, NE. \tag{74.d}$$

Since every ESR can be costed using exactly one row, the sum of e_{nk} for a given n must be 1:

$$\sum_{k=1}^{NE} e_{nk} = 1, \quad n = 1, \ldots, NN. \tag{75}$$

Next we have constraints that make sure that the ESR capacity d_n lies between the minimum and maximum capacity of the selected row of the cost table:

$$LO_k e_{nk} \leq d_n, \, n = 1, \ldots, NN, \; k = 1, \ldots, NE, \tag{76}$$

$$DEe_{nk} + d_n \leq UP_n + DE, \, n = 1, \, \ldots, \, NN, \; k = 1, \, \ldots, \, NE. \tag{77}$$

Finally we have constraints that relate to the bounds for the variables:

$$DE \geq d_n, \quad n = 1, \, \ldots, \, NN, \tag{78}$$

$$d_n \geq 0, \quad n = 1, \, \ldots, \, NN, \tag{79}$$

$$e_{nk} \in \{0, 1\}, \quad n = 1, \, \ldots, \, NN, \quad k = 1, \, \ldots, \, NE. \tag{80}$$

Since there exists a z_{nk} for each ESR and row of cost table combination, a large number of linear decompositions of equation (74) need to be done. This results in poor performance of the model. In the next section we show an improved model that has the same feasible 0-1 set of values but with a tighter LP bound resulting in better performance.

7.2.2 Improved Model (Model-3)

As discussed in the previous section, the main issue with the old model was equation (74), where $z_n k$ is expressed as a product of two variables. In order to decompose the variables, a large number of constraints needed to be added. This is avoided in the new model by not explicitly defining z_{nk}. Instead its relation to e_{nk} and d_n is implicit. In the next section we first show that new model is better in that it has a tighter LP bound than the old model and then we go on to show that the LP for the new model has tight solutions.

The variables remain same for the new model. The first two inequalities of the model provide the bounds for z_{nk} in terms of e_{nk} and the minimum (LO_k) and maximum (UP_k) capacities for each row of the cost table:

$$LO_k e_{nk} \leq z_{nk}, \quad n = 1, \, \ldots, \, NN, \quad k = 1, \, \ldots, \, NE, \tag{81}$$

$$z_{nk} \leq UP_k e_{nk}, \quad n = 1, \, \ldots, \, NN, \quad k = 1, \, \ldots, \, NE. \tag{82}$$

The next equation for the model remains unchanged, it represents the fact that each row of the cost table is chosen exactly once for each ESR:

$$\sum_{k=1}^{NE} e_{nk} = 1, \quad n = 1, \, \ldots, \, NN. \tag{75}$$

Next, we have a similar equation but this time related to the variable z_{nk}. The sum of all z_{nk} values for a given ESR must equal d_n:

$$\sum_{k=1}^{NE} z_{nk} = d_n, \quad n = 1, \, \ldots, \, NN. \tag{83}$$

In fact along with the previous three equations of the model, one can imply that exactly one of the z_{nk} values will be non zero for a specific ESR and therefore will be equal to d_n. This therefore captures the non-linear constraint that equation (74) of the old model captured. The remaining constraints relate to the bounds for the variables:

$$DE \geq d_n, \quad n = 1, \ldots, NN, \tag{78}$$

$$e_{nk} \in \{0, 1\}, \quad n = 1, \ldots, NN, \quad k = 1, \ldots, NE, \tag{80}$$

$$d_n \geq 0, \quad n = 1, \ldots, NN, \tag{79}$$

$$z_{nk} \geq 0, \quad n = 1, \ldots, NN, \quad k = 1, \ldots, NE. \tag{74.a}$$

Let S_1 be the set of points belonging to the old model and S_2 be the set of points belonging to the new model. Let R_1 and R_2 be the set of points corresponding to the LP relaxations of S_1 and S_2 respectively.The continuous bounds for e_{nk} is defined as follows:

$$0 \leq e_{nk} \leq 1, \quad n = 1, \ldots, NN, \quad k = 1, \ldots, NE. \tag{84}$$

As in Section 7.1, we prove that the LP relaxation of the new model is tighter than the LP relaxation of the old model. We next show that R_2 has no fractional corner points and thus cannot be tightened further.

Proposition 2. *R_2 is a strict subset of R_1 i.e. $R_2 \subset R_1$*

As in Proposition 1, we prove R_2 is a strict subset of R_1 in two steps. First we show that R_2 is a subset of R_1 and then we show that R_2 is not equal to R_1.

Consider a point $P \in R_2$. It satisfies the constraints (74.a), (75) and (78)-(84). We prove that it also lies in R_1 by showing that it satisfies the constraints (74.a)-(79) and (84). Constraints (74.a), (75), (78)-(79) and (84) are trivially satisfied since they are common for both sets.

$$For \quad n = 1, \ldots, NN, \quad k = 1, \ldots, NE$$

Proving $(74.b)$: $\quad z_{nk} \leq DEe_{nk}$

$$z_{nk} \leq UP_k e_{nk} \tag{82}$$

$\Rightarrow \quad$ {using $DE \geq UP_k$ (definition)}

$$z_{nk} \leq DEe_{nk}$$

Hence satisfied.

Proving $(74.c)$: $\quad d_n - DE(1 - e_{nk}) \leq z_{nk}$

$$\sum_{k'=1}^{NE} z_{nk'} = d_n \tag{83}$$

$$\equiv \quad \{\text{splitting sum}\}$$

$$z_{nk} + \sum_{k'=1, k' \neq k}^{NE} z_{nk'} = d_n$$

$$\equiv \quad \{\text{rearranging}\}$$

$$\sum_{k'=1, k' \neq k}^{NE} z_{nk'} = d_n - z_{nk} \tag{85}$$

$$\sum_{k'=1}^{NE} e_{nk'} = 1, \tag{75}$$

$$\equiv \quad \{\text{splitting sum}\}$$

$$e_{nk} + \sum_{k'=1, k' \neq k}^{NE} e_{nk'} = 1$$

$$\equiv \quad \{\text{rearranging}\}$$

$$\sum_{k'=1, k' \neq k}^{NE} e_{nk'} = 1 - e_{nk}, \tag{86}$$

$$z_{nk} \leq DEe_{nk} \tag{74.b}$$

$$\Rightarrow \quad \{\text{sum over } k'\}$$

$$\sum_{k'=1, k' \neq k}^{NE} z_{nk'} \leq DE \sum_{k'=1, k' \neq k}^{NE} e_{nk'}$$

$$\equiv \quad \{\text{using (85), (86)}\}$$

$$d_n - z_{nk} \leq DE(1 - e_{nk})$$

$$\equiv \quad \{\text{rearranging}\}$$

$$d_n - DE(1 - e_{nk}) \leq z_{nk}$$

Hence satisfied.

Proving $(74.d)$: $\quad z_{nk} \leq d_n$

$$\sum_{k=1}^{NE} z_{nk} = d_n, \tag{83}$$

$$\Rightarrow \quad \{\text{using } 0 \leq z_{nk} \text{ (74.a)}\}$$

$$z_{nk} \leq d_n$$

Hence satisfied.

$$
\begin{aligned}
&\text{Proving (76)}: \quad LO_k e_{nk} \le d_n \\
&LO_k e_{nk} \le z_{nk} && (81)\\
\Rightarrow\ &\quad \{\text{using } z_{nk} \le d_n \text{ (74.d)}\} \\
&LO_k e_{nk} \le d_n \\
&\text{Hence satisfied.}
\end{aligned}
$$

$$
\begin{aligned}
&\text{Proving (77)}: \quad d_n + DEe_{nk} \le UP_k + DE \\
&d_n - DE(1 - e_{nk}) \le z_{nk} && (74.c)\\
\Rightarrow\ &\quad \{\text{using } z_{nk} \le UP_k e_{nk} \text{ (82)}\} \\
&d_n - DE + DEe_{nk} \le UP_k e_{nk} \\
\Rightarrow\ &\quad \{\text{using } 0 \le e_{nk} \le 1 \text{ (84)}\} \\
&d_n - DE + DEe_{nk} \le UP_k \\
\equiv\ &\quad \{\text{rearranging}\} \\
&d_n + DEe_{nk} \le UP_k + DE \\
&\text{Hence satisfied.}
\end{aligned}
$$

Therefore point $P \in R_1$ and $R_2 \subseteq R_1$.

Next we find a point Q such that $Q \in R_1$ and $Q \notin R_2$. Take a point $Q(z, e, d) = ([d, d], [1/2, 1/2], d)$. Here $(n, m) = (1, 2)$, $(LO_k, UP_k, DE) = ([0, d], [d, 2d], 2d)$ where $d \ge 0$. We show $Q \in R_1$ since it satisfies the constraints (74.a)-(79) and (84).

$74.a$:

$0 \le z_{11}$, replacing values we get $0 \le d$

$0 \le z_{12}$, replacing values we get $0 \le d$

$74.b$:

$z_{11} \le DEe_{11}$, replacing values we get $d \le 2d/2$

$z_{12} \le DEe_{12}$, replacing values we get $d \le 2d/2$

$74.c$:

$d_1 - DE(1 - e_{11}) \le z_{11}$, replacing values we get

$d - 2d(1 - 1/2) \le d$

$d_1 - DE(1 - e_{12}) \leq z_{12}$, replacing values we get

$d - 2d(1 - 1/2) \leq d$

$74.d$:

$z_{11} \leq d_1$, replacing values we get $d \leq d$

$z_{12} \leq d_1$, replacing values we get $d \leq d$

75 :

$e_{11} + e_{12} = 1$, replacing values we get $1/2 + 1/2 = 1$

76 :

$LO_1 e_{11} \leq d_1$, replacing values we get $0 \times 1/2 \leq d$

$LO_2 e_{12} \leq d_1$, replacing values we get $d \times 1/2 \leq d$

77 :

$DEe_{11} + d_1 \leq UP_1 + DE$, replacing values we get

$2d \times 1/2 + d \leq d + 2d$

$DEe_{12} + d_1 \leq UP_2 + DE$, replacing values we get

$2d \times 1/2 + d \leq 2d + 2d$

78 :

$DE \geq d_1$, replacing values we get $2d \geq d$

84 :

$0 \leq e_{11} \leq 1$, replacing values we get $0 \leq 1/2 \leq 1$

$0 \leq e_{12} \leq 1$, replacing values we get $0 \leq 1/2 \leq 1$

Therefore point $Q \in R_1$. To show that $Q \notin R_2$ consider equation 83:

$$z_{11} + z_{12} = d, \text{ replacing values we get } d + d = 2d$$

Therefore $Q \notin R_2$ and $R_2 \subset R_1$. Proposition 2 therefore shows that the LP relaxation of the new model (R_2) has a tighter bound than the LP relaxation of the old model (R_1).

We next show that in fact R_2 has the tightest relaxation possible by showing that a point with fraction value for e_{nk} will never be a corner point.

Proposition 3. *If point $P \in R_2$ has a fractional value for e_{nk}, P cannot be a corner point of R_2.*

Proof. Consider a point $P(z, e, d) \in R_2$ with at least one fractional value for e_{nk} i.e. $0 < e_{n'k'} < 1$ for some n', k'. Let $e_{n'k'} = t$. Construct another point P_1 that has the same components of P for $n \neq n'$. For $n = n'$ take (z, e, d) as follows:

$$For \quad k = 1,\ \ldots,\ NE$$

$$\begin{aligned}
z_{n'k'} &= 0 \\
z_{n'k} &= \frac{z^P_{n'k}}{1-t}, \quad for\ k \neq k' \\
e_{n'k'} &= 0 \\
e_{n'k} &= \frac{e^P_{n'k}}{1-t}, \quad for\ k \neq k' \\
d_{k'} &= \frac{d^P_{n'} - z^P_{n'k'}}{1-t}
\end{aligned}$$

Here $z^P_{n'k'}$, $e^P_{n'k'}$, $d^P_{n'}$ are the corresponding values of point P. We show that $P_1 \in R_2$ since it satisfies all the constraints:

81 :

$$\begin{aligned}
& LO_{k'} e_{n'k'} \leq z_{n'k'} \\
\equiv\ & LO_{k'} \times 0 \leq 0 && \text{(definition)}
\end{aligned}$$

$$\begin{aligned}
& LO_k e_{n'k} \leq z_{n'k}, \quad k \neq k' \\
\equiv\ & LO_k \frac{e^P_{n'k}}{1-t} \leq \frac{z^P_{n'k}}{1-t}, \quad k \neq k' && \text{(definition)} \\
\equiv\ & LO_k e^P_{n'k} \leq z^P_{n'k}, \quad k \neq k' && (0 < t < 1)
\end{aligned}$$

Satisfied since $P \in R_2$.

82 :

$$\begin{aligned}
& z_{n'k'} \leq UP_{k'} e_{n'k'} \\
\equiv\ & 0 \leq UP_{k'} \times 0 && \text{(definition)}
\end{aligned}$$

$$z_{n'k} \leq UP_k e_{n'k}, \quad k \neq k'$$

$$\equiv \frac{z^P_{n'k}}{1-t} \leq UP_k \frac{e^P_{n'k}}{1-t}, \quad k \neq k' \qquad \text{(definition)}$$

$$\equiv z^P_{n'k} \leq UP_k e^P_{n'k}, \quad k \neq k' \qquad (0 < t < 1)$$

Satisfied since $P \in R_2$.

75 :

$$\sum_{k=1}^{NE} e_{n'k}$$

$$= 0 + \sum_{k=1,k\neq k'}^{NE} e_{n'k} \qquad \text{(splitting sum)}$$

$$= \sum_{k=1,k\neq k'}^{NE} \frac{e^P_{n'k}}{1-t} \qquad \text{(definition)}$$

$$= \frac{1-e^P_{n'k'}}{1-t} \qquad \left(\sum_{k=1}^{NE} e^P_{n'k} = 1\right)$$

$$= \frac{1-t}{1-t} \qquad \text{(definition)}$$

$$= 1 \qquad (0 < t < 1)$$

83 :

$$\sum_{k=1}^{NE} z_{n'k}$$

$$= 0 + \sum_{k=1,k\neq k'}^{NE} z_{n'k} \qquad \text{(splitting sum)}$$

$$= \sum_{k=1,k\neq k'}^{NE} \frac{z^P_{n'k}}{1-t} \qquad \text{(definition)}$$

$$= \frac{d^P_{n'}-e^P_{n'k'}}{1-t} \qquad \left(\sum_{k=1}^{NE} z^P_{n'k} = d^P_{n'}\right)$$

$$= d_{n'} \qquad \text{(definition)}$$

84 :

$$\sum_{k=1}^{NE} e^P_{n'k} = 1 \qquad (P \in R_2)$$

$$\equiv e^P_{n'k'} + \sum_{k=1,k\neq k'}^{NE} e^P_{n'k} = 1 \qquad \text{(splitting sum)}$$

$$\equiv t + \sum_{k=1,k\neq k'}^{NE} e^P_{n'k} = 1 \qquad \text{(definition)}$$

$$\equiv \sum_{k=1,k\neq k'}^{NE} e^P_{n'k} = 1 - t \qquad \text{(rearranging)}$$

$$\equiv 0 \leq e^P_{n'k} \leq 1 - t, \quad k \neq k' \qquad (e^P_{n'k} \geq 0)$$

$$\equiv 0 \leq \frac{e^P_{n'k}}{1-t} \leq 1, \quad k \neq k' \qquad (1 - t > 0)$$

$$\equiv 0 \leq e_{n'k} \leq 1, \quad k \neq k' \qquad \text{(definition)}$$

79 :

$$\sum_{k=1}^{NE} z^P_{n'k} = d^P_{n'} \qquad (P \in R_2)$$

$$\equiv z^P_{n'k'} + \sum_{k=1,k\neq k'}^{NE} z^P_{n'k} = d^P_{n'} \qquad \text{(splitting sum)}$$

$$\equiv \sum_{k=1,k\neq k'}^{NE} z^P_{n'k} = d^P_{n'} - z^P_{n'k'} \qquad \text{(rearranging)}$$

$$\equiv d^P_{n'} - z^P_{n'k'} \geq 0 \qquad (z^P_{n'k} \geq 0)$$

$$\equiv \frac{d^P_{n'} - z^P_{n'k'}}{1-t} \geq 0 \qquad (1 - t > 0)$$

$$\equiv d_{n'} \geq 0 \qquad \text{(definition)}$$

$74.a$:

$$z^P_{n'k} \geq 0, \quad k \neq k' \qquad (P \in R_2)$$

$$\equiv \frac{z^P_{n'k}}{1-t} \geq 0, \quad k \neq k' \qquad (1 - t > 0)$$

$$\equiv z_{n'k} \geq 0, \quad k \neq k' \qquad \text{(definition)}$$

Therefore $P_1 \in R_2$.

