

CONSIDERATIONS ON MULTIOBJECTIVE OPTIMIZATION REGARDING
MUNICIPAL WATER SUPPLY SYSTEMS

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY



BY
HAYRETTİN ONUR BEKTAŞ

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF DOCTOR OF PHILOSOPHY
IN
CIVIL ENGINEERING

SEPTEMBER 2020

Approval of the thesis:

**CONSIDERATIONS ON MULTIOBJECTIVE OPTIMIZATION
REGARDING MUNICIPAL WATER SUPPLY SYSTEMS**

submitted by **HAYRETTİN ONUR BEKTAŞ** in partial fulfillment of the requirements for the degree of **Doctor of Philosophy in Civil Engineering, Middle East Technical University** by,

Prof. Dr. Halil Kalıpçılar
Dean, Graduate School of **Natural and Applied Sciences**

Prof. Dr. Ahmet Türer
Head of the Department, **Civil Engineering**

Assoc. Prof. Dr. Nuri Merzi
Supervisor, **Civil Engineering, METU**

Examining Committee Members:

Prof. Dr. Ali Melih Yanmaz
Civil Engineering, METU

Assoc. Prof. Dr. Nuri Merzi
Civil Engineering, METU

Prof. Dr. Elçin Kentel Erdoğan
Civil Engineering, METU

Assoc. Prof. Dr. Yakup Darama
Civil Engineering, Atılım University

Dr. Aslı Numanoğlu Genç
Civil Engineering, TED University

Date: 25.09.2020



I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last name : Hayrettin Onur Bektaş

Signature :

ABSTRACT

CONSIDERATIONS ON MULTIOBJECTIVE OPTIMIZATION REGARDING MUNICIPAL WATER SUPPLY SYSTEMS

Bektaş, Hayrettin Onur
Doctor of Philosophy, Civil Engineering
Supervisor: Assoc. Prof. Dr. Nuri Merzi

September 2020, 250 pages

Optimal water supply system design has always been a challenge for researchers. While most optimization studies focus mainly on pipe diameter sizing, recent works are being done for optimizing other elements of water supply systems.

The design of water supply system can be discussed from either single or multiobjective optimization point of view. Multiobjective optimization does not only consider the costs of the system but also some performance measures such as resilience.

In this study, a multiobjective optimization methodology is developed by using non-dominated sorting genetic algorithms (NSGA-II). Consequently, Pareto trade-off curves for cost vs. network resilience (I_n) are obtained. Different sets of solutions are formed with different population sizes; then performances are measured by hypervolume calculations to determine optimum population size for the system considered. Finally, one solution is selected within the pareto curve by compromise programming to offer a design to the decision maker. Shamir network, Hanoi network and a real distribution network from Ankara are studied as case studies.

It is concluded that constraints play a decisive role on pareto curves in the methodology. When number of constraints are increased, final pareto curve shifts away from the true pareto curve.

Moreover, it is observed that, placing storage tanks in the system have less effect on the network strength when compared to diameter changes.

It is also concluded that a storage tank count limitation is present in Shamir and Hanoi networks. After this case specific limit, it becomes very hard to place another feasible tank to the network and final pareto curves consist of very few points.

Keywords: Multiobjective Optimization, NSGA-II, Network Resilience, Optimum Water Distribution System Design, Ankara

ÖZ

KENTSEL İÇME SUYU SİSTEMLERİNİN ÇOKLU AMAÇ OPTİMİZASYONU ÜZERİNE DEĞERLENDİRMELER

Bektaş, Hayrettin Onur
Doktora, İnşaat Mühendisliği
Tez Yöneticisi: Doç. Dr. Nuri Merzi

Eylül 2020, 250 sayfa

Optimum içme suyu temin sistemleri her zaman araştırmacılar için zorlu bir iş olmuştur. Bu konudaki çalışmaların çoğu boru çapı optimizasyonuna konu edinmişse de yeni çalışmalar içme suyu temin sistemlerinin diğer elemanlarını da işin içine katmaktadır.

İçme suyu temin sistemlerinin optimizasyonu tekli ya da çoklu amaç fonksiyonunun optimizasyonu olarak ikiye ayrılabilir. Çoklu amaç fonksiyonlarının optimizasyonunda, yalnızca maliyet değil, aynı zamanda şebeke performansı da optimize edilir.

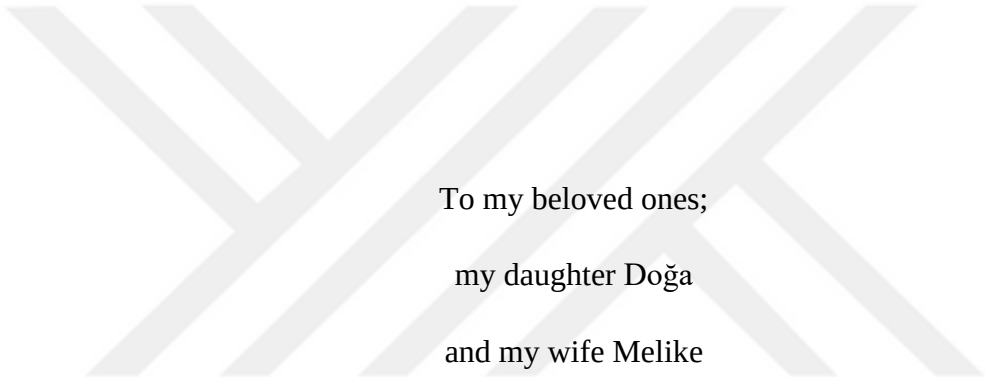
Bu çalışmada, NSGA-II yöntemi kullanılarak bir çoklu amaç optimizasyonu yöntemi geliştirilmiştir. Sonuç olarak maliyet ve şebeke performansı (I_n) arasında pareto eğrileri elde edilmiştir. Çalışılan sistemlerde değişken popülasyon sayıları ile çözümler yaptırılarak pareto eğrilerinin performansları hiper-hacim yöntemi kullanılarak karşılaştırılmış ve her sistem için en uygun popülasyon sayısı belirlenmiştir. En uygun popülasyon ile bulunan pareto eğrisinden uzlaşık programlama yöntemi ile bir tasarım seçilmiştir. Geliştirilen yöntem Shamir sistemi, Hanoi sistemi ve Ankara'dan mevcut bir içme suyu sistemi üzerinde denenmiştir.

Geliştirilen yöntemde kısıtların sonuç pareto eğrilerinin bulunmasında önemli bir rolü olduğu sonucuna varılmıştır. Kısıtların sayısı artırıldıkça, sonuç pareto eğrileri gerçek pareto eğrisinden uzaklaşmaktadır.

Ayrıca sisteme su deposu yerleştirmenin, boru çapı değiştirmeye göre şebeke performansına daha az etkisi olduğu gözlenmiştir.

Ayrıca Shamir ve Hanoi sistemleri göz önüne alındığında, depo sayıları için bir üst sınırın olduğu sonucuna varılmıştır. Sisteme özgü bir sınır olan depo sayılarından sonra, şebekeye olurlu bir depo daha yerleştirmek çok zorlaşmakta, sonuç pareto eğrilerinde çok az sayıda nokta bulunabilmektedir.

Anahtar Kelimeler: Çoklu Amaç Optimizasyonu, NSGA-II, Şebeke Mukavemeti, Optimum Su Dağıtım Şebekeleri Tasarımı, Ankara



To my beloved ones;
my daughter Doğa
and my wife Melike

ACKNOWLEDGMENTS

First thanks go to my supervisor Dr. Nuri Merzi, who gave the very first ideas that started this thesis work. After more than 13 years of cooperation, his supervision transformed into a long joint journey for me. During this journey he enlightened many dark places with his knowledge. I feel really grateful to him.

In the second place, I thank Alter International Engineering and Consulting Services and my dear colleagues in there. They provided a peaceful work environment so that I could generate enough time to complete this work.

I also sincerely thank to thesis monitoring committee members, Dr. Elçin Kentel Erdoğan and Dr. Aslı Numanoğlu Genç. Their innovative ideas encouraged me to perform different trials that I did not think of earlier.

My dear friends Halil Şendil and Dr. Gençer Gençoğlu, you deserve a big appreciation for your earlier studies which I benefitted from so much. After all, I put another brick on top of your previous bricks in a big wall.

Ten years ago, I thanked my extended family for their support in my master's thesis. They get another thank also this time. My dear grandmother Kadriye Sun, I know you are watching me somewhere up there with your pretty smile.

My little princess Doğa, I love you so much. I am sure you will achieve harder goals than me. I will support you as much as I can.

Finally, my love Melike. You gave me so much for the last years. How can I pay for your endless efforts? How can I thank you in a single sentence? Now it's my turn to support you. I feel certain you will succeed in your studies. Thank you.

TABLE OF CONTENTS

ABSTRACT.....	v
ÖZ	vii
ACKNOWLEDGMENTS	x
TABLE OF CONTENTS.....	xi
LIST OF TABLES	xv
LIST OF FIGURES	xvii
LIST OF ABBREVIATIONS	xx
LIST OF SYMBOLS	xxi
CHAPTERS	
1. INTRODUCTION	1
1.1. General	1
1.2. Aim of the Study	3
1.3. Novelty of the Study	3
1.4. Structure of the Thesis	5
2. LITERATURE REVIEW	7
2.1. Optimization Methods	7
2.2. Reliability Measures	10
2.3. Water Distribution Network Parameter Optimization	13
2.4. Constraint Handling	16
2.5. Diversity of Solutions	18
2.6. Performance Measures	22
2.7. Decision Making	24

3. NON-DOMINATED SORTING GENETIC ALGORITHM – II.....	27
3.1. Single vs Multi-Objective Optimizations.....	27
3.2. Domination and Non-Domination Concepts	30
3.3. NSGA-II	33
3.3.1. Fast Non-Dominated Sorting.....	35
3.3.2. Crowded Distance Sorting.....	37
3.3.3. Constrained Tournament Selection	40
3.3.4. Constraint Handling and Constrained Domination	41
3.3.5. Crossover.....	45
3.3.5.1. Binary Crossover	45
3.3.5.2. Simulated Binary Crossover.....	47
3.3.6. Mutation	49
3.3.6.1. Binary Mutation.....	49
3.3.6.2. Real Mutation	50
3.3.7. Pruning and Clustering.....	51
3.4. Performance Determination.....	54
3.5. Decision Making	57
4. OBJECTIVES AND CONSTRAINTS	59
4.1. Resilience	60
4.1.1. Resilience Theory	60
4.1.2. Steady State Simulations vs Extended Period Simulations.....	67
4.1.3. Storage Tank Inclusion.....	68
4.1.4. Example Resilience Calculations	70
4.2. Cost.....	79

4.3. Decision Variables	86
4.4. Constraints	87
4.4.1. Node-Based Constraints.....	91
4.4.2. Pipe-Based Constraints	92
4.4.3. Storage Tank-Based Constraints.....	94
5. METHODOLOGY	101
5.1. Coding.....	101
5.2. Variables	102
5.3. Workflow of the Code.....	107
5.3.1. Initialization	107
5.3.2. Initial Population Formation	107
5.3.3. Iterative NSGA-II Procedure	112
5.3.3.1. Core Mechanism	112
5.3.3.2. Crossover	116
5.3.3.3. Mutation.....	120
5.4. Constraint Handling	123
5.5. Clustering – Hypervolume – Compromise Programming	126
5.6. Stopping Criteria	126
6. CASE STUDIES	127
6.1. 2-Loop Network.....	127
6.1.1. Network Data	127
6.1.2. Implementation	129
6.1.3. Results.....	130
6.1.4. Comparison with Published Works	142

6.2. Hanoi Network	151
6.2.1. Network Data.....	151
6.2.2. Implementation.....	153
6.2.3. Results	154
6.2.4. Comparison with Published Works	165
6.3. Ankara N8.3 Network	170
6.3.1. Network Data.....	170
6.3.2. Implementation.....	176
6.3.3. Results	176
7. CONCLUSIONS	187
7.1. Discussions.....	187
7.2. Suggested Further Studies	190
REFERENCES	193
APPENDICES	205
A. APPENDIX FOR 2L SOLUTIONS	205
B. APPENDIX FOR HAN SOLUTIONS	217
C. APPENDIX FOR N83 SOLUTIONS	238
CURRICULUM VITAE	249

LIST OF TABLES

TABLES

Table 2.1. Volume Comparisons from Countries' Guidelines.....	14
Table 3.1. Sample Values for a Population.....	31
Table 3.2. Sample Values for a Population.....	37
Table 3.3. Sample Elements for Crowding Distance Calculations	39
Table 3.4. Sample Values for Constraint Violations	45
Table 3.5. Sample Binary Coded Pipe Diameters.....	46
Table 3.6. Sample Binary Coded Pipe Diameters.....	46
Table 3.7. Crossover Between Two Members and Resulting Child Members.....	47
Table 3.8. Sample Binary Coded Pipe Diameters.....	50
Table 3.9. Normalized Values of Points used to Create Hypervolume	56
Table 3.10. Normalized Values of Points used for Compromise Programming	58
Table 4.1. Node and Pipe Data for the Sample Network.....	71
Table 4.2. Demand Pattern Coefficients	71
Table 4.3. Node-Pipe Relations and Pipe Diameter Uniformities	71
Table 4.4. 24-Hour EPS Results of the Sample Network	72
Table 4.5. 24-Hour Network Resilience Values of the Sample Network	73
Table 4.6. 24-Hour Network Resilience Values of the Sample Network with Storage	75
Table 4.7. 24-Hour EPS Results of the Sample Network with Storage.....	76
Table 4.8. Unit Pipe Cost Comparison	80
Table 4.9. Tank Costs	80
Table 4.10. Minimum and Maximum Tank Sizes Used in the Study	82
Table 4.11. Reinforced Concrete Storage Tank Costs for 2018	85
Table 4.12. EPS Results of Pipe Parameters for First Modification.....	89
Table 4.13. EPS Results of Node and Tank Parameters for First Modification	90
Table 4.14. EPS Results of Tank Parameters for Second Modification	91

Table 5.1. Summary of variable “pop”	104
Table 5.2. Sample Minimum and Maximum Pipe Diameter Limits	109
Table 6.1. 2-loop Network Characteristics	128
Table 6.2. Summary Results of 2-loop Network	132
Table 6.3. Summary Statistics of 2-loop Network	133
Table 6.4. Pipe Design Values for the Preferred Solutions	135
Table 6.5. Tank Design Values for the Preferred Solutions	135
Table 6.6. Minimum Pressure Comparison	143
Table 6.7. Removed Constraints for Comparison	145
Table 6.8. Hanoi Network Characteristics	152
Table 6.9. Summary Results of Hanoi Network	155
Table 6.10. Summary Statistics of Hanoi Network	156
Table 6.11. Pipe Design Values for the Preferred Solutions	158
Table 6.12. Tank Design Values for the Preferred Solutions	160
Table 6.13. Pipe Diameter Selection Lists for Hanoi Network	166
Table 6.14. Minimum Pressure Comparison	167
Table 6.15. Summary Results of N8.3 Network	177
Table 6.16. Summary Statistics of N8.3 Network	178
Table 6.17. Tank Design Values for the Preferred Solutions	180

LIST OF FIGURES

FIGURES

Figure 2.1. The Two Targets of Multiobjective Optimization.....	19
Figure 2.2. A Sample Plot from a Running Performance Metric	23
Figure 3.1. Constrained Decision Variable Space (Deb 2001)	28
Figure 3.2. Decision Variable and Objective Function Spaces in Multi-Objective Optimizations	29
Figure 3.3. Possible Pareto Fronts in a Bi-Objective Optimization Problem	29
Figure 3.4. A Sample Multi-Objective Function Space	32
Figure 3.5. Solutions Classified into Pareto Optimal Fronts	32
Figure 3.6. Non-dominated Sorting and Crowding Distance Sorting.....	33
Figure 3.7. The Two Goals of Multi-Objective Optimization	34
Figure 3.8. Flowchart of NSGA-II.....	35
Figure 3.9. Crowding Distance Calculation Displayed as Side Lengths of a Cuboid	39
Figure 3.10. Sample Decision Variables with Constraints	43
Figure 3.11. Sample Feasible and Infeasible Objective Function Values	44
Figure 3.12. NSGA-II Steps.....	51
Figure 3.13. Clustering Example Initial Stage.....	53
Figure 3.14. Clustering Example Intermediate Stage	54
Figure 3.15. Clustering Example Final Stage	54
Figure 3.16. Sample Hypervolume Formation.....	56
Figure 3.17. Possible Solutions Compared with the Best Point	58
Figure 4.1. Energy Conservation in a Water Distribution System	61
Figure 4.2. Pipe-Node Relations for C Calculations.....	65
Figure 4.3. A Sample 24-Hour Demand Pattern.....	67
Figure 4.4. A Sample Water Distribution Network Layout.....	70
Figure 4.5. A Sample Water Distribution Network Layout with Storage Tank	74
Figure 4.6. 24-hour Change of Sample I_n values	78

Figure 4.7. 24-hour Change of Sample I_n values and Percent Tank Fullness	78
Figure 4.8. Three Different Types of Water Storage: Reservoirs, Standpipes and Elevated Tanks	81
Figure 4.9. Relative Costs for Different Types of Tanks	82
Figure 4.10. Interpolated Costs for Reservoirs.....	83
Figure 4.11. Interpolated Costs for Standpipes	83
Figure 4.12. Interpolated Costs for Elevated Tanks	84
Figure 4.13. Interpolated 2018 Costs Reinforced Concrete Tanks	85
Figure 5.1. Compiling and Coding Process.....	102
Figure 5.2. Two Main Parts of MATLAB Variable “pop”	103
Figure 5.3. Flowchart for the Complete Methodology.....	105
Figure 5.4. Pipe Parameter Part of the MATLAB Variable “pop”	106
Figure 5.5. Tank Parameter Part of the MATLAB Variable “pop”	106
Figure 5.6. Detailed Flowchart for Initial Population Formation.....	110
Figure 5.7. Detailed Flowchart of Constraint Handling Process.....	115
Figure 5.8. Flowchart Describing Infeasible and Feasible Search Phases	116
Figure 5.9. Sample One Point Binary Crossover for Pipe Diameters	117
Figure 5.10. Sample Crossover for Tank Parameters.....	118
Figure 5.11. Detailed Flowchart for Crossover	119
Figure 5.12. Sample One Point Mutation for Pipe Diameters.....	120
Figure 5.13. Detailed Flowchart for Mutation.....	122
Figure 5.14. Anytown Demand Pattern for EPS Studies.....	125
Figure 6.1. Layout of 2-loop Network.....	128
Figure 6.2. Representations for Final Pareto Curve and Hypervolume Plots	131
Figure 6.3. Pareto Curve for 2L150P1T	134
Figure 6.4. Tank Oscillations for Solution 2L150P1T/29	136
Figure 6.5. Pressure Variations for Solution 2L150P1T/29	137
Figure 6.6. Heat Map of Tank Placements for 2-Loop Network.....	138
Figure 6.7. Member Properties for 2L1T	139
Figure 6.8. Crossover Improvement Example Between Two Iterations	141

Figure 6.9. Comparison with True Pareto Curve	144
Figure 6.10. Tank Fullness Graph of Seven Different Constrained Solutions	147
Figure 6.11. Comparison of Envelope of Solutions with the True Pareto Curve .	148
Figure 6.12. Comparison of Constraint Effects to the True Pareto Curve.....	150
Figure 6.13. Layout of Hanoi Network.....	151
Figure 6.14. Pareto Curve for HAN150P1T	157
Figure 6.15. Tank Oscillations for Solution.....	161
Figure 6.16. Pressure Variations for Solution.....	162
Figure 6.17. Heat Map of Tank Placements for Hanoi Network	163
Figure 6.18. Member Properties for HAN1T.....	164
Figure 6.19. Comparison with Best Estimate Pareto Curve	168
Figure 6.20. Comparison of Constraint Effects to the True Pareto Curve.....	169
Figure 6.21. N8.3 System as a Part of İvedik System.....	172
Figure 6.22. Present Facilities of N8.3 System.....	173
Figure 6.23. Topography of N8.3 Study Area	174
Figure 6.24. Daily Demand Pattern of N8.3 Network	175
Figure 6.25. Pareto Curve for N83250P1T	179
Figure 6.26. Tank Oscillations for Solution N83250P1T/250	181
Figure 6.27. Pressure Variations for Solution N83250P1T/250	182
Figure 6.28. Pressure Distribution for N83250P1T/250 at 12:00	183
Figure 6.29. Pipe Diameters for N83250P1T/250	184
Figure 6.30. Tank Sites for 10 Obtained Solutions.....	185
Figure 6.31. Member Properties for N831T	186

LIST OF ABBREVIATIONS

ABBREVIATIONS

AWWA	American Water Works Association
C-NSGA-II	Clustered Non-Dominated Sorting Genetic Algorithm-II
CV	Constraint Violation
CVCON	Constraint Violation for Constant Water Levels
CVEV	Constraint Violation for Entering Volumes
CVHIV	Constraint Violation for High Velocities
CVHLG	Constraint Violation for Headloss Gradient
CVLOV	Constraint Violation for Low Velocities
CVLSE	Constraint Violation for Level Start-End
CVLV	Constraint Violation for Leaving Volumes
CVOSC	Constraint Violation for Tank Oscillations
CVPRES	Constraint Violation for Pressures
CVVD	Constraint Violation for Volume Differences
DMA	District Metered Area
dll	Dynamic link library
EPS	Extended Period Simulation
HGL	Hydraulic Grade Line
KanGAL	Kanpur Genetic Algorithms Laboratory
MODE	Multi-Objective Differential Evolution
MOPSO	Multi-Objective Particle Swarm Optimization
NPGA	Niched Pareto genetic Algorithm
NSGA	Non-Dominated Sorting Genetic Algorithm
NSGA-II	Non-Dominated Sorting Genetic Algorithm-II
NSGA-III	Non-Dominated Sorting Genetic Algorithm-III
OCV	Overall Constraint Violation
PSO	Particle Swarm Optimization
SA	Simulated Annealing
SMGA	Structured Messy Genetic Algorithm
u-NSGA-III	Unified Non-Dominated Sorting Genetic Algorithm-III
USEPA	United States Environmental Protection Agency
VEGA	Vector Evaluated Genetic Algorithm

LIST OF SYMBOLS

SYMBOLS

<i>con</i>	Constant water level hours
<i>cyc</i>	Allowed tank fill-empty cycle count
<i>D</i>	Tank diameter (m)
<i>E</i>	Tank elevation (m)
<i>ev</i>	Entering volume
<i>H</i>	Tank height (m)
h_i	Head at node <i>i</i>
h_i^*	Head required at node <i>i</i>
H_k	Reservoir head
I_n	Network Resilience
I_r	Resilience Index
kWh	Kilowatt-hour
<i>le</i>	Tank final level
<i>ls</i>	Tank initial level
<i>lv</i>	Leaving volume
MaxIt	Maximum number of iterations
<i>nj</i>	Number of junctions
<i>nn</i>	Number of nodes
<i>non</i>	Number of node-hours having low pressure
<i>np</i>	Number of pipes
nPop	Number of members in a population
<i>npu</i>	Number of pumps
n_{pd}	Number of pipe diameter choices
<i>nr</i>	Number of reservoirs
<i>nt</i>	Number of tanks
<i>osc</i>	Tank fill-empty oscillation count
<i>p</i>	Pressure (m)
<i>pdv</i>	Predefined tank volume
P_{ext}	External power
P_{int}	Internal power
P_{max}	Maximum internal power
P_{tot}	Total power
q_i^*	Demand of member <i>i</i>
Q_k	Reservoir outflow
X_j	Weighted surplus power of node <i>j</i>
X_{max}	Maximum surplus power
z_i^*	Elevation of member <i>i</i>
α	Random value between 0 and 1
γ	Specific weight (N/m ³)
ρ	Density (kg/m ³)
σ	Mutation step size

CHAPTER 1

INTRODUCTION

1.1. General

Municipal water distribution systems are composed of various elements to store, transmit and distribute water from source to consumers. To provide water to consumers, above mentioned elements should be planned, designed, built and operated.