Similar to P_1 construct point P_2 having the same components as P for $n \neq n'$. For $n = n'$ take (z, e, d) as follows:

$$For \quad k = 1, \ \ldots, \ NE$$

$$z_{n'k'} = \frac{z^P_{n'k'}}{t}$$
$$z_{n'k} = 0 \quad for\ k \neq k'$$
$$e_{n'k'} = 1$$
$$e_{n'k} = 0 \quad for\ k \neq k'$$
$$d_{n'} = \frac{z^P_{n'k'}}{t}$$

As before $z^P_{n'k'}$, $e^P_{n'k'}$, $d^P_{n'}$ are the corresponding values of point P. We show that $P_2 \in R_2$ since it satisfies all the constraints:

81 :

$$
\begin{aligned}
& LO_{k'}e^P_{n'k'} \leq z^P_{n'k'} && (P \in R_2)\\
\equiv\ & LO_{k'} \times t \leq z^P_{n'k'} && \text{(definition)}\\
\equiv\ & LO_{k'} \leq \frac{z^P_{n'k'}}{t} && (0 < t)\\
\equiv\ & LO_{k'} \leq z_{n'k'} && \text{(definition)}\\
\equiv\ & LO_{k'} \times 1 \leq z_{n'k'} &&\\
\equiv\ & LO_{k'}e_{n'k'} \leq z_{n'k'} && \text{(definition)}
\end{aligned}
$$

$$
\begin{aligned}
& LO_k e_{n'k} \leq z_{n'k}, \quad k \neq k' &&\\
\equiv\ & LO_k \times 0 \leq 0, \quad k \neq k' && (definition)
\end{aligned}
$$

82 :

$$
\begin{aligned}
& z^P_{n'k'} \leq UP_{k'}e^P_{n'k'} && (P \in R_2)\\
\equiv\ & z^P_{n'k'} \leq UP_{k'} \times t && \text{(definition)}\\
\equiv\ & \frac{z^P_{n'k'}}{t} \leq UP_{k'} && (0 < t)\\
\equiv\ & z_{n'k'} \leq UP_{k'} \times 1 && \text{(definition)}\\
\equiv\ & z_{n'k'} \leq UP_{k'}e_{n'k'} && \text{(definition)}
\end{aligned}
$$

$$
\begin{aligned}
& z_{n'k} \leq UP_k e_{n'k}, \quad k \neq k' &&\\
\equiv\ & 0 \leq UP_k \times 0, \quad k \neq k' && \text{(definition)}
\end{aligned}
$$

75 :

$$
\begin{aligned}
&\sum_{k=1}^{NE} e_{n'k} \\
&\quad = e_{n'k'} + \sum_{k=1,k\neq k'}^{NE} e_{n'k} && \text{(splitting sum)} \\
&\quad = 1 + 0 && \text{(definition)}
\end{aligned}
$$

83 :

$$
\begin{aligned}
&\sum_{k=1}^{NE} z_{n'k} \\
&\quad = z_{n'k'} + \sum_{k=1,k\neq k'}^{NE} z_{n'k} && \text{(splitting sum)} \\
&\quad = \frac{z^P_{n'k'}}{t} + 0 && \text{(definition)} \\
&\quad = d_{n'} && \text{(definition)}
\end{aligned}
$$

84 :

$$
\begin{aligned}
&e_{n'k'} = 1 && \text{(definition)} \\
&e_{n'k} = 0, \quad k \neq k' && \text{(definition)}
\end{aligned}
$$

79 :

$$
\begin{aligned}
&z^P_{n'k'} \geq 0 && (P \in R_2) \\
\equiv\ &\frac{z^P_{n'k'}}{t} \geq 0 && (0 < t) \\
\equiv\ &d_{n'} \geq 0 && \text{(definition)}
\end{aligned}
$$

$74.a$:

$$
\begin{aligned}
&z^P_{n'k'} \geq 0 && (P \in R_2) \\
\equiv\ &\frac{z^P_{n'k'}}{t} \geq 0 && (0 < t) \\
\equiv\ &z_{n'k'} \geq 0 && \text{(definition)}
\end{aligned}
$$

$$
\begin{aligned}
&z_{n'k} \geq 0, \quad k \neq k' \\
\equiv\ &0 \geq 0, \quad k \neq k' && \text{(definition)}
\end{aligned}
$$

Therefore $P_2 \in R_2$.

$$\begin{aligned} P_1 &= ((\frac{z^P_{n'k}}{1-t}, 0), (\frac{e^P_{n'k}}{1-t}, 0), \frac{d'^P_i - z^P_{n'k'}}{1-t}) \\ P_2 &= ((0, \frac{z^P_{n'k'}}{t}), (0, 1), \frac{z^P_{n'k'}}{t}) \\ \Rightarrow P_1(1-t) + P_2 t &= ((z^P_{n'k}, z^P_{n'k'}), (e^P_{n'k}, t), d'^P_i) \\ &= P \end{aligned}$$

Since P can be represented as a linear combination of two other points belonging to R_2, P cannot be a corner point of R_2. This implies that LP relaxation (R_2) of the new model will provide only integer solutions. Therefore, the new model has a tight relaxation. □

7.3 ESR Configuration Improvement

For a given network of nodes and links, one aspect of the problem is to determine the location of ESRs and the set of downstream nodes that are to be served by each ESR. We need a set of constraints to model a valid network configuration. In the following section we repeat the set of constraints that model such a network as laid out earlier. We then show that the model is not tight, i.e. its linear relaxation is not guaranteed to have integral corner points. In Section 7.3.2 we then describe an improved model and prove its tightness.

7.3.1 Initial Model

Consider a tree network of n nodes.

Parameters:

A_n = Set of ancestor nodes of node n, $n = 1, \ldots, NN$.

D_n = Set of descendant nodes of node n, $n = 1, \ldots, NN$.

C_n = Set of child nodes of node n, $n = 1, \ldots, NN$.

P_n = Parent node of node n, $n = 1, \ldots, NN$.

Variables:

s_{nm} = 1 if ESR at n^{th} node serves the demand of m^{th} node, n = 1, ..., NN, $m \in D_n \cup \{n\}$.

Constraints:

We can use the following set of constraints to describe the set of valid network configurations as described earlier:

$$s_{mm} \le s_{nn}, \quad n = 1, \ldots, NN, \quad m \in D_n, \tag{87}$$

$$s_{nm} \le s_{nn}, \quad n = 1, \ldots, NN, \quad m \in D_n, \tag{88}$$

$$\sum_m s_{mn} = 1, \quad n = 1, \ldots, NN, \quad m \in A_n \cup \{n\}, \tag{89}$$

$$s_{nm} \le 1 - s_{oo}, \quad n = 1, \ldots, NN, \quad m \in D_n, \quad o \in D_n \cup A_m, \tag{90}$$

$$s_{nm} \in \{0, 1\}, \quad n = 1, \ldots, NN, \quad m \in D_n \cup \{n\}. \tag{91}$$

Proposition 4. *The linear relaxation of S is not tight.*

Proof. Let the linear relaxation of set S be R. Instead of constraint (91) we will have the following constraint:

$$0 \le s_{nm} \le 1, \quad n = 1, \ldots, NN, \quad m \in D_n \cup \{n\}. \tag{92}$$

Consider a small network of 3 nodes with node 1 as root and node 2 child of node 1, and node 3 child of node 2. For a point to belong to R, the following constraints must be met:

$$s_{22} \le s_{11}, \tag{87.a}$$

$$s_{33} \le s_{11}, \tag{87.b}$$

$$s_{33} \le s_{22}, \tag{87.c}$$

$$s_{12} \le s_{11}, \tag{88.a}$$

$$s_{13} \le s_{11}, \tag{88.b}$$

$$s_{23} \le s_{22}, \tag{88.c}$$

$$s_{11} = 1, \tag{89.a}$$

$$s_{12} + s_{22} = 1, \tag{89.b}$$

$$s_{13} + s_{23} + s_{33} = 1, \tag{89.c}$$

$$s_{13} \le 1 - s_{22}, \tag{90.a}$$

$$0 \le s_{11}, s_{12}, s_{13}, s_{22}, s_{23}, s_{33} \le 1. \tag{92}$$

Since $s_{11} = 1$, we replace its value in the constraints and replace repeating constraints to get the following set:

$$s_{33} \le s_{22}, \tag{87.c}$$

$$s_{23} \le s_{22}, \tag{88.c}$$

$$s_{11} = 1, \tag{89.a}$$

$$s_{12} + s_{22} = 1, \tag{89.b}$$

$$s_{13} + s_{23} + s_{33} = 1, \tag{89.c}$$

$$s_{13} \leq 1 - s_{22}, \tag{90.a}$$

$$0 \leq s_{12}, s_{13}, s_{22}, s_{23}, s_{33} \leq 1. \tag{92}$$

Consider a point P defined as:
$P\{s_{11}, s_{12}, s_{13}, s_{22}, s_{23}, s_{33}\} = \{1, \frac{1}{2}, 0, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}\}$. Since it satisfies all the constraints, $P \in R$. We now show that P cannot be described as a linear combination of two distinct points that belong to R.
Consider two points $Q_1, Q_2 \in R$ such that:

$$P = tQ_1 + (1-t)Q_2, \quad 0 < t < 1$$

$$s_{11}^{Q_1} = s_{11}^{Q_2} = 1 \qquad \{89.a\}$$

$$\begin{aligned} & s_{13}^{P} = 0 && \{\text{definition}\} \\ \Rightarrow & s_{13}^{Q_1} = s_{13}^{Q_2} = 0 && \{92.c\} \end{aligned} \tag{93}$$

$$\begin{aligned} & s_{33} + s_{23} \leq 2s_{22} && \{\text{adding 87.c and 88.c}\} \\ \Rightarrow & 1 - s_{13}^{Q_1} \leq 2s_{22}^{Q_1} && \{89.c\} \\ \Rightarrow & \frac{1}{2} \leq s_{22}^{Q_1} && \{93\} \\ \Rightarrow & \frac{1}{2} \leq s_{22}^{Q_2} && \{\text{Similarly}\} \\ \Rightarrow & s_{22}^{Q_1} = s_{22}^{Q_2} = \frac{1}{2} && \{s_{22}^{P} = \frac{1}{2}\} \end{aligned} \tag{94}$$

$$\begin{aligned} & s_{12} + s_{22} = 1 && \{89.b\} \\ \Rightarrow & s_{12}^{Q_1} = s_{12}^{Q_2} = \frac{1}{2} && \{94\} \end{aligned}$$

$$\begin{aligned} & s_{33} \leq s_{22} && \{87.c\} \\ \Rightarrow & s_{33}^{Q_1} \leq \frac{1}{2} && \{94\} \\ & s_{23} \leq s_{22} && \{88.c\} \\ \Rightarrow & s_{23}^{Q_1} \leq \frac{1}{2} && \{94\} \end{aligned}$$

$$\Rightarrow s_{23}^{Q_1} = s_{33}^{Q_1} = \frac{1}{2} \qquad \{89.c\}$$
$$\Rightarrow s_{23}^{Q_2} = s_{33}^{Q_2} = \frac{1}{2} \qquad \{\text{Similarly}\}$$

Therefore $P = Q_1 = Q_2$. Since P cannot be expressed as a linear combination of two distinct points, P is a corner point of R. And since P contains non-integral values for s_{nm}, relaxation R is not tight. □

7.3.2 Improved Model (Model-4)

A new model is proposed which although maintains the same structure as the initial model, it does so using tighter constraints. The primary insight about the structure is expressed in the second constraint mentioned below. A node i serves its child j if and only if it serves all the nodes downstream of j. Consider set S_2 defined by the following set of constraints:

$$\sum_m s_{mn} = 1, \quad n = 1, \ldots, NN, \quad m \in A_n \cup \{n\}, \tag{89}$$
$$s_{nm} = s_{nk}, \quad n = 1, \ldots, NN, \quad m \in C_n, \quad k \in D_m, \tag{95}$$
$$s_{nm} \in \{0, 1\}, \quad n = 1, \ldots, NN, \quad m \in D_n \cup \{n\}. \tag{91}$$

Proposition 5. *The linear relaxation of S_2 is tight.*

Proof. Let the linear relaxation of set S_2 be R_2. Instead of constraint (91) we will have the following constraint:

$$0 \le s_{nm} \le 1, \quad n = 1, \ldots, NN, \quad m \in D_n \cup \{n\}. \tag{96}$$

We will show that R_2 is tight by showing any point P, with a non-integer component can be expressed as a linear combination of two distinct points from R_2.
Consider a point $P \in R_2$ with $0 < s_{n'n'} = t < 1$ for some n'. Let n' be the first such node in the path from root.

Claim 5.1. $s_{nn} = 1, \quad n \in A_{n'}$

Proof. s_{nn} cannot be fractional since n' is the first such node from root by definition. Assume $s_{nn} = 0$ for some $n \in A_{n'}$.
Let $E_{nn'} = (D_n \cup \{n\}) \cap (A'_n \cup \{n'\})$.

$$s_{nn} = 0$$
$$\equiv \qquad \{\text{using } \sum_m s_{mn} = 1 \ (89)\}$$

$$\sum_m s_{mn} = 1 \qquad m \in A_n$$

$$\sum_m s_{mn'} = 1 \qquad m \in A'_n \cup \{n'\}$$

$\equiv$ {splitting sum}

$$\sum_m s_{mn'} + \sum_k s_{kn'} = 1 \qquad m \in A_n, k \in E_{nn'}$$

$\equiv$ {using $s_{nm} = s_{nk}$ (95)}

$$\sum_m s_{mn} + \sum_k s_{kn'} = 1 \qquad m \in A_n, k \in E_{nn'}$$

$\equiv$ {using $\sum_m s_{mn} = 1$ from above}

$$1 + \sum_k s_{kn'} = 1 \qquad k \in E_{nn'}$$

$\equiv$ {simplifying}

$$\sum_k s_{kn'} = 0 \qquad k \in E_{nn'}$$

$\equiv$ {using $0 \leq s_{nm} \leq 1$ (96)}

$$s_{kn'} = 0 \qquad k \in E_{nn'}$$

$\Rightarrow$ {since $n' \in E_{nn'}$ }

$$s_{n'n'} = 0$$

But this is a contradiction since we know $s_{n'n'}$ is fractional. Therefore s_{nn} cannot be fractional and it cannot be 0.

$$s_{nn} = 1, \quad n \in A_{n'} \tag{97}$$

□

Claim 5.2. $s_{nm} = 0, \quad n \in A_{p'}, \quad m \in D_n, \quad p' = P_{n'}$

Proof.

$$s_{mm} = 1, \qquad m \in A_{n'}$$

$\equiv$ {using $\sum_m s_{mn} = 1$ (89)}

$$s_{nm} = 0, \qquad n \in A_{p'}, \quad m \in A_{n'}, \quad j \in D_n, \quad p' = P_{n'}$$

$\equiv$ {using $s_{nm} = s_{nk}$ (95)}

$$s_{nm} = 0, \qquad n \in A_{p'}, \quad m \in D_i, \quad p' = P_{n'} \tag{98}$$

$\square$

Consider a point Q_1 with $s_{n'n'} = 0$:

$$s_{nm} = s^P_{nm}, \quad m \notin (A_{n'} \cup D_{n'} \cup \{n'\}) \tag{99}$$

$$s_{nm} = \frac{s^P_{nm}}{1-t}, \quad n \in A_{n'}, j \in D_n \tag{100}$$

$$s_{nm} = 0, \quad n \in (D_{n'} \cup \{n'\}), m \in D_n \tag{101}$$

Claim 5.3. *Point $Q_1 \in R_2$*

Proof. We prove that point Q_1 belongs to R_2 by showing it satisfies the constraints (89), (95) and (96).

For nodes that are not downstream or upstream of n', s_{nm} values are same as that of point P. Therefore they satisfy the constraints since P belongs to R_2.

For the rest of the nodes:

For $n \in A_{n'}$:

$$\text{Proving (89)}: \quad \sum_m s_{mn} = 1$$

$$\{\text{using } s_{nn} = 1 \text{ (97)}\}$$

$$s_{nn} = 1, \qquad n \in A_{n'}$$

$$\{\text{using } s_{nm} = 0 \text{ (98)}\}$$

$$s_{mn} = 0, \qquad n \in A_{n'}, \quad m \in A_n$$

$$\equiv \quad \{\text{summing over m}\}$$

$$\sum_m s_{mn} = 1, \qquad n \in A_{n'}$$

Hence satisfied.

$$\text{Proving (95)}: \quad s_{nm} = s_{nk}$$

$$\{\text{using } s_{nm} = s_{nk} \text{ (95)}\}$$

$$s^P_{nm} = s^P_{nk}, \qquad n \in A_{n'}, \quad m \in C_n, \quad k \in D_m$$

$$\equiv \quad \{\text{dividing by } (1-t) \text{ since } t \neq 1\}$$

$$\frac{s^P_{nm}}{1-t} = \frac{s^P_{nk}}{1-t}, \qquad n \in A_{n'}, \quad m \in C_n, \quad k \in D_m$$

$$\equiv \quad \{\text{using } s_{nm} = \frac{s^P_{nm}}{1-t} \text{ (100)}\}$$

$$s_{nm} = s_{nk}, \qquad n \in A_{n'}, \quad m \in C_n, \quad k \in D_m$$

Hence satisfied.

Proving (96) : $\quad 0 \leq s_{mn'} \leq 1$

$$\{\text{using } \sum_m s_{mn} = 1 \ (89)\}$$

$$\sum_m s^P_{mn'} = 1, \qquad m \in A_{n'} \cup \{n'\}$$

$$\equiv \quad \{\text{splitting sum}\}$$

$$\sum_m s^P_{mn'} + s^P_{n'n'} = 1, \qquad m \in A_{n'}$$

$$\equiv \quad \{\text{using } s^P_{n'n'} = t\}$$

$$\sum_m s^P_{mn'} = 1 - t, \qquad m \in A_{n'}$$

$$\equiv \quad \{\text{using } s^P_{mn'} \geq 0 \ (96)\}$$

$$0 \leq s^P_{mn'} \leq 1 - t, \qquad m \in A_{n'}$$

$$\equiv \quad \{\text{dividing by } (1-t) \text{ since } t \neq 1\}$$

$$0 \leq \frac{s^P_{mn'}}{1-t} \leq 1, \qquad m \in A_{n'}$$

$$\equiv \quad \{\text{using } s_{nm} = \frac{s^P_{nm}}{1-t} \ (100)\}$$

$$0 \leq s_{mn'} \leq 1, \qquad m \in A_{n'}$$

Hence satisfied.

For $n \in D_{n'} \cup \{n'\}$:

Proving (89) : $\quad \sum_m s_{mn} = 1$

$$\{\text{using } \sum_m s_{mn} = 1 \ (89)\}$$

$$\sum_m s^P_{mn'} = 1, \qquad m \in A_{n'} \cup \{m'\}$$

$$\{\text{using } s_{nm} = 0 \ (98)\}$$

$$s^P_{kn'} + s^P_{n'n'} = 1, \qquad k = P_{n'}$$

$$\equiv \quad \{\text{using } s^P_{n'n'} = t\}$$

$$s^P_{kn'} = 1 - t, \qquad k = P_{n'}$$

$$\equiv \quad \{\text{using } s_{nm} = \frac{s^P_{nm}}{1-t} \ (100)\}$$

$$s_{kn'} = 1, \qquad k = P_{n'}$$

$$\equiv \quad \{\text{using } s_{nm} = 0 \ (101)\}$$

$$\sum_m s_{mn} = 1, \qquad n \in D_{n'} \cup \{n'\}, m \in A_n$$

Hence satisfied.

$$\text{Proving (95)}: \quad s_{nm} = s_{nk}$$

$$\{\text{using } s_{nm} = 0 \ (101)\}$$

$$s_{nm} = 0 \qquad n \in D_{n'} \cup \{n'\}, m \in D_n$$

$$\equiv$$

$$s_{nm} = s_{nk} \qquad n \in D_{n'} \cup \{n'\}, m \in D_n, k \in C_n$$

Hence satisfied.

$$\text{Proving (96)}: \quad 0 \le s_{nm} \le 1$$

$$\{\text{using } s_{nm} = 0 \ (101)\}$$

$$s_{nm} = 0 \qquad n \in D_{n'} \cup \{n'\}, m \in D_n$$

Hence satisfied.

Therefore point $Q_1 \in R_2$.