The design of water distribution system elements is performed according to either national or international specifications. These specifications define various requirements that should be satisfied. Those requirements may be minimum / maximum pressures in the system, allowable construction materials, limitations to the dimensions (pipe diameters, water storage tank dimensions etc.).

Designing water distribution systems according to design specifications provides the designer a wide variety of different choices; however, most of these choices are not economically feasible. The designer usually uses past engineering experience to decide on a solution that suits his / her needs. However, there may be cheaper alternatives and these alternatives may not be easily seen by pure engineering judgements.

On the other hand, above discussed design choices can be evaluated from another point of view: strength of the designed system. Strength of a water distribution system can be defined by its durability against failure events. A failure event can be anything that will interrupt water supply to consumers. If a system is capable of handling failures without interrupting water supply, then it is a strong system.

Consequently, two questions arise:

- How to express strength of a water distribution system?
- How to increase strength of a system?

Answers to these questions lie in the design process of the system. It is possible to implement a system strength measuring technique depending on the design variables, and by providing alternative designs it is possible to increase or decrease the strength of the system.

When different design choices for distribution systems are thought of, two main distinctions are possible to make; branched and looped networks. Branched networks are cheaper but vulnerable to failures; on the other hand, looped networks will be more expensive having redundant pipes to form loops but will be stronger against failures. From another aspect, water distribution systems having storage facilities will be stronger in case of a water shortage as storage facilities will keep on supplying water for some time. It may even be possible to design stronger systems by including more storage facilities in the system.

At this point, above two concepts should be examined together: How to create a cheap and strong system that does not violate design specifications and provide uninterrupted water supply to end-users? It can be seen that creating stronger designs will result in more expensive solutions. The two main goals discussed above have inverse effects on each other. This leads to a decisive action on how much sacrifice from one goal to achieve more on the other goal? Consequently, a two-objective problem has been defined. The solution to this problem gives many results as a nature of multiobjective problems. It may be concluded that, having many solutions instead of a single solution is advantageous as having alternative solutions creates a flexibility for the decision maker. As the designer is most probably bounded by monetary limitations, he / she would need to make some sacrifices in the designs.

When it is possible to create many trade-off solutions for a problem, it is most welcome to include as many decision variables as possible to increase diversity among solutions. In a water distribution system, including pipe and storage tank parameters as decision variables will create much more solutions when only pipe parameters are considered. Therefore, it is a good idea to optimize pipe and storage tank parameters within the same optimization study.

This thesis study is shaped around concepts introduced above and results with a methodology that finds solutions to optimal water distribution system designs.

1.2. Aim of the Study

Aim of this study is to develop a two-objective optimization methodology for finding optimum designs for water distribution systems. The two conflicting objectives are cost and strength of the water system to be designed. Cost function is based on various construction cost values. Strength of the network is based on a resilience concept. Realization of this two-objective optimization study is possible with numerous constraints. Optimization problem is solved through a widely known evolutionary multiobjective optimization method, Non-dominated Genetic Algorithm-II (NSGA-II).

1.3. Novelty of the Study

This study differs from other similar studies in some aspects. The very first novelty is explicit water storage tank dimension optimization. Many studies are available for water distribution system optimization. However, few studies consider water storage tanks optimized simultaneously with other distribution system components (Farmani et al. 2005; Prasad and Tanyimboh 2008; Walters et al. 1999). Those studies optimizing water storage tanks consider implicit decision variable expressions for water storage tank sizes (such as diameter over height ratio and volume of the cylindrical storage tank). However, this study explicitly uses the tank dimensions

(height, diameter, elevation, etc.). This makes it easier to generate optimal tank designs and make decisions on various choices.

Another novelty is the ability to work on real networks without network simplification (skeletonization). Studies in the literature develop methodologies and test those methodologies on some fictitious water distribution systems. Some studies test their methodologies on real networks but those networks are greatly simplified networks in terms of pipes. This means, developed optimization methodologies are applied to simplified networks that consider only big sized pipes on main avenues but not smaller ones in smaller streets. The study that is presented in this thesis is capable of optimizing all network pipes regardless of diameter that considers every street pipe in the study area. Grouping of pipes are not considered.

Additionally, this study is based on a flexible coding that can provide optimal designs for different technical specifications or different country legislations. This property makes the study applicable to any pressure or tank size / volume requirement. Therefore, the developed methodology can be applied to real water distribution systems to make realistic designs.

This study also includes mechanisms to determine optimum genetic algorithm parameters (population size, crossover and mutation rates) used to generate optimal water distribution system designs. This is achieved by hypervolume comparison approach. Hypervolume is a method used to measure and compare performance of multiobjective optimization techniques (Zitzler and Thiele 1998). Due to case-specific nature of water distribution system optimization, it is hard to decide the genetic algorithm parameters at a first glance. Studies in the literature mainly presume some default values at the beginning and perform later analyses based on those values. There are not widely accepted values for genetic algorithm parameters, but usually low values are used for mutation, and higher values are used for crossover (Gençoğlu 2007, 2015; Wang et al. 2019). In the developed methodology, comparisons can be performed to determine the best choice for any genetic algorithm parameter. One of them (population size) is illustrated in this thesis study.

Finally, this study has a feature that allows automated design selection among pareto optimal solutions process in the absence of a decision maker. Few studies present in the literature focus on decision making part (Salehi et al. 2018). Almost all studies end up giving a set of pareto optimal solutions, but they do not focus on a single preferred solution. This thesis fills a gap in this aspect offering a preferred solution based on compromise programming. Compromise programming is a method to pick a solution among a pareto optimal set of solutions (Yu 1973; Zeleny 1973). By this way, decision maker is provided by an automatically selected solution which can ease his/her decision-making process. In the absence of a decision-making criteria, automatically selected solution can also be used as a reliable design.

1.4. Structure of the Thesis

The thesis is structured into seven chapters. First chapter gives introductory words with aim and novelty of the work. Second chapter goes through the literature under seven sections. In the third chapter, optimization method used in the study (NSGA-II) is explained in detail with simple examples. In fourth chapter, objectives and constraints used in the study is defined and exemplified on a simple network. Fifth chapter presents the full methodology depending on NSGA-II and the objectives defined. In sixth chapter three different networks are tried with the developed methodology and results are presented. In seventh chapter, discussions of the results are made with concluding remarks and future work suggestions.

CHAPTER 2

LITERATURE REVIEW

2.1. Optimization Methods

Optimization studies the area of finding feasible solutions to one or more objective functions. The feasible solutions should be extreme solutions representing such solutions as, minimum cost or maximum strength. Developing methods to find such solutions gain importance in engineering applications.

Single objective optimization mainly involves minimizing the cost for a particular problem. When a second objective is introduced, it is generally maximizing the strength or reliability of the designed problem. Two objective optimization problems are generally named as bi-objective optimization problems; when objectives are four or more they may be named as many-objective optimization problems (Deb and Jain 2014). Generally speaking, when number of objectives are more than one, they are named as multiobjective problems regardless of their objective count.

Numerous water distribution system optimization studies were performed since 1970s. Recent review papers investigate these past work (Maier et al. 2019; Mala-Jetmarova et al. 2017). Within the past studies, it can be traced that multiobjective studies considering water distribution systems are being performed in increasing numbers since 2000s. Some multiobjective optimization techniques are reviewed below.

When multiobjective optimization techniques are being reviewed, Nondominated Sorting Genetic Algorithms (NSGA) appears as one method worth mentioning (Srinivas and Deb 1994). It depends on the nondominated sorting idea which first creates classes of equal weight from an initial population. These classes evolve

through generations to reach a pareto front, which is the main outcome of a multiobjective optimization. This technique has also mechanisms to provide diversity along the pareto curve.

In water distribution system optimization, if a number of water storage tanks are present in the problem as decision variables, length of the chromosome may be variable. Structured Messy Genetic Algorithms (SMGA) are used in some studies to include variable sized decision variables (Halhal et al. 1995, 1997; Walters et al. 1999). In these studies, it is also mentioned that SMGA performance is sensitive to initial population size.

Particle Swarm Optimization (PSO) (Kennedy and Eberhart 1995) and its multiobjective variation Multi-Objective Particle Swarm Optimization (MOPSO) are also methods used in water distribution system optimizations. Some studies using these techniques and focusing only on pipe diameter sizing are present in the literature (Montalvo et al. 2008; Pandey and Mathur 2013).

NSGA methodology is improved by Deb et.al by a number of times. NSGA-II (Deb et al. 2002), NSGA-III (Deb and Jain 2014; Jain and Deb 2014) and u-NSGA-III (Seada and Deb 2015) are some of these improved techniques.

NSGA-II is one of the most widely used techniques especially in water distribution system optimization techniques. One reason for its popularity is that it does not contain any external parameter other than standard genetic algorithm parameters. This property makes NSGA-II easy to implement. The major difference from NSGA is the former method is classified as a non-elitist method while NSGA-II is classified as an elitist method (Deb 2001). Elitist methods are methods discarding worse elements and keeping best ones through the optimization process to reach a solution faster with better diversity. NSGA-II is the optimization method used in this study and will be explained in detail through Chapter 3.

NSGA-III is introduced as an optimization technique suitable for many-objective optimization studies (Deb and Jain 2014; Jain and Deb 2014). While u-NSGA-III is

introduced a method suitable for single, multiple and many objective (more than four) problems. The main difference between NSGA-II and NSGA-III (and u-NSGA-III) is mechanisms used to maintain diversity in objective function space.

Some recent optimization studies on water distribution systems are performed by using Simulated Annealing (SA) technique with four objectives (Marques et al. 2018). Other than the classical objective of cost minimization, three other minimization objectives are included in this study. These are; minimizing pressure deficit, undelivered demand and carbon emissions. When three objectives are considered it is possible to show results on a three-dimensional plot, however it is not possible to plot all four objectives at the same time in this case. So, two-by-two plot comparisons among objectives are made in this study.

Several comparisons among optimization methods are done and in a majority of them NSGA-II steps forward. In the study from Monsef et.al., NSGA-II is reported to perform better than MOPSO and it generated similar pareto curves with Multi-objective Differential Evolution (MODE) (Monsef et al. 2019b). Four benchmark networks are optimized in the study and it is reported that some algorithms perform better than the others depending on the network. This gives an indication that water distribution system optimization is case specific.

NSGA-II method is referred as a “good choice” for two-objective optimization for water distribution systems in a benchmark study comparing 12 different sample networks from the literature and comparing five different multiobjective evolutionary algorithms (Wang et al. 2015). For the studied networks true pareto or best estimate pareto curves are obtained.

In another comparative study (Zitzler and Thiele 1999) it is reported that NSGA achieved best results in some mathematical function optimization problems. It is followed by Vector Evaluated Genetic Algorithm (VEGA) and Niche Pareto Genetic Algorithm (NPGA).

In a recent publication (Wang et al. 2019), parameterization of NSGA-II is discussed for water distribution systems optimizations. In this study, NSGA-II is called as an “industry standard” method in water resources optimization problems. The same study, reviewed many past works done with NSGA-II and emphasized population size as the most important parameter with NSGA-II optimizations. Three benchmark networks are optimized and also optimum parameter determination (population size, crossover probability, mutation probability, number of function evaluations) is performed. The study states, fine tuning of NSGA-II parameters is not widely investigated and generally recommended values are used in the optimization studies.

2.2. Reliability Measures

Dating back to 1970s, optimization of water distribution systems considered a single objective namely minimizing cost. In 2000 two distinct studies focusing on the reliability of water distribution networks are published. One study tries to express reliability in terms of “entropy of the water distribution system” (Tanyimboh and Templeman 2000), whereas the other study defines reliability in terms of an energy based concept, “resilience” (Todini 2000). These concepts and their numerous variations are widely used as a second objective in water distribution network optimization.

In the study of Tanyimboh and Templeman (2000), entropy is defined as a function of “internal pipe flow rates” and reliability is defined as the “ratio of mean value of flow delivered to the flow required”. It is statistically shown that these two are correlated with each other, so flow entropy is defined as a maximization objective for water distribution systems.

On the other hand, Todini (2000) proposes an index named as resilience index, (I_r) which is defined in terms of power. I_r is defined as the ratio of total excess power available at nodes to the maximum power that would be internally dissipated to satisfy constraints in the system. Resilience index is again used as a maximization

objective in many water distribution system optimization studies. A variation of resilience index, network resilience (I_n) (Prasad and Park 2004) is used in this study so resilience concept will be widely explained in Section 4.1.

Comparative studies between entropy and resilience concepts show that resilience and its variations perform better than entropy and its variations. In a study comparing “indirect reliability measures” I_r , I_n and entropy with “direct reliability measures” defined in the study (demand satisfaction rate, average satisfaction during fire conditions and average satisfaction of hydrants), resilience measures are said to have strong positive correlation with the direct reliability measures on the other hand entropy did not have a positive correlation (Creaco et al. 2014).

In a similar comparative analysis, various energy-related indices (including I_r and I_n) are compared with a variation of flow entropy showed that entropy performed poorly compared to energy indices (Liu et al. 2017). Additionally, I_n showed the best performance among others in some cases.

Another study performed on a realistic network done with 24-hour extended period simulation (EPS), it is reported that, I_r outperforms entropy (Cimorelli et al. 2018). Also, it is concluded that entropy cannot be used with limitations on budget in a water distribution system rehabilitation problem.

In a study comparing three reliability measures (I_r , I_n and entropy), a fourth one is introduced from the combination of entropy and I_r (Raad et al. 2010). As a conclusion, fourth mixed measure and I_n are said to be most practical measures, however none of the measures are said to be clearly best.

In a very recent study (Monsef et al. 2019a), deficiency of reliability indicators under abnormal operation conditions is looked for. Five benchmark networks with 20 different cases each are studied and entropy measures showed worst performance. The study also suggests a method to choose which reliability measure is best for the specific network studied.

In another study, a systematic critical review is performed among 21 resilience measures considering 11 criteria (Shin et al. 2018). The review criterion is called as “axiomatic design process”. I_r and I_n are reported to address six criteria among 11 where the maximum is 7 out of 11 again for resilience based measures. It is criticized in the paper that, many of the resilience measures do not consider “long term variation of system resilience” meaning the formulations depend on results for a snapshot time only. When realistic cases like a city’s distribution network is considered, water demands and flows in pipes will be variable through time so temporal effects should be considered.

After the proposal of I_r , various additions to the index are made. One of the most widely used I_r variation is network resilience, I_n (Prasad et al. 2003; Prasad and Park 2004). This variation of resilience index, only adds a diameter uniformity parameter to the original formulation. Adding a diameter uniformity parameter, helps less diameter changes at the nodes. This enables a more uniform distribution of diameters among the network. As mentioned above, detailed explanation for I_n will be provided in Section 4.1 and further.

Similar to diameter uniformity, a loop diameter uniformity parameter (C_u) is formulated and used as a third maximization objective along I_r maximization and cost minimization by researchers including Todini (Creaco et al. 2016a). In this work, cost vs I_r , cost vs I_n and cost vs I_r vs C_u optimization problems are compared by using NSGA-II method. No significant difference is noted as a result in the study. Additionally, it is emphasized that steady-state snapshot simulations are performed rather than long term extended-period simulations (EPS), and water storage tank activities can affect results if EPS is performed.

Considering EPS usage with resilience index, another study offers using statistical measures (minimum, maximum, mean or median) on I_r values to find a single representative value (Creaco et al. 2016b); in this study, resilience index is generalized to be used in pressure-driven modelling and leakage. Authors suggest using median of I_r values computed at each network operation instant.

In a study considering topography effects on average values of I_r are used as representative values (Jeong et al. 2017). In this study, various different topographies (flat, uphill, downhill, convex, concave and irregular) are applied on a fictitious network to see the temporal change in I_r values. It is concluded that biggest deviations in I_r occur in uphill network topographies.

Lastly, another study creates a performance measure from other four measures (Dziedzic and Bryan 2016). These four measures are defined in the study and named as; reliability, vulnerability, resilience and connectivity. A geometric average of these four are used as an alternative measure and it is benchmarked on two networks. According to the reported results, the new measure generally follows similar trend with the four measures defined.

2.3. Water Distribution Network Parameter Optimization

Whether single or multiobjective optimization in water distribution systems is considered, two main focus areas are defined: Pump operation and water quality (Mala-Jetmarova et al. 2017). In both of the focus areas, majority of the studies consider only network pipe diameters as decision variables (Farmani et al. 2004).

Starting from early 2000s, optimization studies considering storage tank parameters with pipe diameters are performed (Batchabani and Fuamba 2014). These studies are generally multiobjective and resolved with some type of genetic algorithms. In the paper by Batchabani and Fuamba (2014), possible optimizable storage tank parameters are discussed. Additionally, a procedure is suggested to choose most suitable approach to include storage tanks in optimization studies. In the same study, typical storage tank design guidelines from various countries (USA, France, South Africa and Quebec/Canada) are also compared. A tabulated comparison is given in Table 2.1. Total storage tank volume is calculated by summing up balancing, fire and emergency volumes.

Table 2.1. Volume Comparisons from Countries' Guidelines

Country	Balancing Volume	Fire Volume(m ³)	Emergency Volume
United States	%35 of peak daily consumption	340	1 day of peak consumption
France	%30 of peak daily consumption	120	2 days of average consumption
South Africa	Included in Emergency Volume	108	2 days of average consumption
Canada Quebec	Included in Emergency Volume	120	1.5 days of average consumption
Turkey (1939)	Not less than %33 of average daily consumption	36 – 144	–
Turkey (1950s)	Not less than %33 and not more than %100 of average daily consumption for gravity fed networks not less than %25 of average daily consumption for pumped networks	Included in Emergency Volume	1/3 – 1/4 days of average consumption
Turkey (1985)	Same as 1950s criteria	18 – 360	–
Turkey (2013)	%33 of average daily consumption	36 – 900	1/10 days of average consumption

(Batchabani and Fuamba 2014; “İçmesuyu Tesisleri Etüt, Fizibilite ve Projelerinin Hazırlanmasına Ait Teknik Şartname” 2013; “Şehir ve Kasaba İçmesuyu Projelerinin Hazırlanmasına Ait Yönetmelik” 1985; “Şehir ve Kasabalara Getirilecek İçme Suyu Projelerinin Hazırlanmasına Ait Talimatname III” n.d.; Özand 1939)

A historical comparison is done for Turkey's regulations from 1939 to 2013 (“İçmesuyu Tesisleri Etüt, Fizibilite ve Projelerinin Hazırlanmasına Ait Teknik Şartname” 2013; “Şehir ve Kasaba İçmesuyu Projelerinin Hazırlanmasına Ait Yönetmelik” 1985; “Şehir ve Kasabalara Getirilecek İçme Suyu Projelerinin Hazırlanmasına Ait Talimatname III” n.d.; Özand 1939). Volume requirements in Turkey's regulations are determined by Bank of Provinces (İller Bankası – formerly Belediyeler Bankası). Findings from this comparison are also included in Table 2.1. From this comparison an increasing trend in fire volumes can be seen. Also, it should

be noted that volumes defined in Table 2.1 are proposed for small towns or villages. Big metropolitan cities' water distribution systems are not limited strictly for the values given.

Almost all studies considering pipe diameter optimization, test their developed methods on benchmark problems. These benchmark problems consist of either a real or a fictitious water distribution system. Two very popular benchmark networks are 2-loop network (or Shamir network) (Alperovits and Shamir 1977) and Hanoi network (Fujiwara and Khang 1990, 1991). These two fictitious and simple networks focus on pipe diameter optimization on a single objective basis of minimizing costs. The mentioned networks are proposed for design purposes. These two networks are also studied extensively in this thesis and will be explained in detail in Sections 6.1 and 6.2. Both of these networks are studied in a multiobjective approach in a recent study (Wang et al. 2015). True pareto curve for 2-loop network and a best estimate of true pareto for Hanoi network are reported to be found in this study.

Another fictitious but very popular benchmark network is Anytown network (Walski et al. 1987). This network is a pumped network unlike 2-loop and Hanoi networks having two storage tanks, a pump station, present and future settlement areas. These properties make this problem a rehabilitation problem rather than a design problem. Original problem is formulated as a single-objective problem however, many researchers modify it as a multiobjective problem and mainly optimize cost vs I_r (Basupi et al. 2014; Basupi and Kapelan 2015; Farmani et al. 2004, 2005, 2006; Prasad and Tanyimboh 2008). Anytown network is another benchmark network benefitted in this thesis by using its unit cost values. Abovementioned six works will also be used for comparison of results.

2.4. Constraint Handling

Real world optimization problems are mostly constrained problems by their nature. Objective functions are solved with accompanying constraint functions. Therefore, constraint solving methodologies are of great importance. Constraints define whether a solution is feasible or infeasible. In many real-world problems, infeasible solutions are physically impossible to implement therefore infeasible solutions should be avoided in the solution set.

In evolutionary multi-optimization, most constraint handling methods can be classified into 5 different categories (Craenen et al. 2001; Deb 2001; Michalewics and Schoenauer 1996):

- Methods based on preserving feasibility of solutions
- Methods based on penalty functions
- Methods making a distinction between infeasible and feasible solutions
- Methods based on decoders
- Other hybrid methods

In most applications, penalty functions are used. However formulating penalty functions can be difficult as they are mostly case specific (Gençoğlu 2007). In this work, a dynamic penalty function (see below) is implemented having a case specific coefficient in the formulation.

Furthermore on penalty functions, widely used 4 groups can be stated below (Yeniay 2005):

- Death penalty, rejecting infeasible solutions
- Static penalty, constant penalty values are applied regardless of the generation
- Dynamic penalty, penalty values changing with increasing generations
- Adaptive penalty, penalty values updated from current population information

An application of penalty functions with multiobjective optimization on water distribution system is performed with search space reduction (Čistý and Bajtek 2009). NSGA-II method is used in the optimization study of Hanoi network. A two-stage genetic algorithm is built and results implicate that proposed methodology shows better performance than traditional genetic algorithm methods.

Deb and Agrawal proposed a niched penalty approach for constraint handling (Deb and Agrawal 1999). It is a parameterless approach making it easy to implement. In this method a distinction is made between feasible and infeasible solutions and infeasible solutions are not evaluated (as infeasible solutions are impossible to implement in practice) to improve computation time.

In a technical report by Kanpur Genetic Algorithm Laboratory (KanGAL) an adaptive constraint handling methodology is proposed (Datta and Deb 2012). In this method at every generation population member is sorted ascending based on their constraint violation values. First half of the population is then considered by taking average of their constraint violation values. And this average value is used to normalize all constraint values to evolve infeasible members' constraint violation values to zero through generations. As a result, infeasible members are avoided through generations.

Some studies handle constraints during selection phase of multiobjective genetic algorithms. In such a study (Coello Coello and Montes 2002), members are compared with each other depending on their feasibility-infeasibility situation. When all possibilities are taken into consideration below four situations are present:

- Both solutions are feasible; solution with a better fitness value is selected.
- One solution is feasible the other one is infeasible; feasible solution is selected.
- Both solutions are infeasible; nondominated solution is selected if the other solution is dominated.
- Both solutions are infeasible and both are either dominated or nondominated; solution with a smaller constraint violation is selected.

Methods of this type do not need any kind of penalty function formulations and are easy to implement as they do not require any parameter.

A very similar method is proposed by Deb and used as constraint handling method in the scope of this thesis. This method is explained as “to devise a penalty function approach that does not require any penalty parameter” (Deb 2000). In this method a constraint violation calculation is proposed for every solution and below three criteria is checked for a pair of solutions during tournament selection:

- Any feasible solution is preferred to any infeasible solution.
- Within two feasible solutions, the one having a better objective function value is preferred.
- Within two infeasible solutions the one having smaller constraint violation is preferred.

This concept is furthermore developed (Deb et al. 2002) under the name of “constrained domination concept” to cover multiobjective optimization problems within NSGA-II scope and its details will be given under Section 3.3.

2.5. Diversity of Solutions

Multiobjective optimization has two targets to achieve. First one is to find solutions as close as possible to the pareto-optimal curve. The second one is to obtain final solutions as diverse as possible from each other. The first target is obvious under the

concept of optimization. Solutions of true or near optimal should be found with no doubt if an optimization study is being carried. However finding diverse solutions in the solution set is entirely specific to multiobjective optimization (Deb 2001). Having diversity in the solution set will provide decision makers a good set of trade-off solutions. So, they will have a big variety of choices to find a solution that suits their needs.