□

Similarly consider point Q_2 with $s_{n'n'} = 1$:

$$s_{nm} = s^P_{nm}, \quad m \notin (A_{n'} \cup D_{n'} \cup \{n'\}) \tag{102}$$

$$s_{nm} = 0, \quad n \in A_{n'}, m \in D_n \tag{103}$$

$$s_{nm} = \frac{s^P_{nm}}{t}, \quad n \in (D_{n'} \cup \{n'\}), m \in D_n \cup \{n\} \tag{104}$$

Claim 5.4. *Point $Q_2 \in R_2$*

Proof. We prove that point Q_2 belongs to R_2 by showing it satisfies the constraints (89), (95) and (96)

For nodes that are not downstream or upstream of n', s_{nm} values are same as that of point P. Therefore they satisfy the constraints since P belongs to R_2.

For the rest of the nodes:

For $n \in A_{n'}$:

$$\text{Proving (89)}: \quad \sum_m s_{mn} = 1$$

$$\{\text{using } s_{nn} = 1 \ (97)\}$$

$$s_{nn} = 1, \qquad n \in A_{n'}$$

$$\{\text{using } s_{nm} = 0 \ (98)\}$$

$$s_{mn} = 0, \qquad n \in A_{n'}, \quad m \in A_n$$

$$\equiv \quad \{\text{summing over m}\}$$

$$\sum_m s_{mn} = 1, \qquad n \in A_{n'}$$

Hence satisfied.

Proving (95) : $\quad s_{nm} = s_{nk}$

$$\{\text{using } s_{nm} = 0 \ (103)\}$$

$$s_{nm} = 0 \qquad n \in A_{n'}, m \in D_n$$

$$\equiv$$

$$s_{nm} = s_{nk} \qquad n \in A_{n'}, m \in D_n, k \in C_n$$

Hence satisfied.

Proving (96) : $\quad 0 \leq s_{nm} \leq 1$

$$\{\text{using } s_{nm} = 0 \ (103)\}$$

$$s_{nm} = 0 \qquad n \in A_{n'}, m \in D_n$$

Hence satisfied.

For $n \in D_{n'} \cup \{n'\}$:

Proving (89) : $\quad \sum_m s_{mn} = 1$

$$\{\text{using } \sum_m s_{mn} = 1 \ (89)\}$$

$$\sum_m s^P_{mn} = 1, \qquad m \in A_n \cup \{n\}$$

$$\equiv \quad \{\text{splitting sum}\}$$

$$\sum_m s^P_{mn} + \sum_k s^P_{kn} = 1, \qquad m \in D_{n'} \cup \{n'\}, k \in A_{n'}$$

$$\equiv \quad \{\text{using } \sum_k s^P_{kn} = 1 - t \ \}$$

$$\sum_m s^P_{mn} = t, \qquad m \in D_{n'} \cup \{n'\}$$

$$\equiv \quad \{\text{dividing by } t \text{ since } t \neq 0 \ \}$$

$$\sum_m \frac{s^P_{mn}}{t} = 1, \qquad m \in D_{n'} \cup \{n'\}$$

$$\equiv \quad \{\text{using } s_{nm} = \frac{s^P_{nm}}{t} \ (104)\}$$

$$\sum_m s_{mn} = 1, \qquad m \in D_{n'} \cup \{n'\}$$

Hence satisfied.

Proving (95) : $\quad s_{nm} = s_{nk}$

$$\{\text{using } s_{nm} = s_{nk} \ (95)\}$$

$$s^P_{nm} = s^P_{nk}, \qquad m \in C_n,\ k \in D_m$$

$$\equiv \quad \{\text{dividing by } t \text{ since } t \neq 0\ \}$$

$$\frac{s^P_{nm}}{t} = \frac{s^P_{nk}}{t}, \qquad m \in C_n,\ k \in D_m$$

$$\equiv \quad \{\text{using } s_{nm} = \frac{s^P_{nm}}{t} \ (104)\}$$

$$s_{nm} = s_{nk}, \qquad m \in C_n,\ k \in D_m$$

Hence satisfied.

Proving (96) : $\quad 0 \leq s_{nm} \leq 1$

$$\{\text{using } \sum_m s_{mn} = 1 \ (89)\}$$

$$\sum_m s^P_{mn} = 1, \qquad m \in A_n \cup \{n\}$$

$$\equiv \quad \{\text{splitting sum}\}$$

$$\sum_m s^P_{mn} + \sum_k s^P_{kn} = 1, \qquad m \in D_{n'} \cup \{n'\}, k \in A_{n'}$$

$$\equiv \quad \{\text{using } \sum_k s^P_{kn} = 1 - t\ \}$$

$$\sum_m s^P_{mn} = t, \qquad m \in D_{n'} \cup \{n'\}$$

$$\{\text{using } 0 \leq s^P_{mn} \ (96)\}$$

$$0 \leq s^P_{mn} \leq t, \qquad m \in D_{n'} \cup \{n'\}$$

$$\equiv \quad \{\text{dividing by } t \text{ since } t \neq 0\ \}$$

$$0 \leq \frac{s^P_{mn}}{t} \leq 1, \qquad m \in D_{n'} \cup \{n'\}$$

$$\equiv \quad \{\text{using } s_{nm} = \frac{s^P_{nm}}{t} \ (104)\}$$

$$0 \leq s_{mn} \leq 1, \qquad m \in D_{n'} \cup \{n'\}$$

Hence satisfied.

Therefore point $Q_2 \in R_2$. □

Claim 5.5. *P is a linear combination of points Q_1 and Q_2 i.e. $P = (1-t)Q_1 + tQ_2$*

Proof. For $m \notin (A_{n'} \cup D_{n'} \cup \{n'\})$:

$$\{\text{using } s^P_{nm} = s^{Q_1}_{nm} \text{ (99) and } s^P_{nm} = s^{Q_2}_{nm} \text{ (102)}\}$$

$$s^P_{nm} = s^{Q_1}_{nm} = s^{Q_2}_{nm}$$

$$\Rightarrow$$

$$s^P_{nm} = (1-t)s^{Q_1}_{nm} + t * s^{Q_2}_{nm}$$

For $n \in A_{n'}, m \in D_n$:

$$s^{Q_1}_{nm} = \frac{s^P_{nm}}{1-t} \qquad \{100\}$$

$$s^{Q_2}_{nm} = 0 \qquad \{103\}$$

$$\Rightarrow$$

$$s^P_{nm} = (1-t)s^{Q_1}_{nm} + t * s^{Q_2}_{nm}$$

For $n \in D_{n'}, m \in D_n$:

$$s^{Q_1}_{nm} = 0 \qquad \{101\}$$

$$s^{Q_2}_{nm} = \frac{s^P_{nm}}{t} \qquad \{104\}$$

$$\Rightarrow$$

$$s^P_{nm} = (1-t)s^{Q_1}_{nm} + t * s^{Q_2}_{nm}$$

Therefore P is a linear combination of points Q_1 and Q_2. □

Since any general point P with a fractional component can be expressed as linear combination of two other points in the set R_2, it implies that such a point P cannot be a corner point and therefore set R_2 is tight. □

This concludes the discussion on the three improvements made to the initial ILP model. Experimental results of the performance of the model after each improvement is presented in Section 7.4. The model in its entirety after the inclusion of all the network components and these improvements is described in Appendix I. Although we have shown the tightness of various subsets of the improved model, the overall set of constraints of the model is still not tight. As such, there remains room for further improvements to the model. In Chapter 8 we describe an initial attempt at an alternative approach to the problem. Instead of using node variables s_{ij} to partition the primary and secondary network, purely edge based variables are used.

7.4 Computational Results

The three pipe cost/ESR cost/ESR allocation improvements were applied sequentially to the initial model (model 1) to give model-2/model-3/model-4 respectively. These four models were tested over eight different networks of varying sizes in order to test their performance and scalability:

- Real World Networks: Three of the networks, Khardi, Shahpur and Mokhada are real life networks from Maharashtra state in India. These regions consist of tribal villages that regularly face extreme water stress during summer months and as a result have to be provided water using tankers.
- Synthetic Networks: The other five networks are artificially created to test the performance of the models across different network sizes (10 to 200). Each of them is a randomly generated branched network. Ranges for the node and link properties are as follows:
  - Number of children nodes: 1 to 5,
  - Elevation (in metres): 100 to 300,
  - Demand (in litres per second): 0.01 to 5,
  - Length of links (in metres): 500 to 5000.

For all four models, the problem statement remains the same, to optimize the total pipe and ESR cost of the network. The number of binary and continuous variables scale with the size of the network. Since all four models solve the problem optimally, the final capital cost of the pipes and ESRs is the same. The performance of each model is measured in terms of three metrics: the total time taken in seconds, the size of the branch and bound tree and the objective value of the LP relaxation. Table 7.1 summarizes the performance of the models. We observe that for each of the eight networks, the time taken improves with each model, resulting in model-4 providing the best performance. Typically the time taken scales with the size of the network. However, this need not always be the case. For example, although gen50 has more nodes (50 nodes) than Mokhada (37 nodes), it is solved in lesser amount of time. This is because apart from the number of nodes being a factor, the network configuration also matters while solving the model.

Table 7.1: Performance of the various models on the eight networks.

Network	Nodes	Binary Variables	Continuous Variables	Objective (Rs)	Metric	Model-1	Model-2	Model-3	Model-4
gen10	10	228	527	2.18E7	Time(s)	1.71	0.96	0.47	**0.32**
					B&B Tree Size	8	14	**2**	10
					LP Obj. (Rs)	7.89E6	7.89E6	**1.02E7**	**1.02E7**
Khardi	11	262	584	4.48E7	Time(s)	0.74	0.36	0.31	**0.24**
					B&B Tree Size	2	**0**	**0**	**0**
					LP Obj. (Rs)	1.14E07	1.16E07	**2.90E07**	**2.90E07**
Shahpur	21	712	1154	4.54E7	Time(s)	11.66	3.36	1.90	**0.72**
					B&B Tree Size	148	27	**4**	**4**
					LP Obj. (Rs)	2.00E7	2.02E7	**2.69E7**	**2.69E7**
Mokhada	37	1848	2066	4.94E7	Time(s)	28.78	7.91	3.48	**2.02**
					B&B Tree Size	965	188	68	**30**
					LP Obj. (Rs)	2.16E7	2.16E7	**3.05E7**	**3.05E7**
gen50	50	3148	2807	9.82E7	Time(s)	373.49	39.31	2.02	**1.27**
					B&B Tree Size	15228	1240	8	**6**
					LP Obj. (Rs)	4.00E7	4.00E7	**5.58E7**	**5.58E7**
gen100	100	11298	5657	2.47E8	Time(s)	197.89	129.75	9.73	**4.18**
					B&B Tree Size	1735	1958	120	**32**
					LP Obj. (Rs)	9.34E7	9.34E7	**1.27E8**	**1.27E8**
gen150	150	24448	8507	3.48E8	Time(s)	2582.35	887.68	18.67	**5.25**
					B&B Tree Size	22401	15498	148	**8**
					LP Obj. (Rs)	8.50E7	8.53E7	**1.35E8**	**1.35E8**
gen200	200	42598	11357	4.55E8	Time(s)	timeout*	timeout*	523.97	**69.80**
					B&B Tree Size	-	-	5034	**858**
					LP Obj. (Rs)	1.12E8	1.12E8	**1.83E8**	**1.83E8**

*timed out after running for 24 hours

Chapter 8

Edge Based Model

Determining the ESR configuration involves two components, partitioning the network into primary and secondary networks and determining the demand each ESR has to serve. In our model, the allocation of nodes to ESRs is done using the node based binary variables s_{ij} which is true if the ESR at the i^{th} node serves the j^{th} node. These variables therefore explicitly determine the demand each ESR serves. An alternative approach to this node based representation of the network is to have an edge based representation. Here instead of the focus being which ESR serves which node, the focus is on which pipes in the network are part of the primary network and which pipes are part of the secondary network.

8.1 Model Details

Consider a tree network of NE edges:

Parameters:

C_i = Set of pipes that are immediately downstream of pipe i, i = 1, . . . , NE.

U_i = Set of pipes that are upstream of pipe i, i = 1, . . . , NE.

D_i = Set of pipes that are downstream of pipe i, i = 1, . . . , NE.

DS = Set of pipes that are immediately downstream of the source.

Variables:

$f_i = 1$ if i^{th} edge belongs to the primary network and $= 0$ if i^{th} edge belongs to the secondary network, i = 1, . . . , e

The primary network connects the source to the ESRs, and the secondary network connects the ESRs to downstream nodes. Therefore pipes starting from the source must belong to the primary network. Also, secondary pipes must be downstream of the primary pipes. And once a pipe is secondary, then any pipes downstream can no longer be primary. We can use the following set of constraints to describe the set S_3 of valid network configurations:

$$f_i = 1, \quad i \in DS, \tag{105}$$

$$f_j \leq f_i, \quad i = 1, \ \ldots, \ NE, \quad j \in C_i, \tag{106}$$

$$f_i \in \{0, 1\}, \quad i = 1, \ \ldots, \ NE. \tag{107}$$

8.2 Tightness Proof

We show that the linear relaxation of S_3 is tight, i.e. it does not contain any corner points with non-integral components.

Proposition 6. *The linear relaxation of S_3 is tight.*

Proof. Let the linear relaxation of set S_3 be R_3. Instead of constraint (107) we will have the following constraint:

$$0 \leq f_i \leq 1, \quad i = 1, \ \ldots, \ NE. \tag{108}$$

We will show that R_3 is tight by showing any point P, with a non-integer component can be expressed as a linear combination of two distinct points from R_3.
Consider a point $P \in R_3$ with $0 < f_i = t < 1$ for some i'. Let i' be the first such edge in the path from source.

Claim 6.1. $f_i = 1, \quad i \in U_{i'}$

Proof. f_i cannot be fractional since i' is the first such edge from source by definition. If $f_i = 0$, then by (106) for all its downstream edges j, $f_i = 0$. But i' is downstream of i and $f_{i'} = t \neq 0$. Therefore f_i cannot be fractional and it cannot be 0.

$$f_i = 1, \quad i \in U_{i'}. \tag{109}$$

□

Consider a point Q_1 with $f_{i'} = 0$:

$$f_i = f_i^P, \quad i \notin (D_{i'} \cup \{i'\}), \tag{110}$$

$$f_i = 0, \quad i \in (D_{i'} \cup \{i'\}). \tag{111}$$

Claim 6.2. *Point $Q_1 \in R_3$*

Proof. For all the edges not downstream of i', constraints (105), (106) and (108) are satisfied since the values are same as point P and $P \in R_3$. Setting $f_i = 0$ for all downstream i also maintains the constraints trivially. Therefore point $Q \in R_3$. $\square$

Similarly consider point Q_2 with $f_{i'} = 1$:

$$f_i = f_i^P, \quad i \notin (D_{i'} \cup \{i'\}), \tag{112}$$

$$f_i = \frac{f_i^P}{t}, \quad i \in (D_{i'} \cup \{i'\}). \tag{113}$$

Claim 6.3. *Point $Q_2 \in R_3$*

Proof. We prove that point Q_2 belongs to R_3 by showing it satisfies the constraints (105), (106) and (108). For edges that are not downstream of i', f_i values are same as that of point P. Therefore they satisfy the constraints since $P \in R_3$. For the rest of the edges:

For $i \in (D_{i'} \cup \{i'\})$: (105) is trivially true since i' (and its downstream edges) cannot be connected to the source since for point P, $f_i \neq 0$.

$$\begin{aligned}
&\text{Proving (106)}: \quad f_j \leq f_i \\
&\quad \{\text{using } f_j^P \leq f_i^P \text{ (106)}\} \\
&f_j^P \leq f_i^P, && i \in (D_{i'} \cup \{i'\}), j \in C_i \\
\equiv &\quad \{\text{dividing by } t \text{ since } t \neq 0\ \} \\
&\frac{f_j^P}{t} \leq \frac{f_i^P}{t}, && i \in (D_{i'} \cup \{i'\}), j \in C_i \\
\equiv &\quad \{\text{using } f_i = \frac{f_i^P}{t} \text{ (113)}\} \\
&f_j \leq f_i, && i \in (D_{i'} \cup \{i'\}), j \in C_i \\
&\text{Hence satisfied.}
\end{aligned}$$

$$\begin{aligned}
&\text{Proving (108)}: \quad 0 \leq f_i \leq 1 \\
&\quad \{\text{using } f_i^P \leq f_{i'}^P \text{ (106) and } 0 \leq f_i^P \text{ (108)}\} \\
&0 \leq f_i^P \leq f_{i'}^P && i \in (D_{i'} \cup \{i'\}) \\
\equiv &\quad \{\text{using } f_i^P = t\ \} \\
&0 \leq f_i^P \leq t && i \in (D_{i'} \cup \{i'\}) \\
\equiv &\quad \{\text{dividing by } t \text{ since } t \neq 0\ \} \\
&0 \leq \frac{f_i^P}{t} \leq 1 && i \in (D_{i'} \cup \{i'\}) \\
\equiv &\quad \{\text{using } f_i = \frac{f_i^P}{t} \text{ (113)}\} \\
&0 \leq f_i \leq 1 && i \in (D_{i'} \cup \{i'\})
\end{aligned}$$

Hence satisfied.

Therefore point $Q_2 \in R_3$. □

Claim 6.4. *P is a linear combination of points Q_1 and Q_2 i.e. $P = (1-t)Q_1 + tQ_2$*

Proof. For $i \notin (D_{i'} \cup \{i'\})$:

$$f_i^P = f_i^{Q_1} = f_i^{Q_2} \qquad \{110, 112\}$$

$$\Rightarrow$$

$$f_i^P = (1-t)f_i^{Q_1} + t * f_i^{Q_2}$$

For $i \in (D_{i'} \cup \{i'\})$:

$$f_i^{Q_1} = 0 \qquad \{111\}$$

$$f_i^{Q_2} = \frac{f_i^P}{t} \qquad \{113\}$$

$$\Rightarrow$$

$$f_i^P = (1-t)f_i^{Q_1} + t * f_i^{Q_2}$$

Therefore P is a linear combination of points Q_1 and Q_2 □

Since any general point P with a fractional component can be expressed as a linear combination of two other points in the set R_3, it implies that such a point P cannot be a corner point and therefore relaxation R_3 is tight.

□

8.3 Computational Results

Table 8.1: Performance of the edge based model.