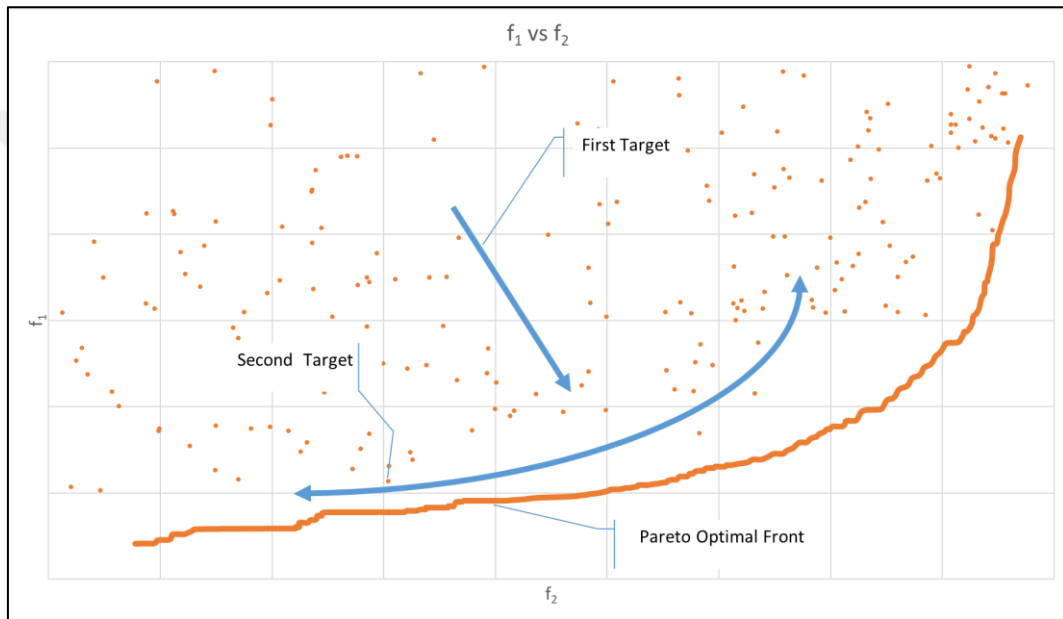


Figure 2.1. The Two Targets of Multiobjective Optimization

To generate and preserve diversity, many methods are present in the literature. These methods mainly try to preserve diverse solutions over generations or they pick one representative solutions from too near solutions (clustering). Diversity generation during pruning stage is one of the applicable methods. Pruning is an essential part in evolutionary multiobjective algorithms that maintains original population size after it was increased by crossover and mutation like operators. In a technical report from KanGAL, two methods for pruning in many-objective problems is proposed (Kukkonen and Deb 2007). In these methods, two different crowding estimation

techniques (nearest-neighbour based) for pruning a set of nondominated set of solutions is offered. In the increased population, every member is assigned a crowding distance value, which is a measure of elements around itself. The extreme solution points are preserved by assigning infinite to their crowding distances. In addition, the population is sorted descending according to the crowding distances. The members of the population are reduced one-by-one starting from the lowest crowding distance element. At every step, crowding distances are recomputed. After reaching the original population size the method stops. It is reported that this method provides better diversity than methods based without nearest neighbour usage (Kukkonen and Deb 2006). It is also reported that best diversity results are achieved in three-objective problems.

Adaptive mutation is used to improve NSGA-II diversity in a study (Carvalho and Araujo 2009). In the study, an adaptive mutation operator is developed to control mutation magnitude, and getting required information from diversity of candidate solutions. The results for five test problems indicate that this method provides better diversity measures from usual mutation technique methods.

In another study aiming to improve NSGA-II diversity, extreme solutions are tried to be preserved with an “agglomerative hierarchical clustering” (Vachhani et al. 2016). The method adds a criterion to NSGA-II methodology in the pruning stage. While NSGA-II selects N best members within $2N$ elements based on ranks, this method proposes a selection based on ranks and clustering method it defines. According to the results, diversity of the proposed method is high when standard NSGA-II is considered.

In another study aiming to increase NSGA-II diversity, neighbourhoods for elements are defined by defining spheres around them (Zheng et al. 2012). Additionally, selection operator is modified to keep extreme solutions and prohibits or postpones adjacent elements to pass into new generations. According to presented results, the proposed method enhanced diversity compared to standard NSGA-II methodology.

Studies aiming to generate diversity among design variables rather than objective functions are also present (Zadorojniy et al. 2012). According to Deb, maintaining diversity in decision variable space mostly means maintaining diversity in the objective function space, it may not be the case in all problems, especially when linearity is not present in the problem (Deb 2001). According to the results of the study mentioned above they obtained the diversity required for both test problems and for real life problems.

In another study about diversity, a diversity maximization approach is proposed (Masin and Bukchin 2008). In this approach, pareto optimal solutions are found by a predefined diversity calculation and a full set of feasible element generation is ensured. With a subset of the feasible front, approach finds next pareto optimal solution combining with the existing ones. The most diverse subset of efficient front is then gathered. The performance of the approach is presented on two benchmark function problems.

In water distribution systems a diversity increase method is proposed (Tanyimboh and Seyoum 2016). The method focuses on feasible and infeasible solutions' existences and simultaneous development based on Pareto dominance. It is emphasized that a boundary search is conducted between feasible and infeasible region. The method is applied on a sample network and results indicate that proposed method is reliable and effective. In the sample network, it is reported that multiple sources, multiple demand categories and EPS were present.

Although methods like NSGA-II has diversity maintaining mechanisms in them, improvements may be necessary in some cases. An improvement in NSGA-II is done by Deb using clustering concept (Deb et al. 2003). The improved version of NSGA-II is named as clustered NSGA-II (C-NSGA-II). The diversity maintaining method used throughout this theses study is the method mentioned here and it will be explained in detail in Section 3.3.7. In short, the method groups nearby solutions to form clusters and use the solution nearest to the centre of gravity of the related cluster as a representative solution.

2.6. Performance Measures

As mentioned in previous section, there are two main goals in multiobjective optimizations. To obtain close solutions to the true pareto solutions and to obtain as diverse solutions as possible. When new methods are proposed to achieve these goals, their performance of achieving the goals should be measured. There exist three ways to measure a multiobjective optimization methods performance (Deb 2001):

- Measures for Pareto-Optimal Closeness
- Measures for Diversity Among Solutions
- Measures for Both Pareto-Optimal Closeness and Diversity Among Solutions

To be able to measure performance using first category true pareto curves should be known. Except some studies reviewed above (Wang et al. 2015), true pareto curve for multiobjective optimization for water distribution systems are not known yet. Their best-known estimates are presented in various studies. Therefore, only works from third category will be reviewed here as one of them is successfully implemented in the thesis work.

Running performance metrics are not classified in the above three categories but they gain importance with the new studies. They are based on a good updated archive of the population members through generations. With that archive information, closeness to pareto and a measure for diversity can be calculated at each generation. As a result, a generation-wise convergence and diversity plot can be obtained (Deb and Jain 2004). A sample plot for the explained concept is given in Figure 2.2 (Deb and Jain 2002). It is seen that, within generations, convergence metric decreases implying that true pareto curve is reached, and after that diversity increases for some generations generating diverse solutions at the end. In this study, some design criteria for running performance metrics are given and some metrics are applied on some test problems.

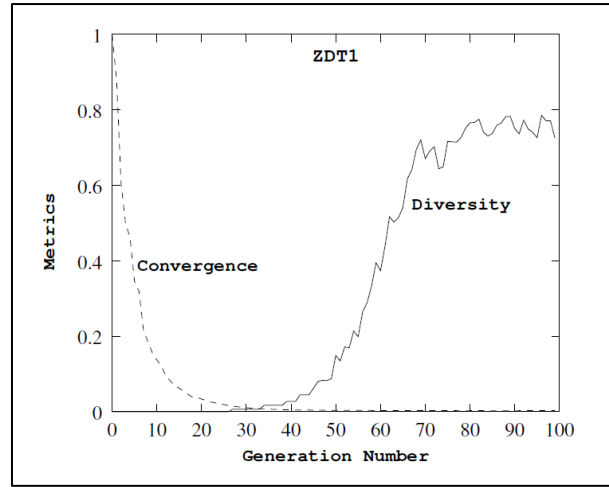


Figure 2.2. A Sample Plot from a Running Performance Metric
(Deb and Jain 2002)

Hypervolume or S-Metric is defined by Zitzler and Thiele to measure performances of different multiobjective evolutionary optimization techniques (Zitzler and Thiele 1998). This metric is in the third category defined above. It measures closeness to pareto optimum and diversity at the same time. It mainly defines a best point (defined from the end points of obtained results) and a worst point defined from worst objective function values obtained throughout the generations. These two reference points and the pareto curve itself are used to form hypervolumes in the objective function space. Then the volume of the obtained hypervolumes are used to compare performances among methods. Obviously, higher hypervolumes will indicate a bigger coverage in the search space. And both convergence and diversity issues will increase this coverage.

In the hypervolume method, it is an important task to choose the reference points and to normalize objective function values. Depending on how these selections are done, hypervolume values will change significantly (Knowles and Corne 2002).

When there are two objectives, hypervolumes will be polygon shapes, and when there are three objectives, hypervolumes will represent 3D shapes. However, after

four and more objectives, hypervolume calculations can be a tough and time-consuming job. Some algorithms for speeding up hypervolume calculations for many-objective optimization studies are present in the literature (While et al. 2006).

Fu et.al used hypervolumes aiming to reduce computational complexity in water distribution system optimization (Fu et al. 2012). A variation of NSGA-II (ϵ -NSGA-II) is used as optimization method and Anytown network is chosen as one of the test networks. More information is provided about hypervolumes in Section 3.4.

2.7. Decision Making

Decision making in general and multi-criteria decision making in particular is not the main topic of this thesis; however, the question of which solution to choose from supplied set of pareto optimal solutions is an important question. Deb categorizes multi-criterion decision making techniques into two: post-optimal and optimization level techniques (Deb 2001). In the first one, final solutions are analysed to prefer one of them. In the second technique, optimization process is steered towards the preferred region of the pareto curve within generations. This second technique needs some information of the designer or decision maker to prefer some parts of the objective functions in order to steer the process. However, if decision makers provide no or little information about their preferences, a choice of solution among pareto optimal solutions can be done by the optimizer.

Zeleny and Yu defined a “Compromise Programming” approach to choose a solution among pareto optimal solutions as a post-optimal technique (Yu 1973; Zeleny 1973). In this approach, given a set of solutions, the best point is set by using extreme solutions in the pareto set. Then every solution is compared to this best point by a distance metric between them (Deb 2001). The solution having smallest distance to the ideal point is chosen as the preferred point.

In a similar manner, the worst point can be defined and every candidate solution is then compared to this worst point. This time, solution having largest distance is preferred (Zeleny 2011).

Multiobjective optimization problems focusing on water distribution systems, especially focusing on real world systems may include hundreds or thousands of decision variables. Including this much information may complicate decision making actions for designers. With improved graphical innovations in computers, some graphical visualization tools are offered to use as a decision support tool. These tools may help designers or decision makers to set their preferences better.

Blasco et.al propose “level diagrams” to overcome abovementioned problems (Blasco et al. 2008). In their work, a final set of pareto optimal solutions are processed to create 2D scatter plots of one of the objectives on y axis, and one of the decision variables on x axis. The produced scatter plots are interactive with each other. Selecting one solution from the plots will highlight the corresponding decision variable and objective functions in the other plots for that particular solution. By this way, visual decisions on decision variables or objective functions become easy, so the solution set can be efficiently downsized to reach some solutions.

Other similar methods are PROMETHEE and GAIA (Mareschal and De Smet 2009). They are described as multicriteria decision aid methods that use outranking. PROMETHEE is defined to be prescriptive, presenting the set of solutions with some ranking to the decision maker. On the other hand, GAIA is defined to be descriptive, presenting important features of the problem and assisting decision maker to make preferences.

Within the scope of this thesis work, it is assumed that no preference information is present for decision making actions. Therefore, simple illustrations of compromise programming explained above are integrated to the methodology. They are explained in more detail in Section 3.5.

CHAPTER 3

NON-DOMINATED SORTING GENETIC ALGORITHM – II

3.1. Single vs Multi-Objective Optimizations

Optimization is usually defined as a search for one or more feasible solutions which are extreme points of one or more objective functions. It is relatively easy to find extreme solutions in a problem when single objective is considered (by visual analysis of corner points in decision variable space, by taking first order derivative of the function, etc.). The main goal is to find minimum or maximum points of the objective function; therefore, the objective is either minimization or maximization. After finding extreme points in the problem, some of them are local extreme points and only one of them is global extreme point.

Things get complicated when two-or more objectives are considered. When there is more than one objective, usually they are conflicting objectives. With conflicting objectives, it becomes harder to find single global optimum solutions. Rather, a set of equal-optimum solutions can be found. This set of equal-optimum solutions are called pareto optimal solutions. A formal definition for pareto optimal solutions is provided in Section 3.2.

It can be noted that, in most engineering optimization works, when single objective is considered it is usually minimization of costs on the other hand when two objectives are considered those are minimization of costs and maximization of strength. These two objectives are mostly conflicting objectives meaning that stronger solution means a more expensive solution and vice versa. Therefore, when one objective is improved, the other one will be declined.

In single objective optimization with two decision variables, it is possible to draw a decision variable space (Figure 3.1), however, in multiobjective optimization problems, both a decision variable space and an objective function space can be

plotted (Figure 3.2). These spaces are mapped with each other through some functions defining the problem. In addition, these spaces are usually bounded by functions called constraints. In problems involving two decision variables, the optimum solution is one of the corner points of the feasible region (Mays and Tung 2002). However, in multiobjective studies, it is not possible to talk about a single optimum solution but a group of equal optimum solutions. Depending on the objectives, optimum solutions form some curves when two objectives are considered. They are illustrated in Figure 3.3.

Most real-world applications are constrained optimizations. So, to be able to find applicable solutions, constraints should not be violated. Any solution violating one or more constraint is not an applicable solution and it should be avoided. With these explanations it can be said that, solutions satisfying all constraints are called feasible (or efficient) and any solution violating at least one constraint is called infeasible (or inefficient).

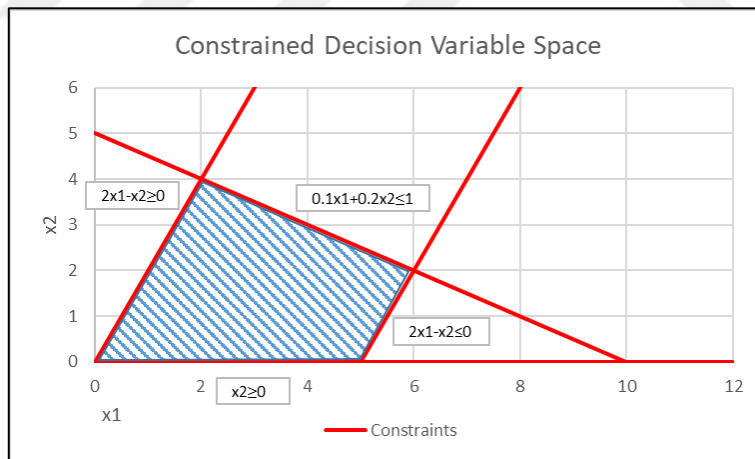


Figure 3.1. Constrained Decision Variable Space (Deb 2001)

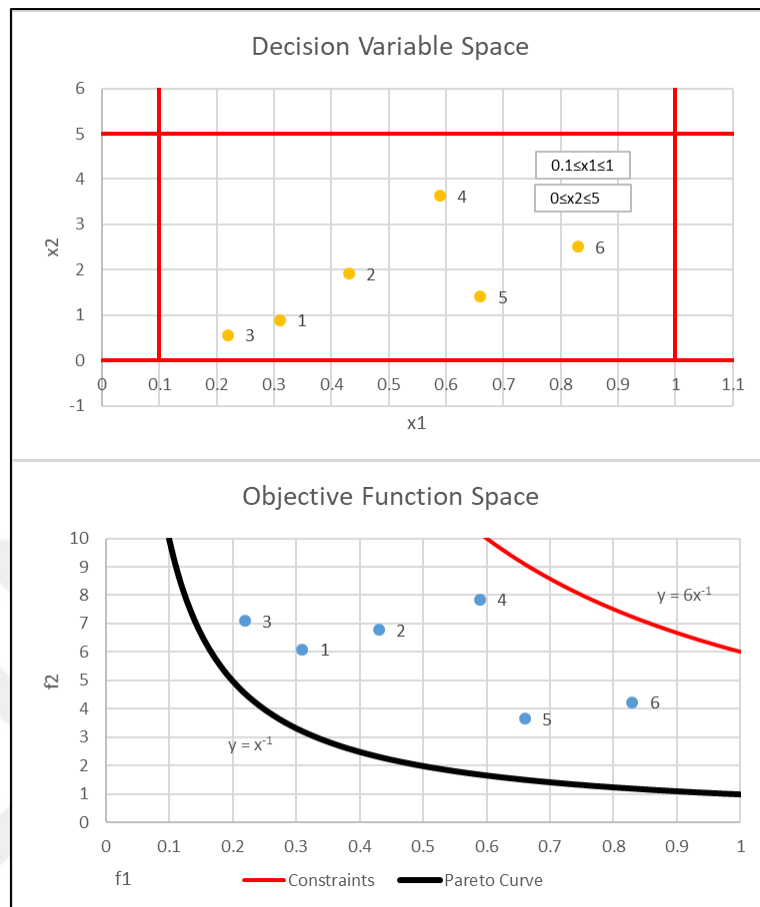


Figure 3.2. Decision Variable and Objective Function Spaces in Multi-Objective Optimizations (Deb 2001)

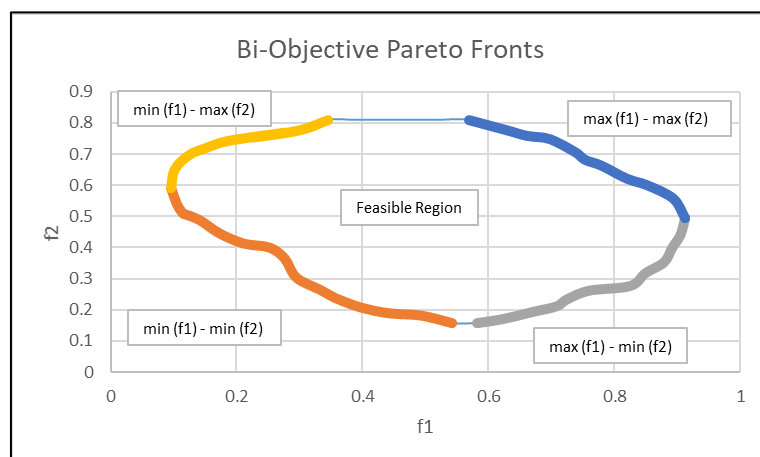


Figure 3.3. Possible Pareto Fronts in a Bi-Objective Optimization Problem

3.2. Domination and Non-Domination Concepts

Before explaining non-dominated sorting, domination concept should be explained. To explain the concept three definitions are given below from Deb (2001).

Domination: “A solution $x^{(1)}$ is said to dominate the other solution $x^{(2)}$, if both conditions 1 and 2 are true:

1. The solution $x^{(1)}$ is no worse than $x^{(2)}$ in all objectives, or $f_j(x^{(1)})$ is not better than $f_j(x^{(2)})$ for all $j = 1, 2, \dots, M$.
2. The solution $x^{(1)}$ is strictly better than $x^{(2)}$ in at least one objective, or $f_j(x^{(1)})$ is better than $f_j(x^{(2)})$ for at least one $j \in \{1, 2, \dots, M\}$.”

Non-dominated set: “Among a set of solutions P , the nondominated set of solutions P' are those that are not dominated by any member of the set P .”

Globally Pareto-optimal set: “The non-dominated set of the entire feasible search space S is the globally Pareto-optimal set.”

To explain NSGA-II mechanisms, sample parent and child population values are considered in Table 3.1 (Deb 2001). First six elements (1 – 6) in the table are parent members and remaining six elements (a – f) are child members that will be produced in crossover and mutation phases. x_1 and x_2 are decision variables for the members and f_1 and f_2 are objective function values. Both objectives are minimization objectives. Objectives are defined below:

$$\text{Minimize } f1(x) = x1$$

$$\text{Minimize } f2(x) = \frac{1 + x2}{x1}$$

$$\text{subject to } 0.1 \leq x1 \leq 1$$

$$0 \leq x2 \leq 5$$

Table 3.1. Sample Values for a Population

Solution	x1	x2	f1	f2
1	0.31	0.89	0.31	6.1
2	0.43	1.92	0.43	6.79
3	0.22	0.56	0.22	7.09
4	0.59	3.63	0.59	7.85
5	0.66	1.41	0.66	3.65
6	0.83	2.51	0.83	4.23
a	0.21	0.24	0.21	5.9
b	0.79	2.14	0.79	3.97
c	0.51	2.32	0.51	6.51
d	0.27	0.87	0.27	6.93
e	0.58	1.62	0.58	4.52
f	0.24	1.05	0.24	8.54

Considering only first six parent members, it is possible to say:

- Element 1 dominates elements 2 and 4, it cannot dominate 3, 5 and 6.
 - Element 1 must be in a better front than 2 and 4.
- Element 2 dominates element 4, it cannot dominate 1, 3, 5 and 6.
 - Element 2 must be in a better front than 4
- Element 3 dominates element 4, it cannot dominate 1, 2, 5 and 6.
 - Element 3 must be in a better front then 4
- Element 4 cannot dominate any element.

- Element 5 dominates element 6, it cannot dominate 1, 2, 3 and 4.
 - Element 5 must be in a better front than 6
- Element 6 cannot dominate any element.

Depending on the above criteria, elements 1, 3 and 5 are in the 1st front, elements 2 and 6 are in the 2nd front and element 4 is in the 3rd front. A simple procedure to define fronts from domination of elements is defined in Section 3.3.1. Parent elements in Table 3.1 and defined fronts are given in Figure 3.4 and Figure 3.5.

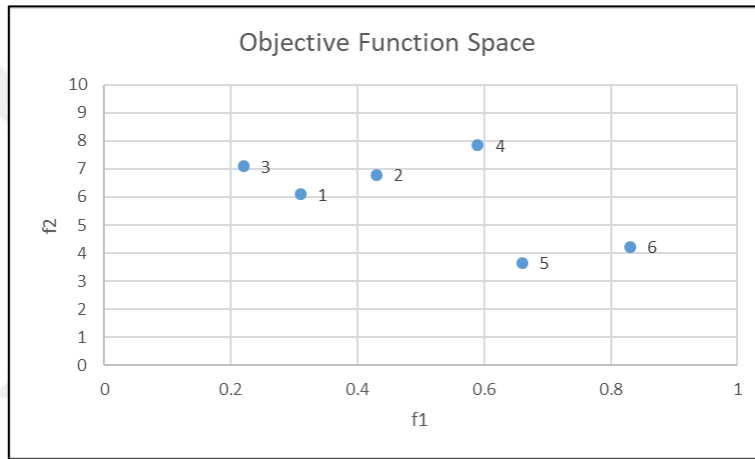


Figure 3.4. A Sample Multi-Objective Function Space

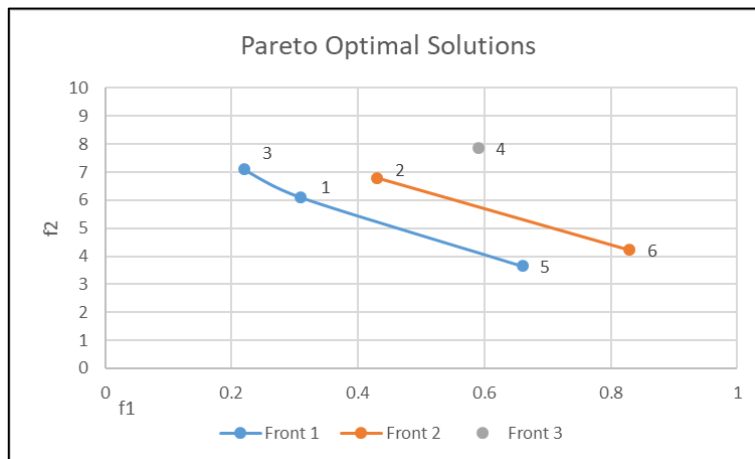


Figure 3.5. Solutions Classified into Pareto Optimal Fronts

3.3. NSGA-II

Among genetic algorithm variations, a method called NSGA stands one-step ahead with respect to the others (Srinivas and Deb 1994). NSGA is further improved under the name NSGA-II (Deb et al. 2002). Developers of NSGA-II explain it as an elitist evolutionary multiobjective optimization method, which sorts the population into various fronts obtaining a very quick convergence of the pareto optimal values.