Network	**Nodes**	**Model-1**	**Model-2**	**Model-3**	**Model-4**	**Edge Model**
gen10	10	1.71	0.96	0.47	**0.32**	0.40
Khardi	11	0.74	0.36	0.31	**0.24**	0.39
Shahpur	21	11.66	3.36	1.90	**0.72**	0.91
Mokhada	37	28.78	7.91	3.48	**2.02**	3.03
gen50	50	373.49	39.31	2.02	**1.27**	1.86
gen100	100	197.89	129.75	9.73	**4.18**	4.97
gen150	150	2582.35	887.68	18.67	**5.25**	16.14
gen200	200	timeout*	timeout*	523.97	**69.80**	100.84

*timed out after running for 24 hours

Table 8.1 displays the performance of the edge based model in comparison to the four models from Chapter 7 in terms of time taken in seconds. The performance of the edge based model is between that of model-3 and model-4. Although we prove that the LP relaxation of the set of constraints described by S_3 is tight, the LP relaxation objective for the overall model is worse. This is due to changes in other constraints of the model. Computing the demand d_n that an ESR at node n serves now becomes more complicated since in this model only edge based variables are considered. Instead of the linear equation for d_n in the original model (38), we now have the following equation:

$$d_n = \sum_{j \in D_i} DE_m * f_i * (1 - f_j), \quad n = 1, \ \ldots, \ NN \quad i = I_n, \quad j = O_n = I_m. \tag{114}$$

This equation is non-linear since it contains the product of two variables f_i and f_j. Since both are binary the product term can be linearized but this introduces many constraints which deteriorates the performance. Therefore the model implemented in the JalTantra system is model-4. Details of the JalTantra system are provided in the next chapter.

Chapter 9

JalTantra System Description

In this chapter we describe the system JalTantra, that we created for the design and optimization of piped water networks. The aim of the system is to be optimal and fast while at the same time providing an interface that is simple and intuitive to use. We discuss the main features of the system and highlight how it achieves the objectives we laid out in Section 1.1. Details regarding how to use the system can be found in Appendix II.

9.1 First Iteration: A Desktop JAVA Application

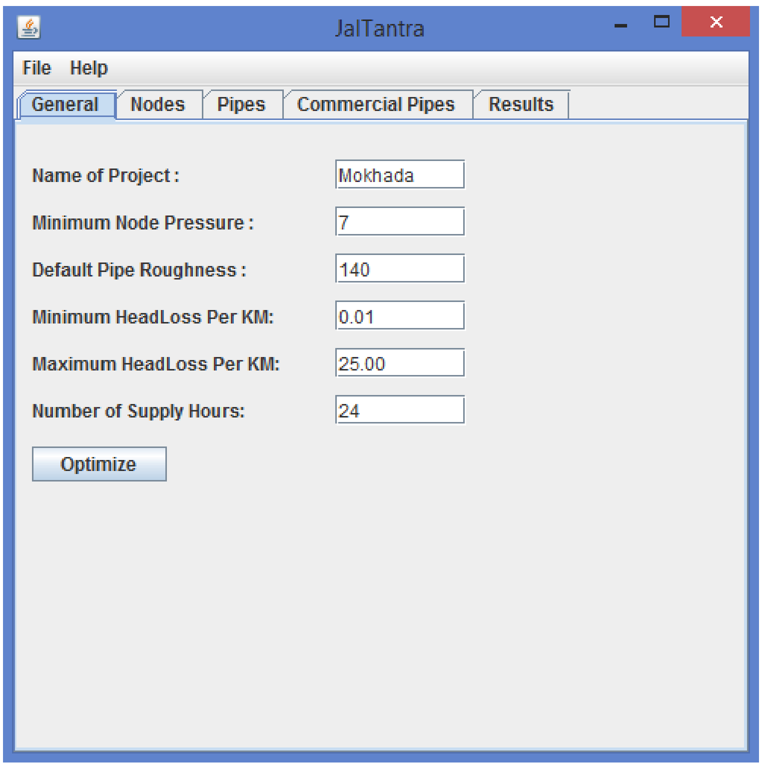


Figure 9.1: General Tab of JalTantra desktop Java application

Government engineers tasked with design of water networks in India work on systems with operating systems ranging from Windows XP to Windows 10. We wanted to develop the system using cross-platform technology that would not only work on the range of machines currently being used by real world practitioners but also be future proof. To that end, the first version of JalTantra was developed as a JAVA [35] desktop application.

The application provided a user interface through which the user could input details of the network as seen in figure 9.1. Input could also be provided via XML files. This version only solved the pipe diameter selection problem described in Chapter 4. Once the network was entered, the user could start the optimization and the final results could be seen on a separate result tab as seen in figure 9.2. For the optimization we first used the linear solver lpsolve 5.5 [9]. This was subsequently replaced with a faster linear solver library GLPK 4.55 [26]. For both lpsolve and GLPK, Java ILP 1.2a [36] was used as the Java interface to the solver libraries.

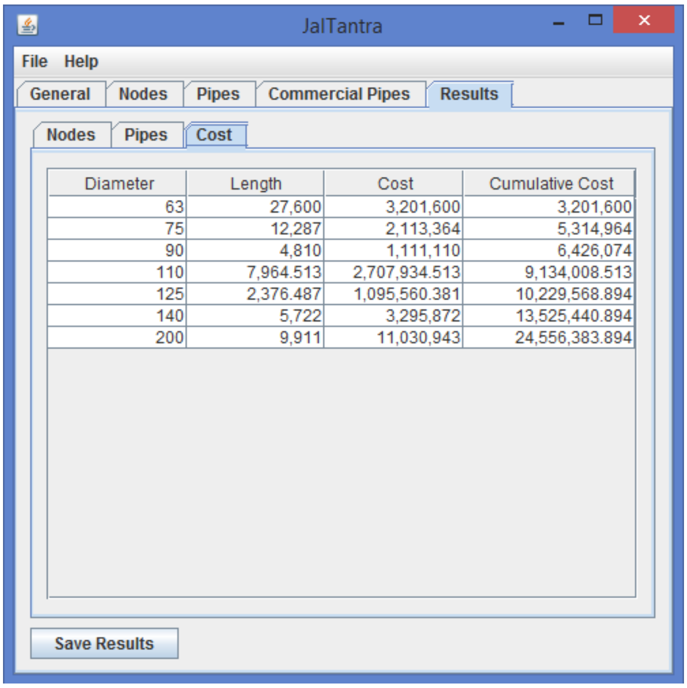


Figure 9.2: Results Tab of JalTantra desktop Java application

Once developed, feedback on JalTantra was sought from designers from the government offices of MJP. We did this by one on one interactions as well as holding multiple tutorial sessions where the engineers could get a hands on experience with the system. During these tutorials, we found that JAVA was not installed on these machines more often than not. Even when installed, sometimes the PATH variables would not be set correctly and thus the machine would not be able to launch JalTantra. We had to frequently intervene and help the engineers with the initial setup due to these issues. We realised issues could also arise when JalTantra or JAVA was updated in the future and a reinstallation would be required by the user. We also wanted to include a GIS based integration in future versions of JalTantra, which would significantly ease in input of network details. With the inclusion of more network components like tanks, pumps and valves, some of the very old machines might have a hard time running the optimization. With these things in mind, we decided to migrate JalTantra to a web system.

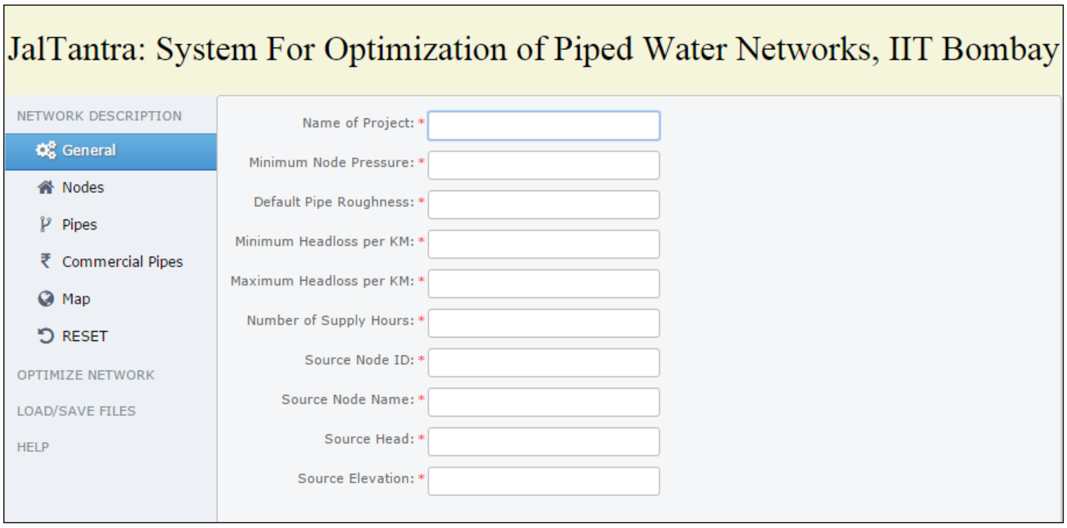


Figure 9.3: JalTantra Web System

9.2 JalTantra 2.0: A Web System

Given the installation, maintenance and scalability issues for a desktop application, we decided to instead develop JalTantra as a web application as seen in figure 9.3. Now all that would be required from the user is a simple web browser. Any updates to the system could be pushed without disrupting the user experience. A web application allows JalTantra to be used on any system. The system is available at http://www.cse.iitb.ac.in/jaltantra . It includes links to usage instructions, a video tutorial and sample input/output files. The system can also be downloaded and run on a local server for offline use. The local server version requires JAVA 7 [35].

9.3 System Architecture

JalTantra is a web application with three principal components as shown in figure 9.4:

Client Interface: The client interface houses the logic used to interact with the user. Its built using javascript libraries jQuery [37] and w2ui [72]. The dagre [17] javascript library was used to generate the spatial coordinates for the network when saving the network as an EPANET file. Once the user has provided all the network details after a preliminary validation step, the network is converted into JSON objects and a request is made to the web server.

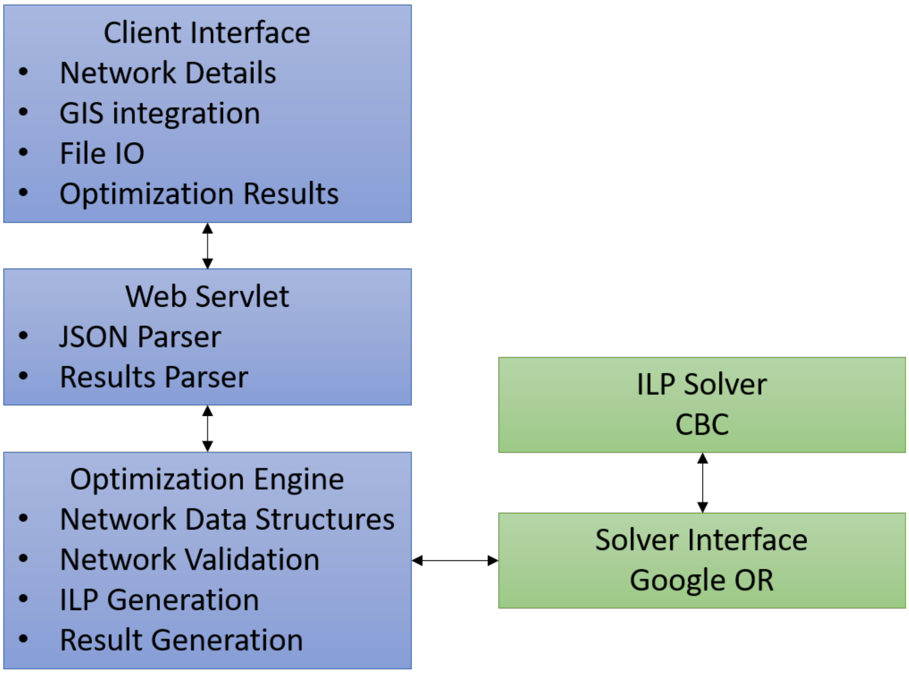


Figure 9.4: JalTantra System Architecture

Web Servlet: A JAVA web servlet acts as the intermediary between the web interface and the optimization engine. Upon receiving a web request, it unpacks the JSON string objects and converts them into JAVA objects and then forwards these to the optimization engine. Upon receiving the results from the optimization engine, it converts the results into JSON objects and sends a response to the web client.

Optimization Engine: The optimization engine houses the core logic of the system. It is here where the provided network is validated and then the optimization problem is generated as an Integer Linear Program (ILP). This program is then provided to a third party ILP solver. Since a lot more constraints and variables were added with the ESR integration, instead of using GLPK, we moved on to a faster solver CBC [22]. We used the Google OR Tools [28] JAVA interface to call the CBC solver. The abstraction provided by the Google OR tools interface allows us to use more powerful commercial solvers in the future without any code rewrite.

9.4 Data Handling

One of the objectives that we had set out when designing JalTantra was that it should be easy to use in terms of input of data, since that is one of the most tedious steps in the design process. Additionally there should be interoperability between JalTantra and legacy systems used by engineers to ease their transition to JalTantra. Data input to JalTantra can be entered manually in the various tabs discussed above or via importing files in various formats. Importing a file will populate all the fields in the tabs of the system.

- BRANCH Files: .bra file format used by BRANCH. Users of BRANCH software can easily test JalTantra using their existing files.
- XML Files: .xml file format. Standard XML format to allow easy input from future third party sources.
- Excel Files: .xls file format. Suited for users having data already present in Excel files.

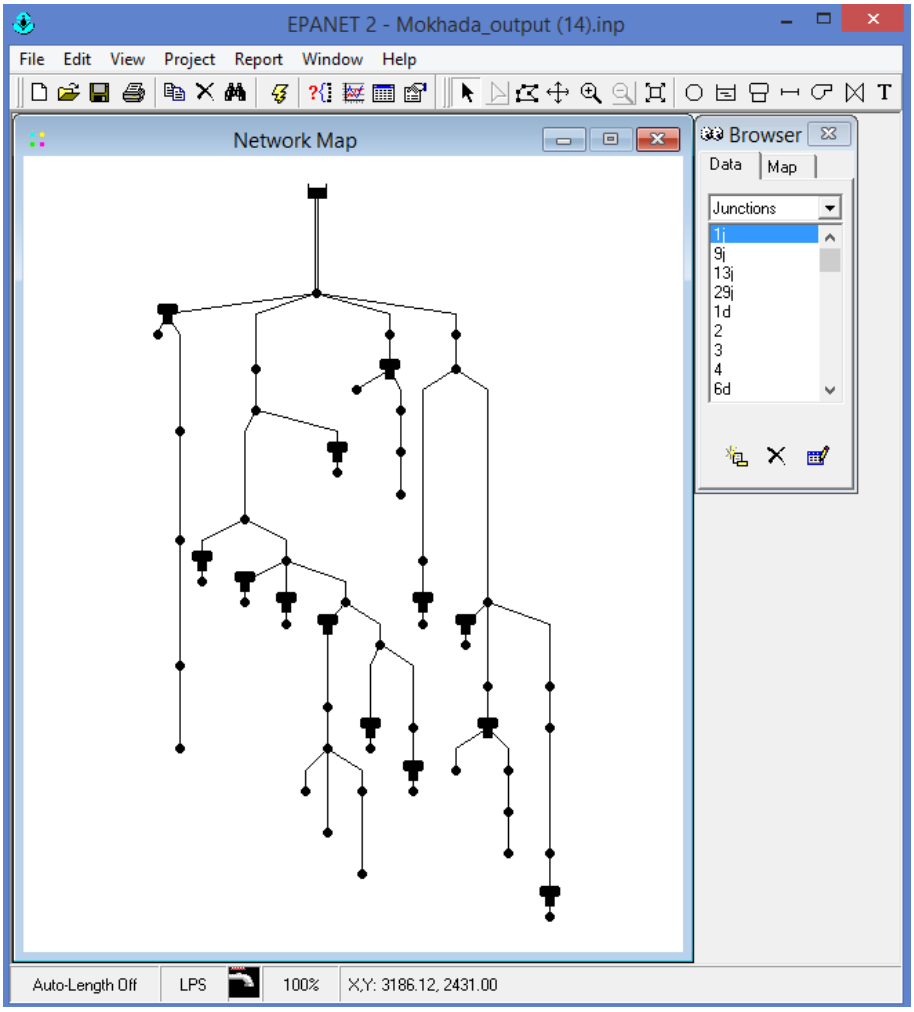


Figure 9.5: EPANET for Mokhada Network

The output of the optimization can be saved as an Excel File. Another option that is included is the ability to export the network as an EPANET file. This is useful since EPANET is the de-facto software used for running water networks simulations. Previously it would be a tedious process to transfer the output of JalTantra or BRANCH and create the EPANET network manually. Figure 9.5 shows the EPANET file generated for the Mokhada network.

9.5 GIS Integration

As mentioned earlier (and can be seen in the figures above), to describe the network we need to provide elevation data for nodes as well as the lengths of the links connecting them. These values are measured using physical surveys at location to ensure accuracy before doing the final design. But for earlier prototype designs to gauge feasibility of the design GIS data is used. This data is looked up and then entered manually. This can be a very tedious process, especially for link distances since the link must typically be manually drawn along the road network. As part of our web system we have integrated google map based GIS [27] which allows the user to add nodes on the map. Links between the nodes can be simply added using the google directions service without having to manually enter the entire path. The tool also allows to view the elevation profile of the paths generated. This is then used to add dummy nodes along the path at points of high elevation. Information like elevations and distances can then be extracted directly into the node/pipe information screens. A sample network created using the GIS tool is shown in figure 9.6.

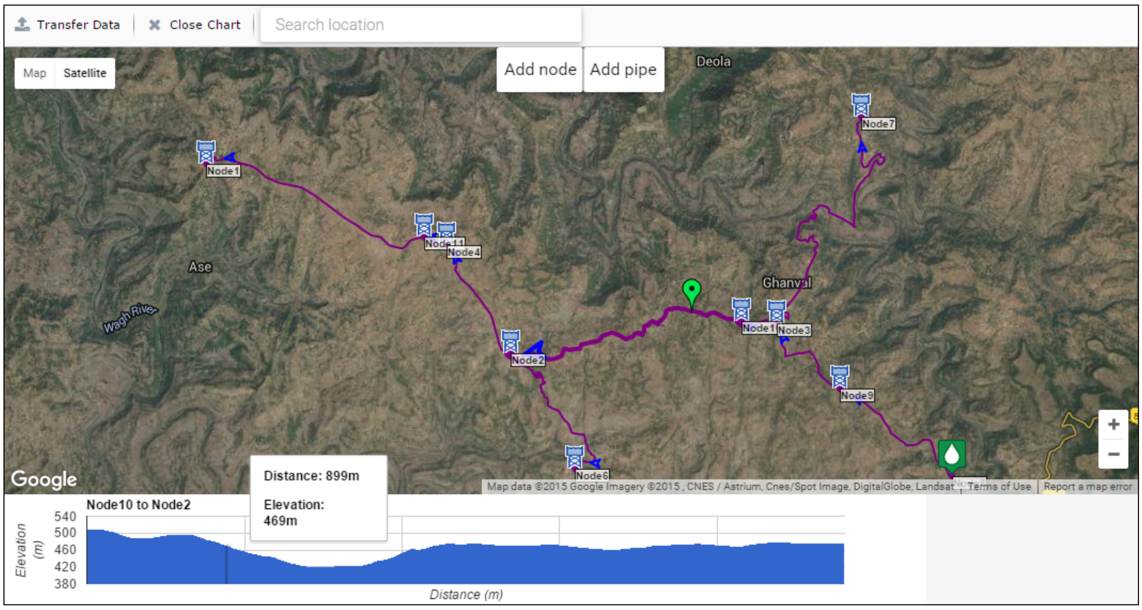


Figure 9.6: GIS Tool in JalTantra

Chapter 10

Conclusion and Future Work

10.1 Conclusion

With this thesis, our objective was to create a system that will help government engineers and other users in the design and optimization of piped water networks. We looked at the cost optimization of rural drinking water schemes in particular. These schemes consist of several network components like pipes, tanks, pumps and valves. We motivated the importance of considering all these components while optimizing as opposed to just the pipe diameters. We presented an ILP model that was used to solve the problem.