The main difference between NSGA and NSGA-II is the second algorithm is an elitist algorithm. NSGA-II has the tools to preserve good solutions to next generations while rejecting worst solutions for a faster convergence.

Figure 3.6 shows the two mechanisms in NSGA-II, applied to the population members to reach pareto optimality. The first one is non-dominated sorting and the second one is crowding distance sorting. While non-dominated sorting pushes the solutions towards the ideal pareto curve, crowding distance sorting lays the solutions apart from each other to cover a better spread along the curve. The two mechanisms are explained in detail with sample calculations in Sections 3.3.1 and 3.3.2.

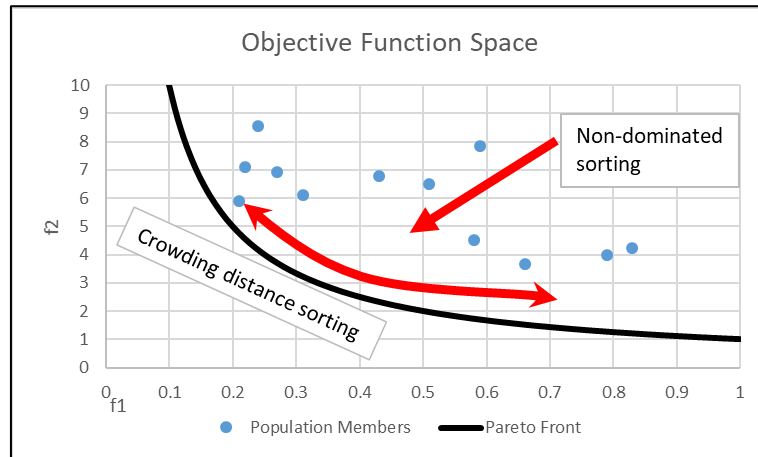


Figure 3.6. Non-dominated Sorting and Crowding Distance Sorting

The two mechanisms defined above try to achieve two goals. They are obtaining a solution set close to pareto optimal curve and getting solutions distributed well on this curve. In Figure 3.7, two different sets of solutions that could not achieve these goals are presented. The first set given in orange has a good distribution of solutions but failed to converge on the pareto optimal solution curve. The second set given in yellow achieved to converge on the pareto optimal front but failed to cover the whole curve.

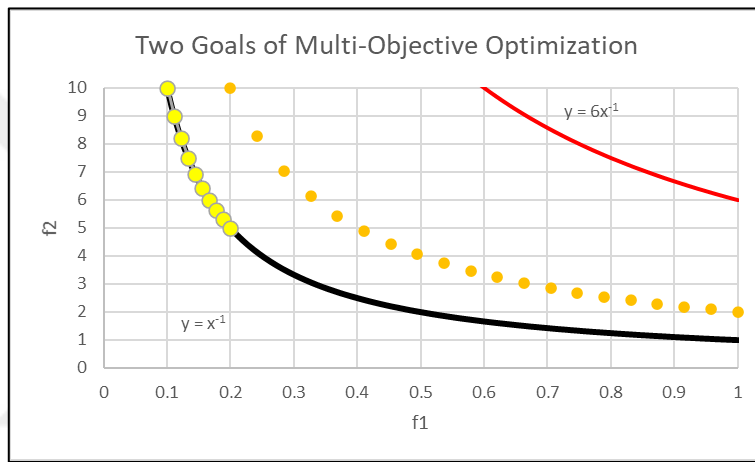


Figure 3.7. The Two Goals of Multi-Objective Optimization

In Figure 3.8, a flowchart of NSGA-II is provided. Procedure starts with creating a random population of size N . Afterwards every member is evaluated to obtain their objective function values. With the obtained objective function values, ranking is performed. It should be noted that ranking is a two-stage operation composed of non-dominated sorting and crowding distance sorting. After ranking, selection, crossover and mutation operations are applied to obtain a child population of size N . The child population is also evaluated for objectives and combined with parent population to have population size of $2N$. This combined population is again ranked, and half of them is pruned to return the original population size of N (elitism). The procedure returns to selection step to run for another generation until the stopping criteria is

met. At the final stage, pareto optimal solutions are found for the multiobjective optimization problem.

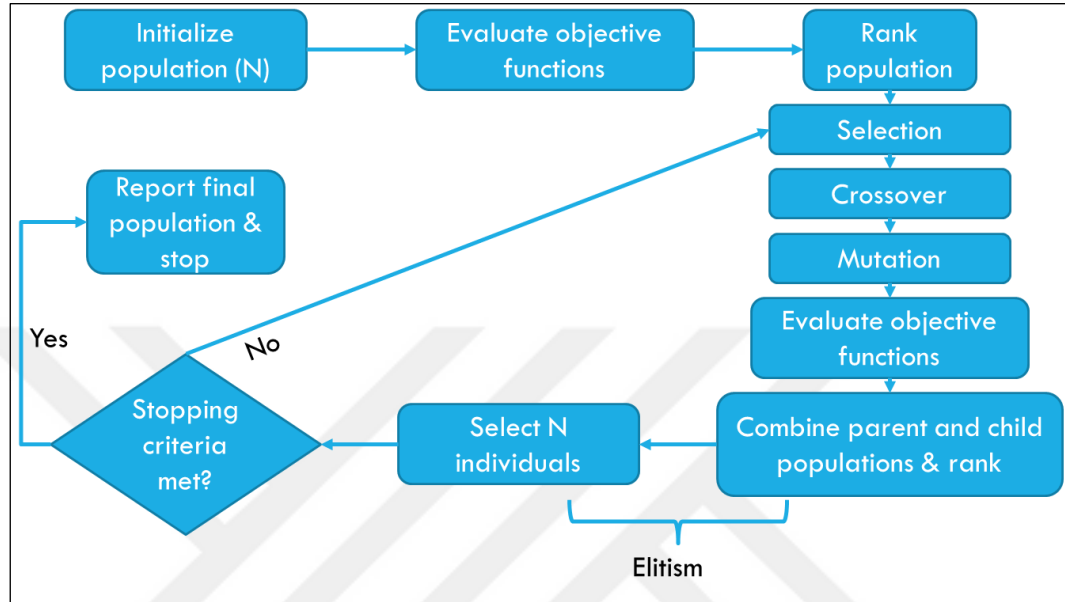


Figure 3.8. Flowchart of NSGA-II

Ranking phase is explained in Sections 3.3.1 and 3.3.2, selection phase is explained in Section 3.3.3, crossover phase is explained in Section 3.3.5, mutation phase is explained in 3.3.6 and elitism phase is explained in Section 3.3.7.

3.3.1. Fast Non-Dominated Sorting

Fast non-dominated sorting procedure starts checking the first member, p , of the population one-by-one with other members, q , in terms of domination. Two separate entities are calculated simultaneously during these checks. First entity records the number of members dominating p , second entity is the list of members that p dominates. Simply saying that two-way checks are made with each member whether p dominates q and whether q dominates p .

Every member having a domination count zero will become the first non-dominated front of the population. Now dominated sets of every first-front members are investigated and those elements' domination counts are reduced by one. If any element's domination count reduces to zero then they will become part of the second front. This procedure is continued until every member is assigned a front number starting from one. As a result, all population is sorted by front numbers indicating domination levels (Deb et al. 2002).

To explain the procedure in step-by-step detail, parent population from Table 3.1 is used. As all N members of a population is compared with each other one by one, there are totally $N-1$ passes in the mechanism. In this sample case of 6-member comparison, there are 5 passes. The passes are explained below:

Pass 1: 1 dominates 2 and 4. Dominated count of 2 and 4 are increased by one.

Pass 2: 2 dominates 4. Dominated count of 4 is increased by one.

Pass 3: 3 dominates 4. Dominated count of 4 is increased by one.

Pass 4: 4 cannot dominate any elements. No changes are done in the status.

Pass 5: 5 dominates 6. Dominated count of 6 is increased by one.

Pass 6: 6 cannot dominate any elements. No changes are done in the status.

Following tabulated results are achieved:

Table 3.2. Sample Values for a Population

Solution	f1	f2	Number of Dominations	Dominated Set of Elements	Front
1	0.31	6.1	0	2, 4	1
2	0.43	6.79	1	4	2
3	0.22	7.09	0	4	1
4	0.59	7.85	3	—	3
5	0.66	3.65	0	6	1
6	0.83	4.23	1	—	2

With Table 3.2, elements 1, 3 and 5 are marked with Front 1 and elements dominated by 1st front (2, 4 and 6) are revisited to decrease their number of dominations. This time elements 2 and 6 reach zero on number of dominations so they become the next front that is Front 2. Lastly, remaining element 4 is marked with 3rd front.

Finally, all the population is sorted in an ascending way according to their fronts and crowding distance sorting phase is applied to the resulting population.

3.3.2. Crowded Distance Sorting

Second method shown in Figure 3.6 is crowding distance sorting. This method helps preserving diversity of the members in each front by promoting members that are far from each other, and eliminating nearby elements. This is achieved by calculating a crowding distance value for each member for every front separately. For the calculation of this distance, the population is sorted according to first objective function values. After this sorting, first and last elements are assigned a crowding distance of infinite. Starting from second member, normalized distances are calculated and summed up from first objective function values. Normalization is done by dividing absolute value of differences of maximum and minimum objective function values for that specific objective. Same procedure is repeated for other objective functions and overall crowding distance is calculated as the sum of

individual crowding distance values corresponding to each objective. Simply speaking, the crowding distance for a member i will be the sum of normalized values of the short and long sides of the cuboid shown in Figure 3.9 (Deb 2001).

Crowding distance for element i , for one objective is defined by the formula:

$$CD_i = \frac{f^{i+1} - f^{i-1}}{f_{max} - f_{min}} \quad (3.1)$$

To illustrate the calculation, 4 sample values from Table 3.1 will be used. The values are from elements f, 2, c and 6 (Table 3.3). These 4 elements form 3rd front when all 12 values are applied non-dominated sorting. An additional data needed is the minimum and maximum values of f_1 and f_2 . They are provided as below. (Deb 2001)

$$f1_{min} = 0.1 \quad f1_{max} = 1 \quad f2_{min} = 0 \quad f2_{max} = 60$$

As the procedure says, extreme elements are given a crowding distance of infinity. Extreme elements are defined as having $\min(f_1) - \max(f_2)$ and $\min(f_2) - \max(f_1)$. They are elements f and 6.

For element 2, crowding distance is calculated from its neighbouring elements f and c. It is the sum of cuboid side lengths covering 2:

$$CD_2 = \frac{0.51 - 0.24}{(1 - 0.1)} + \frac{8.54 - 6.51}{(60 - 0)} = 0.334$$

Similarly, 2 and 6 are considered for element c:

$$CD_c = \frac{0.83 - 0.43}{(1 - 0.1)} + \frac{6.79 - 4.23}{(60 - 0)} = 0.487$$

Table 3.3. Sample Elements for Crowding Distance Calculations

Solution	f1	f2	Crowding Distance
f	0.24	8.54	∞
2	0.43	6.79	0.334
c	0.51	6.51	0.487
6	0.83	4.23	∞

Finally, the members in the same front are sorted in a descending way according to their crowding distance values.