Although optimal, the model took a significant amount of time for larger networks, 45 minutes for a network with 150 nodes. We then described a series of three improvements of the model. For each improvement we proved that the improved model is tighter than the initial model. We then presented the performance results of the three improved models along with the initial model over eight networks of various sizes. The 150 node network now takes only 5 seconds to solve. This enables practitioners to consider greater number of iterations of the design for large networks, since each iteration can be optimized in a matter of seconds.

We have implemented our solution in a water network design web system JalTantra, available publicly at http://www.cse.iitb.ac.in/jaltantra . JalTantra also has GIS integration for ease of adding network details. It handles legacy file formats such as BRANCH and Excel files. This helps users to transition to JalTantra smoothly. After the optimization, it can also output the network details as an EPANET file which is used to simulate and validate the network.

We held several tutorial and demo sessions with government engineers from the MJP and received positive feedback. MJP has officially adopted it as one of the software packages to be used in the design of water supply schemes. MEETRA which is responsible for the training of MJP engineers, has integrated JalTantra into its curriculum. Additionally, details of the JalTantra system has been included in a compendium titled "Improving the performance of rural water supply and sanitation sector in Maharashtra".

10.2 Future Work

In the present work we have looked at the design and optimization of piped water networks and in particular rural networks. Some of the assumptions include branched network, fixed demands, minimizing capital and operational cost. Each of these assumptions could be relaxed to generalize the problem further, leading to future work in various directions:

Looped Networks: Although rural networks are branched, urban and semi-urban networks are typically looped. A natural extension of the current work would be extend the model to include looped networks.

Alternative Objective Criteria: The introduction of operational cost opens up interesting possibilities in determining the objective function. Currently we use the standard technique of considering the present value of the entire operational cost and simply adding it to the capital cost. One line of future work would lie in considering alternative objective functions. In particular, because operational cost is often the cause of schemes becoming obsolete, it might be desirable to consider only the operational cost as the minimization objective and constrain the capital cost so that it does not exceed current government norms.

Scheduling and Operation: Currently the demands in the network are assumed to be fixed and are simultaneously satisfied. More complicated and dynamic demand patterns could be studied. By including operational considerations the demands need not be simultaneously satisfied but could be staggered throughout the day.

EPANET Integration: Currently once the network components have been determined by JalTantra, the network details are entered into EPANET to verify the solution. The EPANET hydraulic library could be integrated into JalTantra to ease this process and provide the user with a unified experience.

Integrating more design steps: We started with just pipe diameter selection and iteratively automated more steps from the design process and included them in JalTantra. This process could continue further. For example, currently the demands are calculated using population forecasting and then fed into the system. This could be integrated into JalTantra itself, while also leveraging GIS technology to map the demand more accurately.

Cost Allocation: A common source of scheme failure is the refusal of payments by users due to a mismatch in the perceived benefits and payment demanded. Currently all users are assumed equal, when in reality there is great variance in the quality of service received in the case of multi village schemes. Analysing this disparity and allocating cost accordingly is an interesting problem in the design of multi village piped water schemes.

Appendices

Appendix I

Complete ILP Model

Drinking water distribution networks consist of various components. To optimize the cost of such networks, several inputs must be considered, and for each component several parameters must be determined. We first explicitly formulate the problem that we are attempting to solve and then provide details of the ILP model used to solve the problem.

I.1 Problem Formulation

Inputs:

- General: primary/secondary supply hours, minimum/ maximum headloss per km, maximum water speed
- Source node: head
- Node: elevation, water demand, minimum pressure requirement
- Link: start/end node, length
- Existing Pipes: start/end node, length, diameter, parallel allowed, roughness
- Commercial Pipes: diameter, roughness, cost per unit length
- Tanks: maximum tank heights, tank capacity factor, nodes that must/must not have tanks, capital cost table
- Pumps: minimum pump size, efficiency, design lifetime, capital/energy cost, discount/interest rate, pipes that cannot have pumps
- Valves: location, pressure rating

Outputs:

- Length and diameter of pipe segments for each link
- Partitioning the set of links into primary and secondary network
- Location, height and size of Tanks
- Set of nodes being served by each Tank
- Location and power of Pumps

Objective:

- Minimize total capital cost (pipe + tank + pump) and total energy cost (pump)

Constraints:

- Pressure at each node must be at least the minimum pressure specified
- Water demand must be met at each node

I.2 Model Details

The pipe diameter selection in the model is represented by the continuous variable l_{ij} which represents the length of the j^{th} pipe diameter component of the i^{th} link in the network. This determines the capital cost of the pipes. The tank allocation is represented by the binary variable s_{nm} which is one if the tank at the n^{th} node in the network provides water to the m^{th} node in the network. The choice of tank allocation variables fixes the total demand that each tank serves i.e. the variable d_n. This in turn determines the capital cost of the tanks. Apart from the cost considerations, each node n must also have its minimum pressure constraint satisfied. The head at each node, h_n is dependent on the headloss hl_i in the links of the network. This headloss depends on the pipe variables l_{ij} and the tank variables s_{nm} mentioned earlier. In addition, the introduction of pumps/valves increases/decreases the headloss respectively. The details of the parameters, variables, objective function and constraints of the model are as follows:

Parameters:

- NL: Number of links in the network
- NP: Number of commercial pipe diameters
- D_j: Diameter of j^{th} commercial pipe diameter

- C_j: Cost per unit length of j^{th} commercial pipe diameter
- NN: Number of nodes in the network
- NE: Number of rows in the tank cost table
- B_k: Base cost of the k^{th} row of the tank cost table
- UN_k: Unit cost of the k^{th} row of the tank cost table
- UP_k: Upper limit capacity for the k^{th} row of the tank cost table
- LO_k: Lower limit capacity for the k^{th} row of the tank cost table
- CP: Capital cost of pumps per unit kW
- EP: Energy cost of pumps per unit kWh
- DF: Discount factor for the energy cost over the entire scheme lifetime
- PH: Number of hours of water supply in the primary network
- SH: Number of hours of water supply in the secondary network
- Y: Lifetime of scheme in years
- $INFR$: Inflation rate
- $INTR$: Interest rate
- L_i: Length of the i^{th} link
- PR_n: Minimum pressure required at node n
- E_n: Elevation of the n^{th} node
- DE_n: Water demand of the n^{th} node
- DE: The total water demand of the network
- VH_i: Head reduction by valve in i^{th} link
- HL^p_{ij}: Headloss for the j^{th} diameter of the i^{th} link, if i is part of the primary network
- HL^s_{ij}: Headloss for the j^{th} diameter of the i^{th} link, if i is part of the secondary network
- FL^p_i: Flow in i^{th} link if i is part of the primary network

- FL_i^s: Flow in i^{th} link if i is part of the secondary network
- R_j: Roughness of j^{th} commercial pipe diameter
- T_{min}: Minimum tank height allowed
- T_{max}: Maximum tank height allowed
- ρ: Density of water
- g: Acceleration due to gravity
- η: Efficiency of pump
- PP_{min}: Minimum pump power allowed
- PP_{max}: Maximum pump power allowed
- A_n: Set of nodes that are ancestors of node n
- D_n: Set of nodes that are descendants of node n
- C_n: Set of child nodes of node n
- P_n: Parent node of node n
- I_n: Incoming link for node n
- O_n: Set of outgoing links from node n

Continuous Variables:

- l_{ij}: Length of the j^{th} pipe component of the i^{th} link
- l_{ij}^p: Length of the j^{th} pipe component of the i^{th} link, if link i is part of the primary network
- l_{ij}^s: Length of the j^{th} pipe component of the i^{th} link, if link i is part of the secondary network
- hl_i: Total headloss across link i
- d_n: Total demand served by tank at node n
- z_{nk}: Total demand served by tank at node n, if costed by the k^{th} row of the tank cost table

- p_i: Power of pump installed at link i
- p_i^p: Power of pump installed at link i, if link i is part of primary network
- p_i^s: Power of pump installed at link i, if link i is part of secondary network
- ph_i: Head provided by pump at link i
- h_n: Water head at node n
- t_n: Height of tank at node n
- h'_{ni}: Effective head provided to link i by its starting node n

Binary Variables:

- e_{nk}: 1 if tank at n^{th} node is costed by the k^{th} row of tank cost table, 0 otherwise
- f_i: 1 if link i is part of the primary network, 0 if part of the secondary network
- es_{ni}: 1 if source of water for link i is its immediate upstream node n, 0 otherwise
- s_{nm}: 1 if tank at node n is source for node m, 0 otherwise
- pe_i: 1 if a pump is installed at link i, 0 otherwise

Objective Function: The objective function is simply the sum of capital cost of the pipes, tanks, pumps and valves used in the network. In addition, we also have the operational cost of the pumps. This operational cost is computed as the present value of the total cost over the scheme lifetime.

$$\begin{aligned} O(.) = \quad & \sum_{i \in S_{\sim E}}^{NL} \sum_{j=1}^{NP} C_{ij}(D_{ij}) l_{ij} \quad + \quad \sum_{i \in S_E}^{NL} \sum_{j=1}^{NP} L_i C_{ij}(D_{ij}) p_{ij} \\ + \quad & \sum_{n=1}^{NN} \sum_{k=1}^{NE} e_{nk} * (B_k + UN_k * (d_n - LO_k)) \\ + \quad & \sum_{i=1}^{NL} CP * p_i \quad + \quad EP * DF * \left(\sum_{i=1}^{NL} PH * p_i^p + \sum_{i=1}^{NL} SH * p_i^s \right) , \end{aligned}$$

$$where \quad DF = \quad \sum_{n=1}^{Y} \left(\frac{1 + INFR}{1 + INTR} \right)^{n-1} .$$

Constraints:

- The length of each link segment l_{ij}, is the sum of the primary and secondary components, i.e. l^p_{ij} and l^s_{ij} respectively:

$$l_{ij} = l^p_{ij} + l^s_{ij}, \quad i = 1, \ldots, NL, \quad j = 1, \ldots, NP.$$

- For a given link i, either all l^p_{ij} are 0 or all l^s_{ij} are 0, depending on the value of f_i. And the sum of the non-zero components must equal the length of the link L_i.

$$\sum_{j=1}^{NP} l^p_{ij} = L_i f_i, \quad i = 1, \ldots, NL,$$
$$\sum_{j=1}^{NP} l^s_{ij} = L_i(1 - f_i), \quad i = 1, \ldots, NL.$$

- The pressure at each node must exceed the minimum pressure required:

$$PR_n \leq h_n - (E_n + t_n), \quad n = 1, \ldots, NN \ .$$

- Across every link i there is headloss hl_i. This headloss depends on the flow, length and diameter of the pipe chosen. We use the Hazen-Williams equation [74] to calculate the headloss. The headloss across a link also depends on the pump and valve installed across it, if any. The valves are input parameters to the model since they are manually fixed. The constraints related to the pump head ph_i are described further below. The flow through the link depends on whether the link is part of the primary or secondary network:

$$\begin{aligned}
hl_i =& \sum_{j=1}^{NP}(HL^p_{ij} l^p_{ij} + HL^s_{ij} l^s_{ij}) - ph_i + VH_i, && i = 1, \ldots, NL,\\
HL^p_{ij} =& \frac{10.68 * \left(\frac{FL^p_i}{R_j}\right)^{1.852}}{D_j^{4.87}}, && i = 1, \ldots, NL, \quad j = 1, \ldots, NP,\\
HL^s_{ij} =& \frac{10.68 * \left(\frac{FL^s_i}{R_j}\right)^{1.852}}{D_j^{4.87}}, && i = 1, \ldots, NL, \quad j = 1, \ldots, NP,\\
FL^s_i =& FL^p_i * \frac{PH}{SH}, && i = 1, \ldots, NL.
\end{aligned}$$

- The head h_n at each node n is calculated by the effective head h'_{mi} provided by its parent node m and the headloss hl_i across the link connecting two nodes. The effective head in turn depends on whether the link i has the tank at the starting node m as its source. This is represented by the binary variable es_{mi}:

$$h_n = h'_{mi} - hl_i, \quad n = 1, \ldots, NN, \quad m = P_n, \quad i = I_n,$$

$$
\begin{aligned}
h'_{mi} = &\ (t_m + E_m) * es_{mi} + h_m * (1 - es_{mi}), && m = 1, \ldots, NN, \quad i \in O_m, \\
es_{mi} = &\ s_{mm} * (1 - f_i), && m = 1, \ldots, NN, \quad i \in O_m.
\end{aligned}
$$

- Next, we look at the constraints related to the tank allocation. The first tank constraint is to ensure that every tank height is between parameters T_{min} and T_{max}.

$$T_{min} \leq t_n \leq T_{max} \ .$$

- We then look at the constraints that model allocation of demand nodes to tanks. s_{nm} is 1 if tank at node n serves the demand of node m. A node n serves its child m if and only if it serves all the nodes downstream of m:

$$s_{nm} = s_{nk}, \quad n = 1, \ldots, NN, \quad m \in C_n, \quad k \in D_m \ .$$

- For every node n, only one upstream node m can serve its demand.

$$\sum_{m \in A_n \cup \{n\}} s_{mn} = 1, \quad n = 1, \ldots, NN \ .$$

- The total demand d_n served by node n is the sum of the demands of the downstream nodes that it serves i.e. all m such that $s_{nm} = 1$.

$$d_n = \sum_{m \in D_n \cup \{n\}} s_{nm} * DE_m, \quad n = 1, \ldots, NN \ .$$

- For a node n, its incoming pipe i will have primary flow only if the node serves itself.

$$f_i = s_{nn}, \quad n = 1, \ldots, NN, \quad i = I_n \ .$$

- Next, we have the constraints that relate the demand that a tank serves to its cost variables e_{nk}. Note that we require z_{nk} in our objective function to replace the non linear term $e_{nk} \times d_n$. The following two inequalities of the model provide the bounds for z_{nk} in terms of e_{nk} and the minimum(LO_k) and maximum(UP_k) capacities for each row of the cost table:

$$
\begin{aligned}
LO_k e_{nk} \leq z_{nk}, \quad n = 1, \ldots, NN, \quad k = 1, \ldots, NE, \\
z_{nk} \leq UP_k e_{nk}, \quad n = 1, \ldots, NN, \quad k = 1, \ldots, NE.
\end{aligned}
$$

- Since every tank can be costed using exactly one row of the table, the sum of e_{nk} for a given n must be 1:

$$\sum_{k=1}^{NE} e_{nk} = 1, \quad n = 1, \ldots, NN \ .$$

- Next, we have a similar equation but this time related to the variable z_{nk}. The sum of all z_{nk} values for a given tank must equal d_n:

$$\sum_{k=1}^{NE} z_{nk} = d_n, \quad n = 1, \ldots, NN.$$

- Next, we look at constraints related to pumps. The pump power p_i relates to the pump head ph_i in the following way:

$$\begin{aligned}
p_i &= \quad p_i^p + p_i^s, && i = 1, \ldots, NL,\\
p_i^p &= \frac{(\rho * g * FL_i^p * ph_i)}{\eta} * f_i, && i = 1, \ldots, NL,\\
p_i^s &= \frac{(\rho * g * FL_i^s * ph_i)}{\eta} * (1 - f_i), && i = 1, \ldots, NL\ .
\end{aligned}$$

- Finally, the pump power for each pump must lie between minimum and maximum allowed pump power. This is implemented using the binary variable pe_i.

$$PP_{min} * pe_i \leq p_i \leq PP_{max} * pe_i, \quad i = 1, \ldots, NL\ .$$

Appendix II

JalTantra Usage Details

The web interface of JalTantra is split into two panels as can be seen in Figure 9.3. The sidebar panel on the left enables the user to switch through the various tabs of the system. Clicking on any of them fills the main panel on the right with the relevant content. We go through each of the tabs available and briefly discuss their functionality below. For each tab we list the inputs that the user may provide. The format for each input is:

{Name of the Input}:{Variable Type} {Unit of Input} {Mandatory?} {Definition}

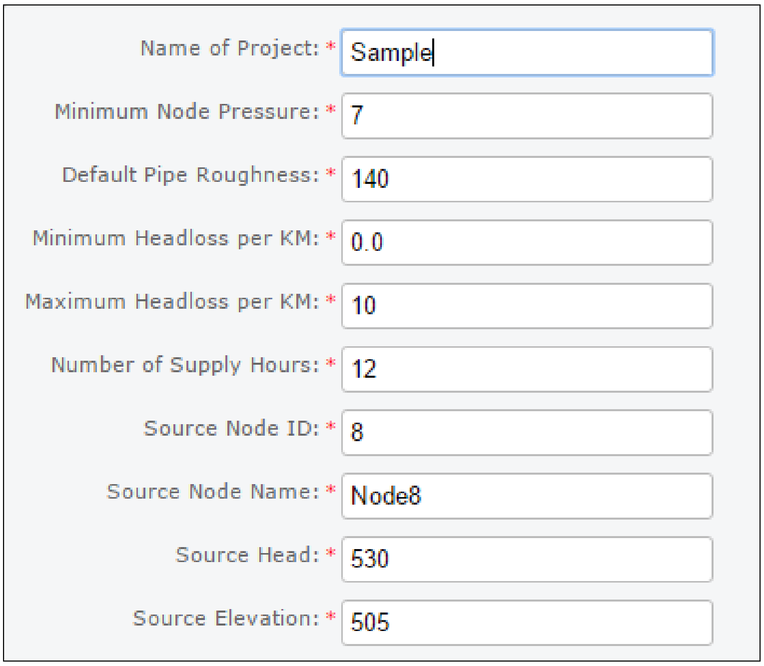


Figure II.1: General Tab of JalTantra

General Tab: Figure II.1 displays the content of the general tab. As the name suggests, this tab is used to enter the general information about the network. The fields to be entered are:

- Name of Project: Text (Your Project Name)

- Minimum Node Pressure: Double (metre) (The default minimum pressure that must be maintained at all nodes)

- Default Pipe Roughness: Double (The default pipe roughness that is used to calculate headloss in the pipes)

- Minimum Headloss per KM: Double (metre) (The minimum headloss per km that is allowed in each pipe)

- Maximum Headloss per KM: Double (metre) (The maximum headloss per km that is allowed in each pipe)

- Maximum Speed of Water: Double (metre per second) (not mandatory, default:no constraint) (Maximum speed of water in metres per second that is allowed in a pipe)

- Maximum Pressure in Pipe: Double (metre) (not mandatory, default:no constraint) (Maximum pressure of water in metres that should exist in a pipe. Note that this constraint is not strictly enforced, but a warning is provided in the results if pressure is exceeded)

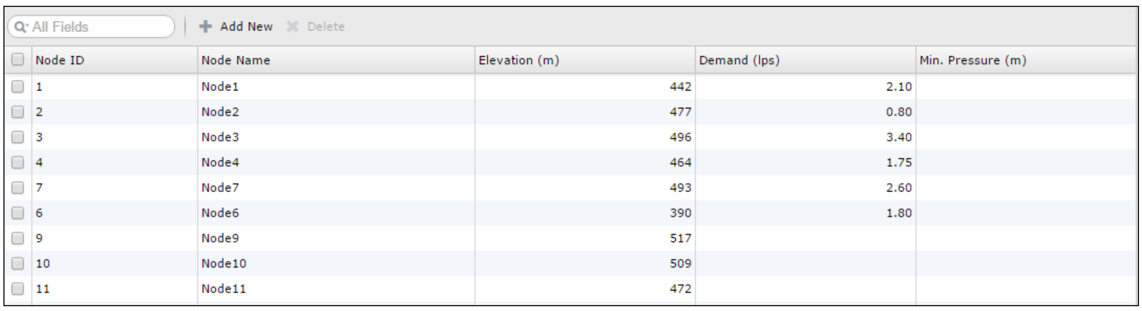
- Number of Supply Hours: Double (Number of hours in a day that water is supplied for. For example if supply hours is 12 it corresponds to a peak factor of 2)

- Source Node ID: Integer (The unique node id of the source)

- Source Node Name: String (The name of the source node)

- Source Head: Double (metre) (The constant water head provided by the source)

- Source Elevation: Double (metre) (The elevation of the source)



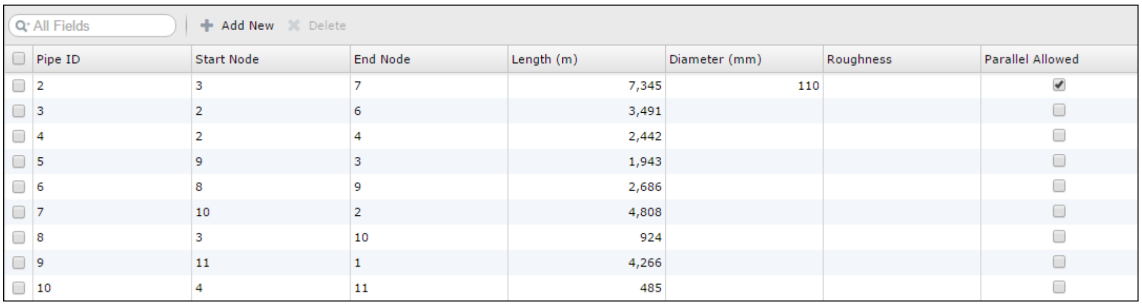
All Fields | Add New | Delete

Node ID	Node Name	Elevation (m)	Demand (lps)	Min. Pressure (m)
1	Node1	442	2.10	
2	Node2	477	0.80	
3	Node3	496	3.40	
4	Node4	464	1.75	
7	Node7	493	2.60	
6	Node6	390	1.80	
9	Node9	517		
10	Node10	509		
11	Node11	472		

Figure II.2: Nodes Tab of JalTantra

Nodes Tab: Figure II.2 displays the content of the node tab. This tab is used to enter the information about the nodes in the network. The fields to be entered are:

- NodeID: Integer (unique nodeID of your node)
- Node Name: String (Name of your node)
- Elevation: Double (metre) (elevation of your node)
- Demand: Double (litres per second) (not mandatory, default:0) (demand of your node)
- Min. Pressure: Double (metre) (not mandatory, default: from general tab) (minimum pressure that needs to maintained at the node)
- Add New: (Add extra row corresponding to one node)
- Delete: (Remove selected rows)



All Fields | Add New | Delete

Pipe ID	Start Node	End Node	Length (m)	Diameter (mm)	Roughness	Parallel Allowed
2	3	7	7,345	110		☑
3	2	6	3,491			☐
4	2	4	2,442			☐
5	9	3	1,943			☐
6	8	9	2,686			☐
7	10	2	4,808			☐
8	3	10	924			☐
9	11	1	4,266			☐
10	4	11	485			☐

Figure II.3: Pipes Tab of JalTantra

Pipes Tab: Figure II.3 displays the content of the pipe tab. This tab is used to enter the information about the pipes in the network. The fields to be entered are:

- PipeID: Integer (unique pipeID of your pipe)
- Start Node: Integer (nodeID of the node at the start of your pipe)
- End Node: Integer (nodeID of the node at the end of your pipe)
- Length: Double (metre) (length of your pipe)
- Diameter: Integer (millimeter) (not mandatory, default:to be calculated) (diameter of your pipe)
- Roughness: Double (not mandatory, default:from general tab) (roughness of pipe that is used to calculate the headloss in the pipe)
- Parallel Allowed: Boolean (not mandatory, default: false) (If a new pipe is allowed to be placed in parallel to an existing pipe.)

- Add New: (Add extra row corresponding to one pipe)
- Delete: (Remove selected rows)

All Fields   Add New   Delete

Diameter (mm)	Roughness	Cost (Rs)
63		116
75		172
90		231
110		340
125		461
140		576
160		750
180		945
200		1,113
225		1,430
250		1,762
280		2,210
315		2,794

Figure II.4: Commercial Pipes Tab of JalTantra

Commercial Pipes Tab: Figure II.4 displays the content of the Commercial Pipes tab. This tab is used to enter the information about the discrete set of commercially available pipe diameters in the network. The fields to be entered are:

- Diameter: Integer (millimeter) (unique diameter of the commercial pipe)
- Roughness: Double (not mandatory, default:from general tab) (Roughness of the commercial pipe)
- Cost: Double (Rs per metre) (cost per metre of the commercial pipe)
- Add New: (Add extra row corresponding to one commercial pipe)
- Delete: (Remove selected rows)

ESR Tab: Figure II.5 displays the content of the ESR tab. This tab is used to enter the information about the ESRs in the network. The fields to be entered are:

- Secondary Network Supply Hours : Double (Number of hours in day that an ESR provides water to its secondary network)
- ESR Capacity Factor: Double (Size of ESR in relation to the daily demand it serves. For e.g. a value of 0.5 means that the ESR capacity is half the daily demand)

- Maximum ESR Height: Double (metre) (not mandatory, default:no constraint) (Maximum height of ESR in metres)
- Allow ESRs at zero demand nodes: Boolean (not mandatory, default:false) (Allow ESRs to placed at nodes with zero demand. Note that if zero demand nodes are not allowed optimization will be significantly faster)
- Nodes with ESRs: (not mandatory, default:no constraint) (List of Nodes that must have ESRs)
- Nodes without ESRs: (not mandatory, default:no constraint) (List of Nodes that must not have ESRs)

ESR Cost Tab: Within the ESR Tab, there is a separate ESR cost tab as shown in Figure II.6. This tab is used to enter the information about the cost of ESRs. The fields to be entered are:

- Minimum Capacity: Double (litre) (default: max capacity of previous row) (Minimum capacity of the row in the ESR cost table)
- Maximum Capacity: Double (litre) (Maximum capacity of the row in the ESR cost table)
- Base Cost: Double (Rs) (default: calculated from previous row) (base cost of ESR having capacity between the minimum and maximum capacity)
- Unit Cost: Double (Rs per litre) (unit cost of ESR having capacity between the minimum and maximum capacity)
- Add New: (Add extra row for the ESR cost table)

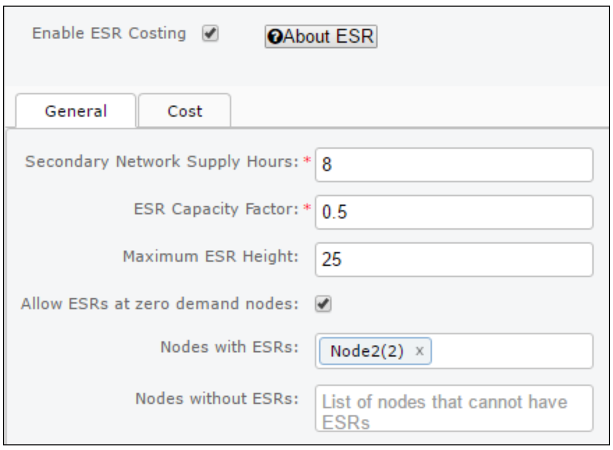


Figure II.5: ESR Tab of JalTantra

Enable ESR Costing ☑ About ESR

General | Cost

All Fields | Add New | Delete

Minimum Capacity (l)	Maximum Capacity (l)	Base Cost (Rs)	Unit Cost (Rs)
0	25,000	0	24.47
25,000	50,000	611,750	12.96
50,000	75,000	935,750	9.64
75,000	100,000	1,176,750	8.64
100,000	150,000	1,392,750	7.23
150,000	200,000	1,754,250	6.03
200,000	300,000	2,055,750	5.40
300,000	400,000	2,595,750	5.12
400,000	1,500,000	3,107,750	4.32
1,500,000	2,000,000	7,859,750	3.92

Figure II.6: ESR cost Tab of JalTantra

- Delete: (Remove selected rows)

Pump Tab: Figure II.7 displays the content of the Pump tab. This tab is used to enter the information about the pumps in the network. The fields to be entered are:

- Minimum Pump Size : Double (kW) (Size of Minimum Pump in kW that can be installed)
- Pump Efficiency: Double (Efficiency of pump expressed as a %)
- Capital Cost per kW: Double (Rs) (Capital cost of the pump per kW installed capacity)
- Energy Cost per kWh: Double (Rs) (Energy cost per kWh energy consumed)
- Design Lifetime: Integer (Number of years for which pumps will be operational)
- Discount Rate: Double (not mandatory, default: 0) (Discount rate is the interest rate expressed as a %. More the discount rate lesser is the energy cost of the pump)
- Inflation Rate: Double (not mandatory, default: 0) (Inflation rate is the rate by which prices rise expressed as a %. More the inflation rate greater is the energy cost of the pump)

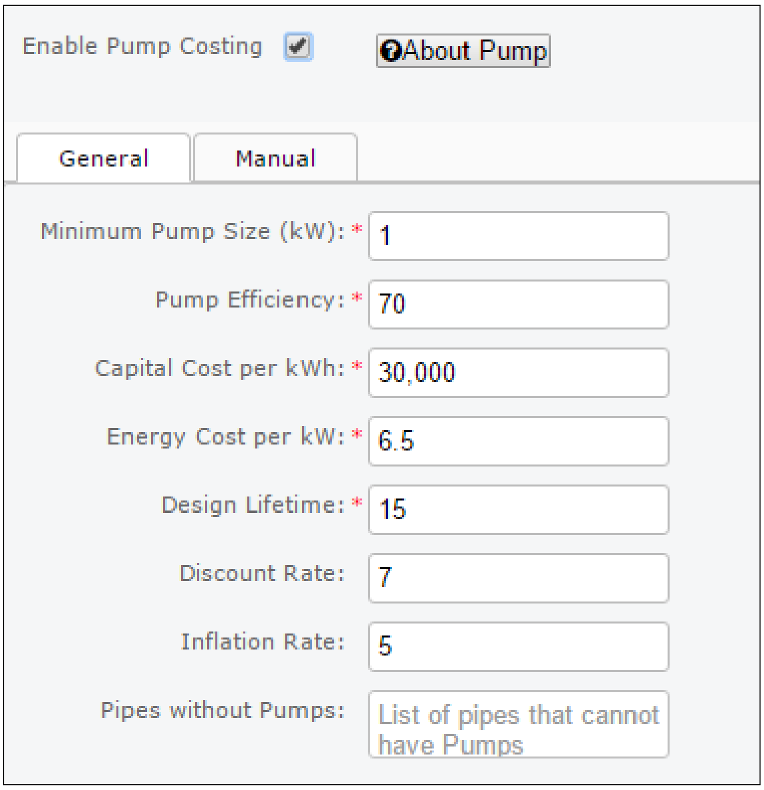


Figure II.7: Pump Tab of JalTantra

- Pipes without Pumps: (not mandatory, default:no constraint) (List of pipes that cannot have Pumps)

Nodes Pipes Cost

All Fields

Node ID	Node Name	Demand (lps)	Elevation (m)	Head (m)	Pressure (m)	Min. Pressure (m)
1	Node1	4.20	442.00	455.28	13.28	7.00
2	Node2	1.60	477.00	491.16	14.16	7.00
3	Node3	6.80	496.00	518.03	22.03	7.00
4	Node4	3.50	464.00	480.01	16.01	7.00
6	Node6	3.60	390.00	476.58	86.58	7.00
7	Node7	5.20	493.00	502.49	9.49	7.00
8	Node8	0.00	505.00	530.00	25.00	7.00
9	Node9	0.00	517.00	524.00	7.00	7.00
10	Node10	0.00	509.00	516.00	7.00	7.00
11	Node11	0.00	472.00	479.00	7.00	7.00

Figure II.8: Node Results Tab of JalTantra

Results Tab: Figure II.8 displays the content of the Results tab. This tab displays the results of the optimization. The results are broken down into details in various subtabs:

- Node Tab: Nodewise results including head and pressure values for each node

- Pipe Tab: Pipewise results including flow, diameter, headloss, headlossperkm and cost for each pipe

- Cost Tab: Cost results for each diameter of commercial pipe
- ESR Tab: Cost results for each ESR to be constructed
- Pump Tab: Cost results for each Pump to be installed

Map Tab: The available options available in the Map tab are summarized below:

- Transfer Data: Transfer the node and pipe information to the nodes and pipes tab
- Close Chart: Deselect the currently selected pipe and close its associated elevation chart
- Search Location: Search for a location on the map (type in the location and select from the dropdown menu or press enter. Can also enter lat, long information)
- Add Node: Add a node to the map (Click on the button then click on a point on the map. Once added you can modify the node name and id and also change location by either entering the lat/long info or manually moving the node. You can also set if this node is the source node or an ESR node)
- Add Pipe: Add a pipe between two nodes on the map (Click on the button then click on two existing nodes on the map one by one)
- Right click the map: Provides options to add/edit/delete nodes, add/delete/split pipes or close the elevation chart
- Right click a node: Provides options to delete or edit the node. (Deleting the node removes all pipes connected to it)
- Right click a pipe: Provides options to delete/split a pipe or close the elevation chart. (Splitting a pipe adds a node at the split point and creates two pipes instead of the original one)

Appendix III

Karegaon Scheme Redesign

III.1 Background

Located at about 120 km from Mumbai on its North East side, Mokhada Taluka, is a block in Palghar district of Maharashtra state. In spite of high rainfall in the range of 3000 mm to 4000 mm, there are currently around 30 villages (70 habitations) in the Taluka that are perpetually dependent on tanker supplied water during pre-summer and summer months. Irony is that this tribal dominated area with a hilly terrain also has the distinction of being the biggest supplier of drinking water to Mumbai city. It hosts two big reservoirs namely, Upper Vaitarna and Middle Vaitarna, supplying over 1000 million litres of water per day (MLD). To augment water supply to Mumbai by 455 MLD, the Middle Vaitarna project on Vaitarna River in this Taluka was recently commissioned. The construction of the dam on Middle Vaitarna submerged the source of the Karegaon Rural Water Supply Scheme which supplies drinking water to four villages besides Karegaon village. When Karegaon scheme was revamped because of submergence of its assets in the backwater of the dam, the people in the neighbouring water-scarce villages were upset over the fact that they were not included in the scope of redesigned scheme. They do not object to their water being taken away as long as their need of drinking water is addressed.

III.2 CTARA Proposal for Redesign of Karegaon Scheme

The coverage area of revamped Karegaon scheme was restricted to five villages, namely, Karegaon, Kaduchiwadi, Kochale, Bhasmyachiwadi, and Karegaon Ashramshala. In 2013, a study was undertaken by CTARA, IIT Bombay [33] to evaluate the feasibility of augmentation of scope of Karegaon scheme to include a cluster of 13 tanker fed villages in its neighbourhood. The primary objective of the study was evaluation of techno economic feasibility of a multi village water supply scheme (MVS) to supply drinking water to the cluster of tanker fed villages in the neighbourhood of Karegaon scheme in Mokhada Taluka. A step by step process following guidelines and protocols used by Maharashtra

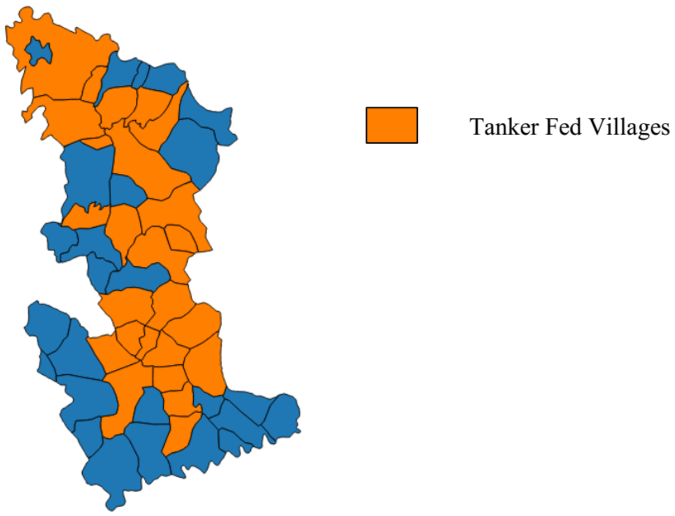


Figure III.1: Tanker Fed Villages in Mokhada Taluka

Jeevan Pradhikaran (MJP) in their design process was used for this purpose. The secondary objective was to understand the process thoroughly and identify where and how the process could be improved.

III.3 Scheme Design Process and Design Options

III.3.1 Source Selection

There were two alternative sources under consideration, Middle Vaitarna reservoir as per current assumption of Karegaon scheme. The other alternative was to consider Upper Vaitarna reservoir as source. The Full Supply Level of middle Vaitarna is 285 m while the same for Upper Vaitarna is 603 meters. Hence the latter offers considerable advantage over the former due to its high elevation in terms of savings not only in capital cost but also in energy costs. Earlier studies indicated that the scheme based on Middle Vaitarna as source is not feasible. This is because several villages are at higher elevations (average elevation: 360m) than the former source (285m elevation). Hence, Upper Vaitarna was chosen as the source.

III.3.2 Population and Demand Forecast

Population for the year 2001 and 2011 was got from census data [12]. Using this data, projections for the year 2030 were made using the incremental and geometric methods (assuming construction in year 2015 and a 15 year design). To get our projection we then take the average of the two methods. This gives us a total projected population of 48407 for the year 2030.

Incremental method:

Projected population = Current population + decadal growth * (no. of years/10)

Geometrical method:

$$\text{Projected population} = \text{Current population} * (1 + \text{decadal growth rate})^{(\text{no. of years}/10)}$$

Drinking water demand was estimated from the projected population at 40 lpcd (litres per capita per day). This gives a total demand of 1.94 million litres per day (MLD) for the 17 villages. A further 20% loss factor was added to the demand. This gives us a total supply requirement of 2.32 MLD. Village wise population and demand is detailed in table III.1 below.

Table III.1: Details of the tanker fed villages in Mokhada Taluka

Village Name	**Population (2011)**	**Population (2041 est.)**	**Demand (litres per day)**	**Elevation (m)**
Kiniste	939	2957	1,18,280	460
Udhale	1064	2281	91,240	422
Jogalwadi	812	2245	89,800	426
Khodala	2807	7721	3,08,840	434
Sayade	1770	3006	1,20,240	420
Gomghar	1228	1536	61,440	419
Shirasgaon	526	1454	58,160	243
Dolhare	1141	1838	73,520	380
Nashera	733	4179	1,67,160	354
Adoshi	923	1060	42,400	202
Dhamanshet	1241	3431	1,37,240	384
Palsunde	1365	3774	1,50,960	393
Pathardi	661	4153	1,66,120	192
Kochale	609	1684	67,360	345
Ashramshala	750	2073	82,920	350
Karegaon	1196	3306	1,32,240	350
Kaduchiwadi	618	1709	68,360	350
Total	**18,381**	**48,407**	**19,36,280**	

III.3.3 WTP and MBR

Water Treatment Plant is generally designed for a capacity of total daily demand. This includes the demand calculation along with any losses. The cost estimation for the WTP was done by extrapolating from the WTP cost of the Khardi scheme [69]. In the Khardi scheme a WTP of capacity 1 MLD costs Rs 26 lakhs. For us the daily demand is 2.3 MLD. Accordingly we priced our WTP at 26 * 2.3 * 1.1 = Rs 66 lakhs. The 1.1 factor is to account for an inflation of 10% in prices. The MBR was designed with a capacity of one third the daily demand i.e. 0.77 MLD. The cost estimation for the MBR was done from the MJP schedule of rates. This gave us a cost of about Rs 30 lakhs.

III.3.4 Pumping Machinery and Rising Main

Water has to be pumped at two places: from Source to WTP and from WTP to MBR. The lowest draw level of source is 595m. The elevation of WTP and MBR is 631m and 640m respectively. The diameter of the rising mains is calculated using the flow of water required and assuming a velocity of 1.25 m/s.

$$
\begin{aligned}
flow &= \frac{dailydemand}{time} \\
\Rightarrow flow &= \frac{2.3MLD}{12hours} =\sim 53.8\ lps \\
dia &= 2 * \sqrt{\frac{flow}{\pi * vel}} \\
\Rightarrow dia &= 2 * \sqrt{\frac{53.8\ lps}{\pi * 1.25\ m/s}} =\sim 234mm
\end{aligned}
$$

We round up this value and use a diameter of 250mm.

Raw water rising main: The distance is of 500m and diameter required is 250mm. Due to the high flow we choose D.I. pipe for the rising main. This gives us a cost of $\sim$ Rs 13.3 lakhs for the rising main. Pure water rising main: The distance is of 650m and the diameter is 250mm as before. We again use D.I. pipe which gives us a cost of $\sim$ Rs 17.3 lakhs.

Pumping machinery: Pump capacity required is calculated using the flow rate of water that needs to be pumped and the head difference. For raw water rising main it is 53.8 lps and head difference of 50m. Assuming an efficiency of 70% we get a power requirement of $\sim$ 50hp. Similarly for the pure water rising main, we have a flow of 53.8 lps and head difference of 25m which gives us a pump power requirement of 25hp. Cost for the two pumps is calculated by comparing it with the cost of a 100hp pump in Khardi scheme which was 22 lakhs. So our 50hp and 25hp pumps cost 11 lakhs and 5.5 lakhs respectively.

III.3.5 ESRs

Water is distributed to all the ESRs in the network by gravity. All the villages are marked on Google Earth. ESR is generally designed for 50% of the daily demand serviced by the ESR. The staging height is varied until an optimum point is reached. The cost of ESR goes up along with increase in staging height but it is compensated by reduction in pipe size of the secondary network due to availability of higher head and corresponding reduction in cost. Similarly for a higher height the primary network needs to provide a higher head and thus pipes with larger diameter are required which increases the primary network cost. In the present study this optimization is not considered and all the cost estimations are done based on the staging height of 10m. The choice of ESR location crucially decides the cost of the network. This is because fixing the ESR, fixes the primary and secondary networks. Several ESR configurations must be considered. The cost of ESRs rises sub linearly with increasing capacity. For example a 1 lakh litre capacity ESR costs Rs 13.12 lakhs (at 2011-12 prices) and a 2 lakh litre capacity ESR costs Rs 18.87 lakhs. Therefore to minimize ESR cost the optimum option is to have one big ESR for the entire network. But with additional number of ESRs the piping cost goes down. So the optimum ESR configuration can go anywhere from one to the number of villages in the network. In the present study we place an ESR at each village. But this may not be always true since it depends on the actual pipe network and its configuration. The ESR capacity and cost for each village are detailed in table III.2 below.

Table III.2: Details of the Elevated Storage Reservoirs in the Mokhada Scheme

Village Name	**Population (2041 est.)**	**Demand (l per day)**	**Elevation (m)**	**ESR Capacity (l)**	**Cost of ESR (Rs)**
Kiniste	2957	1,18,280	460	71000	11,54,340
Udhale	2281	91,240	422	55000	10,01,700
Jogalwadi	2245	89,800	426	54000	9,92,160
Khodala	7721	3,08,840	434	186000	19,22,840
Sayade	3006	1,20,240	420	73000	11,73,420
Gomghar	1536	61,440	419	37000	7,61,080
Shirasgaon	1454	58,160	243	35000	7,31,400
Dolhare	1838	73,520	380	45000	8,79,800
Nashera	4179	1,67,160	354	101000	13,97,875
Adoshi	1060	42,400	202	26000	5,97,840
Dhamanshet	3431	1,37,240	384	83000	12,56,100
Palsunde	3774	1,50,960	393	91000	13,19,700
Pathardi	4153	1,66,120	192	100000	13,91,250
Kochale	1684	67,360	345	41000	8,20,440
Ashramshala	2073	82,920	350	50000	9,54,000
Karegaon	3306	1,32,240	350	80000	12,32,250
Kaduchiwadi	1709	68,360	350	42000	8,35,280
Total	**48407**	**19,36,280**		**11,70,000**	**1,84,21,475**

III.3.6 Primary Distribution Network

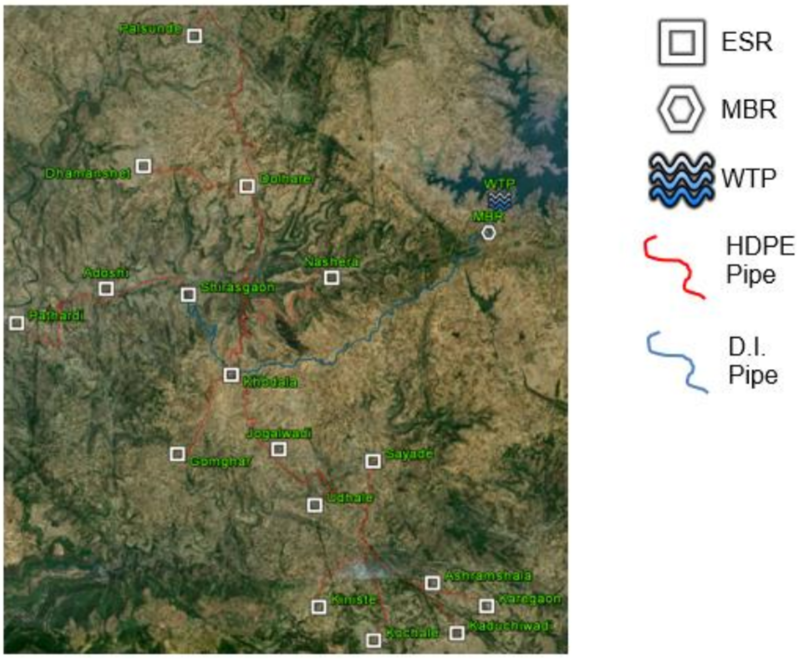


Figure III.2: Network Layout of the Mokhada Scheme

The pipe network is marked on Google Earth connecting the MBR to the ESRs. The pipes are laid out on the existing road network. We can now extract the distance details from Google Earth. Dummy nodes are marked to take care of changing elevation along paths. This is because water not only has to reach the end point but also it has to meet the minimum head requirement of 7m at the highest point along the path. Now one of the most important considerations is what should be the diameter of the pipe that needs to be laid out so that sufficient head is maintained at all points of the network. Smaller the diameter, smaller is the cost but greater the headloss which may cause points in the network to not receive water. To determine the pipe diameters we use a software BRANCH [53]. Given the pipe network and the elevation and demand of the nodes, it gives us the diameters of the pipes that need to be used to satisfy the demand while maintaining the head requirements. The choice of pipe quality depends on the maximum pressure that the pipe is under throughout its length. Greater the pressure, greater the thickness of the pipe required and more the cost of the pipe. If pressure is high and one can do with lower pressures then pressure reducing valves or break-pressure reservoirs can be utilized. This would lower downstream pressures resulting in lower pipe costs. Pipes in the network are HDPE pipes rated to withstand 80m of head. Where higher head is required we use D.I. K-9 pipes which can withstand upto 400m of head (marked * in table III.3 below). The final layout of WTP, MBR, ESRs and pipes is shown in figure III.2.

Table III.3: Details of the Pipe Network in the Mokhada Scheme

Sr. No.	From Node	To Node	Peak Flow (lps)	Dia (mm)	Headloss (m)	HL/km (m)	Length (m)	Cost (1000 Rs)
1*	MBR	Khodala	53.76	200	59.48	12.95	4595	8523.73
				225	38.78	7.29	5316	9861.18
2	Khodala	3	21.39	160	21.84	6.98	3130	1993.81
3	3	2	18.90	140	12.09	10.64	1137	557.04
				160	8.07	5.55	1455	926.96
4	2	Udhale	2.53	63	7.26	12.58	577	56.55
5	2	1	16.37	140	7.74	8.15	950	465.5
6	3	Jogalwadi	2.49	63	3.47	12.22	284	27.83
7	Khodala	21	1.70	63	3.77	6.02	626	61.35
8	21	Gomghar	1.70	63	14.65	6.03	2430	238.14
9*	Khodala	Shirasgaon	7.41	90	60.60	16.17	3747	3900.63
10	Shirasgaon	Adoshi	5.79	75	66.61	24.90	2675	390.55
11	Adoshi	25	4.61	75	40.83	16.33	2500	365.00
12	25	Pathardi	4.61	75	32.67	16.33	2000	292.00
13	Khodala	4	14.68	125	25.23	11.57	1181	462.95
*			14.68	125	25.23	11.57	1000	1594.68
14	4	22	4.64	75	77.69	16.53	4700	686.20
15	22	Nashera	4.64	75	13.03	16.54	788	115.05
16	4	Dolhare	10.04	110	59.44	10.68	5568	1609.15
17	Dolhare	23	3.81	75	23.08	11.48	2010	293.46
18	23	Dhamanshet	3.81	75	16.68	11.48	1453	212.14
19	Dolhare	24	4.19	90	11.27	5.64	2000	392.00
20	24	26	4.19	90	17.80	5.63	3160	619.36
21	26	Palsunde	4.19	75	16.61	13.69	1214	177.20
				90	1.39	5.64	246	48.28
22	1	Sayade	3.34	63	29.26	21.04	1391	136.32
23	1	29	13.03	140	5.93	5.34	1110	543.90
24	29	30	5.16	110	1.87	3.12	600	173.40
25	30	Kiniste	3.29	90	7.20	3.60	2000	392.00
26	30	32	1.87	63	15.11	7.20	2100	205.80
27	32	Kochale	1.87	63	1.80	7.20	250	24.50
28	29	31	7.87	90	36.16	18.08	2000	392.00
29	31	Ashramshala	2.30	63	4.22	10.55	400	39.20
30	31	Karegaon	3.67	75	21.42	10.71	2000	292.00
31	31	Kaduchiwadi	1.90	63	13.33	7.41	1800	176.40
						Total	**68393**	**36246.26**

*Links where D.I. K-9 pipes are used to withstand upto 400m of pressure.

III.3.7 Verification of Network using EPANET

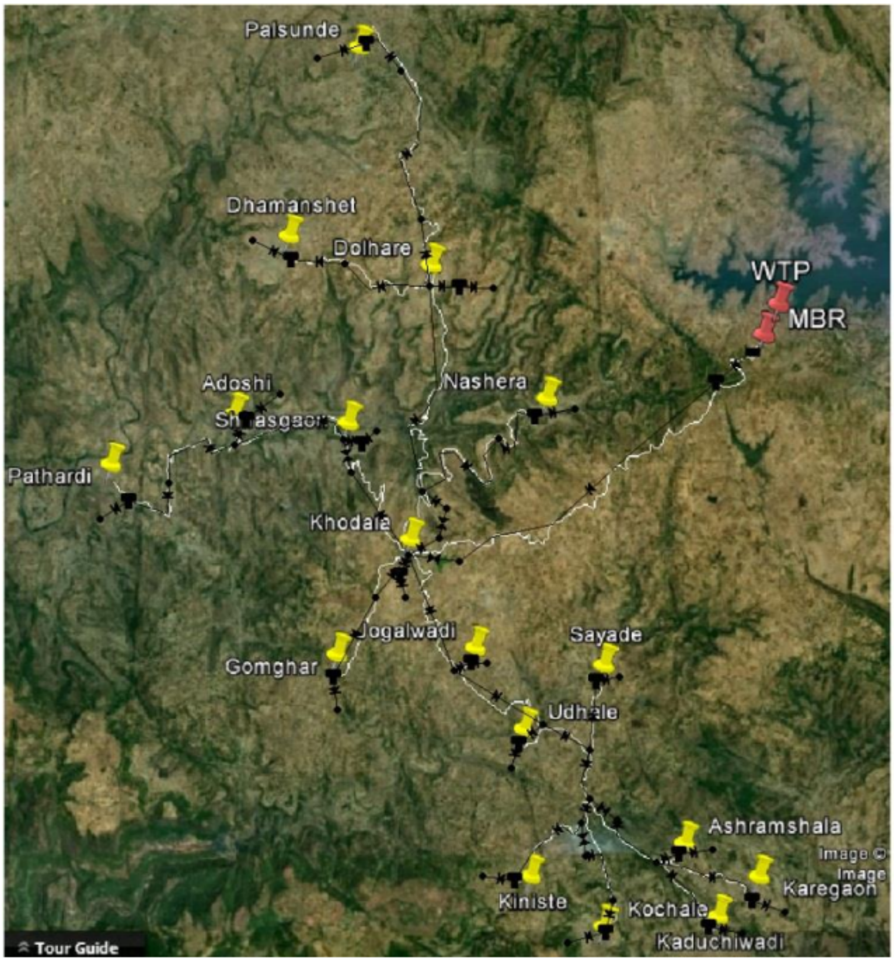


Figure III.3: EPANET model of the Mokhada Scheme Network

The network design was verified by using EPANET [68]. EPANET is a software tool that models the flow of water in pressurized piped networks. After completing the sizing and locations of the pipes and ESRs we construct the network in EPANET to verify whether sufficient head is being realized at all nodes. EPANET allows analysing how the various ESRs in the network fill up and empty during the daily life cycle. This helps indicates if there are any "problem" nodes where sufficient head is not being met. As can be seen in figure III.4, upstream villages get filled up first. Khodala in particular being the first village in the network gets filled up within the first hour. Other villages gradually fill up and empty during the demand period which is hours 9-12 and 21-24. Note that supply occurs at hours 1-6 and 13-18.

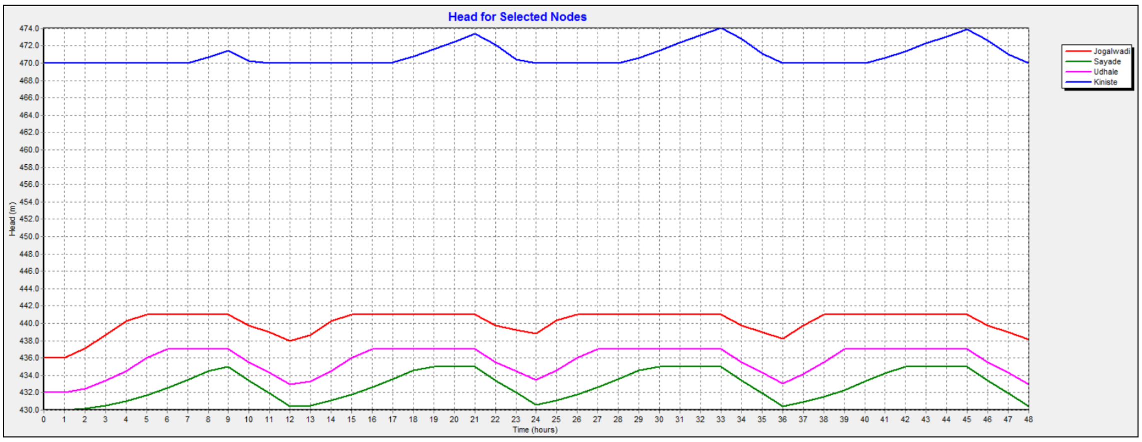


Figure III.4: Water Heads at various nodes over time in the Mokhada Scheme

III.3.8 Capital Cost

The proposed multi village scheme for supplying piped water to 17 from Upper Vaitarna has an estimated capital cost of Rs. 13.99 crores. This amounts to a per capita cost of Rs. 2890 per for a design population of 48407. The cost breakup across various components is detailed in table III.4.

Table III.4: Details of the Capital cost of the Mokhada Scheme

Sr. No.	**Cost Component**	**Base Cost (Rs)**	**Misc. Factor**	**Net Cost (Rs)**
1	Jack Well	25,00,000	1	25,00,000
2	WTP	66,45,313	1	66,45,313
3	Raw water rising main (Length 500m, Dia. 250mm)	13,28,405	1.479	19,64,711
4	Pure water rising main (Length 650m, Dia. 250mm)	17,26,927	1.379	23,81,432
5	MBR	30,59,408	1.151	35,21,379
6	Cost of raw water pump (50HP)	11,00,000	3.185	35,03,500
7	Cost of pure water pump (25HP)	5,50,000	2.652	14,58,600
8	Excavation	2,05,17,900	1.273	2,61,19,287
9	Piping	3,86,99,749	1.273	4,92,64,781
10	ESRs	2,02,63,623	1.142	2,31,41,057
11	Tertiary Network	1,59,06,008	1	1,59,06,008
12	M.S.E.B.	20,00,000	1	20,00,000
13	Land Acquisition	15,00,000	1	15,00,000
14	Total Cost			13,99,06,066
15	Cost per capita			2890.20

References

[1] Alperovits, Elyahu, and Uri Shamir. "Design of optimal water distribution systems." Water resources research 13.6 (1977): 885-900.

[2] Assembly, G. (2015). Sustainable Development Goals. (SDGs), Transforming our world: the, 2030.

[3] Babayan, A., Kapelan, Z., Savic, D., and Walters, G. (2005) Least-cost design of water distribution networks under demand uncertainty. Journal of Water Resources Planning and Management, 131(5):375–382.

[4] Bentley (2019) Water distribution and analysis design software. Accessed August 28, 2019, https://www.bentley.com/en/products/product-line/hydraulics-and-hydrology-software/watergems.

[5] Bhaduri, A., Bogardi, J., Siddiqi, A., Voigt, H., Vörösmarty, C., Pahl-Wostl, C., and Kremer, H. (2016). Achieving sustainable development goals from a water perspective. Frontiers in Environmental Science, 4, 64.

[6] Bhave, P. R. (1979) Selecting pipe sizes in network optimization by LP. Journal of the Hydraulics Division, 105(7):1019–1025.

[7] Bhave, P. R. (1981). "Node flow analysis of water distribution systems." J. Transp. Eng., 107(4), 457–467.

[8] Bhave, P. R. and Lam, C. F. (1983) Optimal layout for branching distribution networks. Journal of Transportation Engineering, 109(4):534–547.

[9] Berkelaar, M., Eikland, K., and Notebaert, P. (2006). lpsolve, version 5.5. Eindhoven University of Technology, http://sourceforge. net/projects/lpsolve.

[10] Bouchart, F. and Goulter, I. (1991) Reliability improvements in design of water distribution networks recognizing valve location. Water Resources Research, 27(12):3029–3040.

[11] Calhoun, C. A. (1971) Optimization of pipe systems by linear programming. In Tullis, J. P., editor, Control of Flow in Closed Conduits, pages 175–192. Colorado State University, Ft. Collins.

[12] Chandramouli, C., and Registrar General. "Census of India 2011." Provisional Population Totals. New Delhi: Government of India (2011).

[13] Chase, D. V., and Ormsbee, L. E. (1989). "Optimal Pump Operation of Water Distribution Systems with Multiple Storage Tanks." Proc. Conf. on Water Resources Planning and Management, ASCE, New York, 733-736.

[14] da Conceicao Cunha, Maria, and Luisa Ribeiro. "Tabu search algorithms for water network optimization." European Journal of Operational Research 157.3 (2004): 746-758.

[15] Cunha, M. D. C. and Sousa, J. (1999) Water distribution network design optimization: Simulated annealing approach. Journal of Water Resources Planning and Management, 125(4):215–221.

[16] Dandy, G. C., Simpson, A. R., and Murphy, L. J. (1996) An improved genetic algorithm for pipe network optimization. Water Resources Research, 32(2):449–458.

[17] Dagre javascript library,(Jan, 2017), https://github.com/cpettitt/dagre/wiki

[18] De Corte, Annelies, and Kenneth Sörensen. "HydroGen: an artificial water distribution network generator." Water resources management 28.2 (2014): 333-350.

[19] Duan, N., Mays, L. W., and Lansey, K. E. (1990) Optimal reliability-based design of pumping and distribution systems. Journal of Hydraulic Engineering, 116(2):249–268.

[20] Eiger, G., Shamir, U., and Ben-Tal, A. (1994) Optimal design of water distribution networks. Water Resources Research, 30(9):2637–2646.

[21] Eusuff, Muzaffar M., and Kevin E. Lansey. "Optimization of water distribution network design using the shuffled frog leaping algorithm." Journal of Water Resources Planning and Management 129.3 (2003): 210-225.

[22] Forrest J, Lougee-Heimer R (2014) CBC user guide. Accessed July 15, 2019, https://pubsonline.informs.org/doi/pdf/10.1287/educ.1053.0020.

[23] Fujiwara, O. and Dey, D. (1988) Method for optimal design of branched networks on flat terrain. Journal of Environmental Engineering, 114(6):1464–1475.

[24] Fujiwara, Okitsugu, and Debashis Dey. "Two adjacent pipe diameters at the optimal solution in the water distribution network models." Water Resources Research 23.8 (1987): 1457-1460.

[25] Fujiwara, O., Jenchamahakoon, B., and Edirisinghe, N. C. P. (1987) A modified linear programming gradient method for optimal design of looped water distribution networks. Water Resources Research, 23(6):977–982.

[26] GLPK, (GNU Linear Programming Kit), (Accessed April, 2015), https://www.gnu.org/software/glpk/

[27] Google Maps Platform (2017) Maps JavaScript API: Overview. Accessed January 3, 2017, https://developers.google.com/maps/documentation/javascript/.

[28] Google OR Tools, (Jan, 2017), https://developers.google.com/optimization/

[29] Goulter, I. C. and Bouchart, F. (1990) Reliabilty-constrained pipe network model. Journal of Hydraulic Engineering, 116(2):211–229.

[30] Goulter, I. C. and Coals, A. V. (1986) Quantitative approaches to reliability assessment in pipe networks. Journal of Transportation Engineering, 112(3):287–301.

[31] Gupta, R., & Bhave, P. R. (1994). Reliability analysis of water-distribution systems. Journal of Environmental Engineering, 120(2), 447-461.

[32] Halhal, D., Walters, G. A., Ouazar, D., & Savic, D. A. (1997). Water network rehabilitation with structured messy genetic algorithm. Journal of water resources planning and management, 123(3), 137-146.

[33] Hooda, Nikhil, Desai, Rajaram, and Damani Om. "Design and Optimization of Piped Water Network for Tanker Fed Villages in Mokhada Taluka" Technical Report No. TR-CSE-2013e-55, 2013, Dept. of Computer Science and Engineering, IIT Bombay.

[34] Jacoby, S. L. S. (1968) Design of optimal hydraulic networks. Journal of the Hydraulics Division, 94(3):641–661.

[35] Java 7, (Sep, 2014), http://java.com/en/download/index.jsp

[36] Java ILP – Java interface to ILP solvers, (April, 2015), http://javailp.sourceforge.net/

[37] jQuery JavaScript Library. https://jquery.com/

[38] Karmeli, D., Gadish, Y., and Meyers, S. (1968) Design of optimal water distribution networks. Journal of Pipeline Division, 94(1):1–10, 1968.

[39] Karmarkar, Narendra. "A new polynomial-time algorithm for linear programming." Proceedings of the sixteenth annual ACM symposium on Theory of Computing. ACM, 1984.

[40] Lansey, Kevin E., and Larry W. Mays. "Optimization model for water distribution system design." Journal of Hydraulic Engineering 115.10 (1989): 1401-1418.

[41] Lansey, K. E., Duan, N., Mays, L. W., and Tung, Y. K. (1989) Water distribution system design under uncertainties. Journal of Water Resources Planning and Management, 115(5):630–645.

[42] Lejano, Raul P. "Optimizing the layout and design of branched pipeline water distribution systems." Irrigation and Drainage Systems 20.1 (2006): 125-137.

[43] Mahmoud, H. A., Savić, D., and Kapelan, Z. (2017). New pressure-driven approach for modeling water distribution networks. Journal of Water Resources Planning and Management, 143(8), 04017031.

[44] Maier, H. R., Simpson, A. R., Zecchin, A. C., Foong, W. K., Phang, K. Y., Seaha, H. Y., and Tan, C. L. (2003) Ant colony optimization for design of water distribution systems. Journal of Water Resources Planning and Management, 129(3):200–209.

[45] Maier, H. R., Kapelan, Z., Kasprzyk, J., Kollat, J., Matott, L. S., Cunha, M. C., Dandy, G. C., Gibbs, M. S., Keedwell, E., Marchi, A., Ostfeld, A., Savic, D., Solomatine, D. P., Vrugt, J. A., Zecchin, A. C., Minsker, B. S., Barbour, E. J., Kuczera, G., Pasha, F., Castelletti, A., Giuliani, M., and Reed, P. M. (2014). "Evolutionary Algorithms and Other Metaheuristics in Water Resources: Current Status, Research Challenges and Future Directions." Environmental Modelling and Software, 62, 271-299.

[46] Mala Jetmarova, Helena, Nargiz Sultanova, and Dragan Savic. "Lost in optimisation of water distribution systems? A literature review of system operation." Environmental Modelling and Software 93 (2017): 209-254.

[47] Mala-Jetmarova, Helena, Nargiz Sultanova, and Dragan Savic. "Lost in optimisation of water distribution systems? A literature review of system design." Water 10.3 (2018): 307.

[48] Mandry, J. E. (1967) Design of pipe distribution systems for sprinkler projects. Journal of Environmental Engineering, 93(3):243–257.

[49] Maharashtra Jeevan Pradhikaran (MJP), Schedule of Rates for Konkan Region: Civil Works for Year 2018-19, (2018).

[50] Montesinos, P., Garcia-Guzman, A., and Ayuso, J. L. (1999) Water distribution network optimization using a modified genetic algorithm. Water Resources Research, 35(11):3467–3473.

[51] Nicklow, J., Reed, P., Savic, D., Dessalegne, T., Harrell, L., Chan-Hilton, A., Karamouz, M., Minsker, B., Ostfeld, A., Singh, A., and Zechman, E. (2010). "State of the Art for Genetic Algorithms and Beyond in Water Resources Planning and Management." Journal of Water Resources Planning and Management, ASCE, 136(4), 412-432.

[52] Papadimitriou, Christos H. "On the complexity of integer programming." Journal of the ACM (JACM) 28.4 (1981): 765-768.

[53] Prasad M. Modak, Juzer Dhoonia, "A Computer Program in Quick BASIC for the Least Cost Design of Branched Water Distribution Networks." UNDP/ World Bank, Asia Water Supply and Sanitation Sector Development Project, RAS/86/160, December 1991.

[54] Robinson, R. B. and Austin, T. A. (1976) Cost optimization of rural water systems. Journal of the Hydraulics Division, 102(8):1119–1134.

[55] Samani, Hossein MV, and Alireza Mottaghi. "Optimization of water distribution networks using integer linear programming." Journal of Hydraulic Engineering 132.5 (2006): 501-509.

[56] Sankar, G. S., Kumar, S. M., Narasimhan, S., Narasimhan, S., and Bhallamudi, S. M. (2015). Optimal control of water distribution networks with storage facilities. Journal of Process Control, 32, 127-137.

[57] Savić, D., and Ferrari, G. (2014). Design and performance of district metering areas in water distribution systems. Procedia Engineering, 89, 1136-1143.

[58] Savic, Dragan A., and Godfrey A. Walters. "Genetic algorithms for least-cost design of water distribution networks." Journal of Water Resources Planning and Management 123.2 (1997): 67-77.

[59] Sayyed, M. A., Gupta, R., and Tanyimboh, T. T. (2014). "Modelling pressure deficient water distribution networks in EPANET." Procedia Eng., 89, 626–631.

[60] Schaake, J.C.; Lai, D. Linear Programming and Dynamic Programming Application to Water Distribution Network Design; Report No. 116; Hydrodynamics Laboratory, Department of Civil Engineering, Massachusetts Institute of Technology: Cambridge, MA, USA, 1969.

[61] Schwarz, J., Meidad, N., and Shamir, U. (1985). "Water Quality Management in Regional Systems." Scientific Basis for Water Resources Management, IAHS Publ. no. 153, Proceedings of the Jerusalem Symposium, September 1985.

[62] Shamir, U. (1974) Optimal design and operation of water distribution systems. Water Resources Research, 10(1):27–36, 1974.

[63] Simpson, A. R., Dandy, G. C., and Murphy, L. J. (1994) Genetic algorithms compared to other techniques for pipe optimization. Journal of Water Resources Planning and Management, 120(4):423–443.

[64] Su, Y., Mays, L. W., Duan, N., and Lansey, K. E. (1987) Reliability-based optimization model for water distribution systems. Journal of Hydraulic Engineering, 114(12):1539–1556.

[65] Swamee, P. K., Kumar, V., and Khanna, P. (1973) Optimization of dead end water distribution system. Journal of Environmental Engineering Division, 99(2):123–134.

[66] Thornton, Julian, R. Sturm, and G. Kunkel. Water loss control manual. New York: McGraw-Hill, 2002.

[67] Tuttle, G. W. "The economic velocity of transmission of water through pipes." Engineering Record 32.15 (1895): 258-258.

[68] United States Environmental Protection Agency (2017) EPANET: Application for modeling drinking water distribution systems. Accessed June 1, 2019, http://www.epa.gov/nrmrl/wswrd/dw/epanet.html.

[69] V. Choudhary, O. Damani, R. Desai, A. Joshi, M. Kanwat, M. Kaushal, S. Kharole ,Kharole, Y. Pawde, P. Rathore, and M. Sohoni. "Redesigning Khardi Rural Piped Water Network Scheme for Sustainability" Technical Report No. TR-CSE-2013-56, 2013, Dept. of Computer Science and Engineering, IIT Bombay.

[70] Van Zyl, J. E., Savic, D. A., and Walters, G. A. (2004). "Operational Optimization of Water Distribution Systems Using a Hybrid Genetic Algorithm." Journal of Water Resources Planning and Management, ASCE, 130(2), 160-170.

[71] Vyas, Janki H., Narendra J. Shrimali, and Mukesh A. Modi. "Optimization of Dhrafad Regional Water Supply Scheme using Branch 3.0.", IJIRSET 2014

[72] w2ui JavaScript Library. http://w2ui.com/web/

[73] Watanatada, T. (1973) Least-cost design of water distribution systems. Journal of the Hydraulics Division, 99(9):1497–1513.

[74] Williams, Gardner Stewart, and Allen Hazen. "Hydraulic tables." (1933).

[75] Wu, Z. Y. and Simpson, A. R. (2001) Competent genetic-evolutionary optimization of water distribution systems. Journal of Computing in Civil Engineering, 15(2):89–101.

[76] Xu, C. and Goulter, I. C. (1999) Reliabilty-based optimal design of water distribution networks. Journal of Water Resources Planning and Management, 125(6):352–362.

[77] Yates, D. F., A. B. Templeman, and T. B. Boffey. "The computational complexity of the problem of determining least capital cost designs for water supply networks." Engineering Optimization 7.2 (1984): 143-155.

[78] Young, B. (1994) Design of branched-water-supply network on uneven terrain. Journal of Environmental Engineering, 120(4):975–979.

[79] Yugandhara Lad, J S Main, and S D Chawathe. "Optimization of Hydraulic Design of Water Supply Tree Network Based on Present Worth Analysis." Journal of Indian Water Works Association, 2012.

List Of Publications

Published

- Hooda, Nikhil, and Om Damani. ”A system for optimal design of pressure constrained branched piped water networks.”, 18th Conference on Water Distribution System Analysis, WDSA (2016).
- Hooda, Nikhil, and Om Damani. ”Inclusion of tank configurations as a variable in the cost optimization of branched piped-water networks.” Drinking Water Engineering and Science 10.1 (2017): 39-44.

Accepted

- Hooda, Nikhil, and Om Damani, “JalTantra: A System for Design and Optimization of Rural Piped Water Networks”, INFORMS Journal on Applied Analytics
- Hooda, Nikhil, Ashutosh Mahajan, and Om Damani, ”A Series of ILP Models for the Optimization of Water Distribution Networks”, SADHANA

Acknowledgements

I would like to thank my supervisor, Prof. Om P. Damani for his constant guidance, encouragement and patience throughout the duration of my study. Conversations with him were always enlightening, be they on the research topic or about the world in general. I have relied on his advice from before the research work started and I am sure I will continue to do so in the future as well. I thank the members of my research progress committee, Prof. Milind Sohoni, Prof. Ashutosh Mahajan and Prof. Sundar Vishwanathan. Their diverse expertise provided invaluable feedback which helped direct and refine the research work.

Special thanks are in order for Prof. Puru Kulkarni, Raj Desai, Abhishek Sinha, Ashish Karale and Vaishali Bharambe from the Centre for Technology Alternatives for Rural Areas (CTARA). The work done by the team at CTARA was the genesis for this thesis. The students of the Technology and Development Supervised Learning (TDSL) course did the field survey and analysis of existing and planned water schemes, providing us with an invaluable resource for understanding the design process. In particular I thank Raj sir for being the interface between the practical and the theoretical. It was through him that we were able to interact with numerous government engineers and understand the challenges they face. I thank the engineers from Maharashtra Jeevan Pradhikaran, Santosh Gawankar, Mahesh Patil, Brajesh Sengar and Pradeep Gokhale, who listened to us patiently, answered our questions and then provided invaluable feedback on the JalTantra system.

I thank Dipak Chaudhri for his guidance on all things research at IITB. I thank the office staff at Computer Science Department at IIT Bombay for their assistance. In particular, for their patience and prompt help, I thank Vijay Ambre and Sunanda Ghadge.

I thank the Ministry of Human Resource Development (MHRD), Government of India for the PhD assistantship and Tata Consultancy Services (TCS) for supporting me through a research fellowship.

I thank Vijay Arya and Pratyush Kumar from IBM, Bangalore for hosting me during my summer intern.

Most of all, I thank my parents, Surender and Rachna Hooda and my sister Garima Hooda. My research work would not be possible without their constant unconditional love and support. I hope I can live upto their expectations.