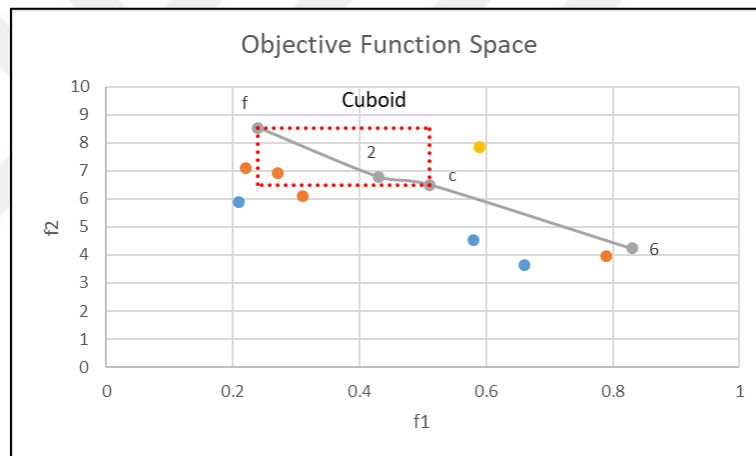


Figure 3.9. Crowding Distance Calculation Displayed as Side Lengths of a Cuboid

Depending on the values, c has a greater crowding distance therefore it is more valuable than element 2 even if they are in the same pareto optimal front. The reason is that c is in a less crowded region and has a bigger chance to preserve diversity.

3.3.3. Constrained Tournament Selection

In the methodology of NSGA-II, binary tournament selection is recommended as a selection procedure. In single objective optimization, binary tournament selection takes place by randomly selecting two members and comparing them according to their fitness values. Member having better fitness wins the tournament and is selected as the first parent for crossover operation. Same procedure is repeated to obtain the second parent and then they are mated in the crossover phase.

Usual binary tournament selection is further developed to make it compatible with constraints. The development includes a three-rule principle checking elements' feasibility, front number and crowding distances. The checks are performed just after selecting two random elements for tournament. These rules are as follows:

1. If two feasible members are entering the tournament;
 - a. If they have different fronts;
 - Member having lower front wins the tournament.
 - b. If they are in the same front;
 - Member having bigger crowding distance wins the tournament.
2. If one feasible one infeasible member enters the tournament;
 - Feasible member wins the tournament.
3. If two infeasible members enter the tournament;
 - Member having lower overall constraint violation wins the tournament.

Constraint handling and overall constraint violation calculation is explained in Section 3.3.4.

For one crossover operation two parent members are required. Then, above procedure should be run two times to select two parents for mating.

3.3.4. Constraint Handling and Constrained Domination

Deb gives a classification of constraint handling methods in genetic algorithms as follows (Deb 2000):

1. Methods based on preserving feasibility of solutions
2. Methods based on penalty functions
3. Methods making distinctions between feasible and infeasible solutions
4. Methods based on decoders
5. Hybrid methods

In the same article, a constraint handling method belonging to both second and third categories is proposed which is also the same constraint handling method in the article where NSGA-II is defined (Deb et al. 2002).

According to the proposed method, tournament selection operator is used where two solutions are compared at a time, the below criteria is applied:

- Any feasible solution is preferred to any infeasible solution.
- Between two feasible solutions, the one having a better objective function is preferred.
- Between two infeasible solutions, the one having a smaller constraint violation is preferred.

This method provides a penalty parameterless constraint handling strategy. However, it is suitable for single-objective optimizations because of the second criteria. To overcome this situation, Deb proposed to compare crowding distances described above. Therefore, between two feasible solutions, the one having a higher crowding distance is preferred.

In addition, the definition of domination needs to be revised in order to adopt multiobjective optimizations. Domination concept is turned into constrained-domination. It is defined as follows (Deb et al. 2002):

A solution i is said to constrained-dominate a solution j , if any of the following conditions is true:

- Solution i is feasible and solution j is not.
- Solutions i and j are both infeasible, but solution i has a smaller overall constraint violation.
- Solutions i and j are feasible and solution i dominates solution j .

Lastly, a solution is said to be feasible if it satisfies all the constraints. If one or more constraints are violated, the solution becomes infeasible and treated accordingly with the above rules.

After these definitions, the calculation concepts of overall constraint violations can be explained. In the above articles, Deb uses the term “overall constraint violation” and in his book (Deb 2001) gives numerical examples for these calculations.

Before the calculation, if the constraints are equality type of constraints, they need to be changed into inequality type of constraints:

$$g(x) \geq h \quad (3.2)$$

These types of constraints are transformed into below formulation by dividing both sides by h and making the inequality greater than or equal to zero:

$$\frac{g(x)}{h} - 1 \geq 0 \quad (3.3)$$

When all constraints are transformed into this form, they are normalized and they become summable with other constraints.

Overall constraint violation (OCV) term then becomes sum of all the normalized constraint violations (Deb 2001).

Consider, two more constraints are introduced to the sample optimization problem defined in Section 3.2. The constraints are given below and plotted in Figure 3.10:

$$-x_2 + 9x_1 \geq 1$$

$$x_2 + 9x_1 \geq 6$$

The constraints are normalized as described making formulations for overall constraint violations:

$$OCV_1 = -x_2 + 9x_1 - 1$$

$$OCV_2 = \frac{x_2 + 9x_1}{6} - 1$$

After these constraints, the objective function space becomes like in Figure 3.11. Some of the members become infeasible with the defined constraints. New pareto curve is shown with a continuous line, while original one is shown with a dashed line. Members 1, 2 and 3 are now in the infeasible space. Overall constraint violation values now become an indicator of infeasibility. Infeasible members having lower overall constraint violations are preferred against infeasible members having higher violations as defined in Section 3.3.3.

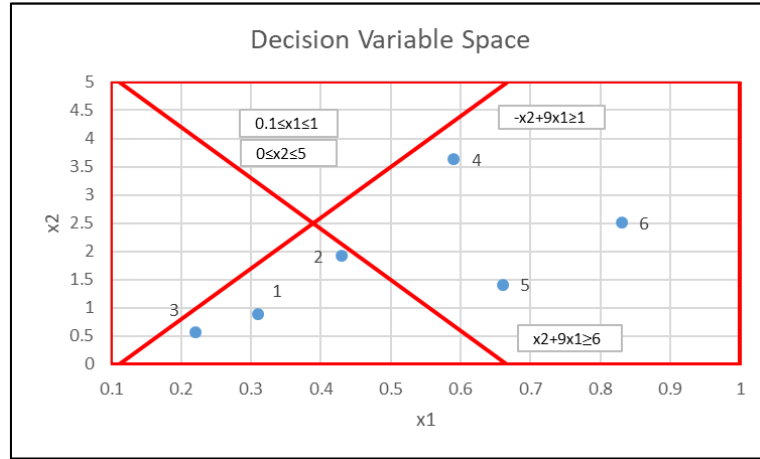


Figure 3.10. Sample Decision Variables with Constraints

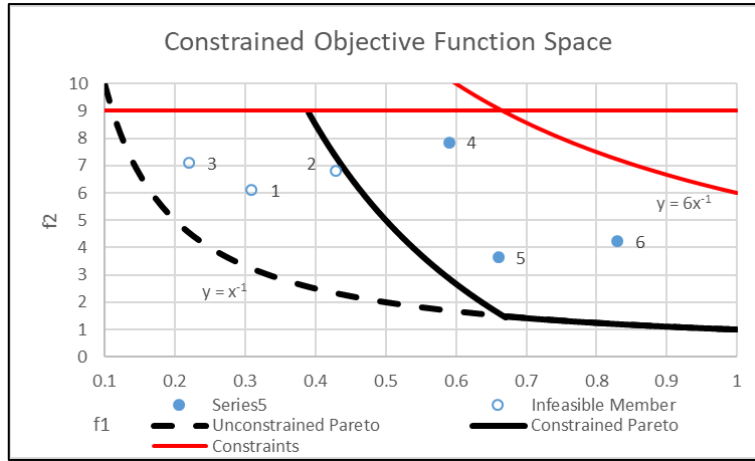


Figure 3.11. Sample Feasible and Infeasible Objective Function Values

Overall constraint violation for elements are shown below:

For element 1;

$$OCV_1 = -0.89 + 9 \times 0.31 - 1 = 0.9 \geq 0$$

$$OCV_2 = \frac{0.89 + 9 \times 0.31}{6} - 1 = -0.386 < 0$$

For element 2;

$$OCV_1 = -1.92 + 9 \times 0.43 - 1 = 0.95 \geq 0$$

$$OCV_2 = \frac{1.92 + 9 \times 0.43}{6} - 1 = -0.035 < 0$$

For element 3;

$$OCV_1 = -0.56 + 9 \times 0.22 - 1 = 0.42 \geq 0$$

$$OCV_2 = \frac{0.56 + 9 \times 0.22}{6} - 1 = -0.576 < 0$$

As seen, some constraints are not violated having OCVs greater than zero. Those values are considered as zero violation, and for every member OCV values are added to obtain a total OCV given in Table 3.4.

Table 3.4. Sample Values for Constraint Violations

Solution	x1	x2	f1	f2	OCV ₁	OCV ₂	OCV _{Tot}
1	0.31	0.89	0.31	6.1	0.386	0.000	0.386
2	0.43	1.92	0.43	6.79	0.035	0.000	0.035
3	0.22	0.56	0.22	7.09	0.576	0.000	0.576
4	0.59	3.63	0.59	7.85	0.000	0.000	0.000
5	0.66	1.41	0.66	3.65	0.000	0.000	0.000
6	0.83	2.51	0.83	4.23	0.000	0.000	0.000

According to the results, member 2 has the lowest constraint violation therefore in a case of comparison with members 1 or 3, it will be preferred.

3.3.5. Crossover

Crossover is the main operator to create offspring solutions from parent solutions in NSGA-II methodology. There are many types of crossover for genetic algorithms but only two of them which are included in this thesis study will be investigated under this section. These are binary crossover and simulated binary crossover.

3.3.5.1. Binary Crossover

To perform binary genetic operations to a population, the members have to be defined in a binary manner. For example, 8 commercial pipe diameters can be defined in a binary coding as follows (Table 3.5). For any number of possible decision variables, 2^n is the length of the binary representation called string where n represents number of bits. In this sample case n is 3.

Table 3.5. Sample Binary Coded Pipe Diameters

Diameter (mm)	Bit 1	Bit 2	Bit 3	String
100	0	0	0	000
150	0	0	1	001
200	0	1	0	010
250	0	1	1	011
300	1	0	0	100
350	1	0	1	101
400	1	1	0	110
450	1	1	1	111

Using binary coded strings, a pipe network having four pipes can be represented (Table 3.6).

Table 3.6. Sample Binary Coded Pipe Diameters

Network	Pipe 1	Pipe 2	Pipe 3	Pipe 4	Chromosome
1	100	200	150	400	000010001110
2	300	250	450	350	100011111101
3	150	350	300	300	001101100100
4	100	450	150	450	000111001111

After selection phase, two members are chosen for crossover. Let's assume those two members are members 1 and 4. For a single point binary crossover, a random number less than or equal to the length of binary chromosome is defined. This upper bound is 12 in this sample case. Let's assume that random number came out to be 5. It means that, first parent's chromosome will be cut at 5th bit, and second parent's chromosome will be cut from 7th bit. Then they will be exchanged to preserve the original chromosome length and two new members will be formed. The parts that are going to be exchanged in between are shown bold and exchanged parts are shown

in italic in Table 3.7. As a result, two new child members are created and are given in Table 3.7.

Table 3.7. Crossover Between Two Members and Resulting Child Members

Network	Pipe 1	Pipe 2	Pipe 3	Pipe 4	Chromosome
1	100	200	150	400	0000 10001110
4	100	450	150	450	00011100 1111
C1	250	400	150	400	011110001110
C2	100	450	100	150	000111000001

In this example, a single point binary crossover is illustrated. Also crossing is performed with the chromosome beginning of first element and chromosome ending of the second element. Any other way for the given decisions can also be possible for crossovers. More than one-point crossovers or exchanging different portions of member's chromosomes are all possible.

3.3.5.2. Simulated Binary Crossover

For decision variables difficult to code in binary format, they can be used as real numbers without any coding. For such cases simulated binary crossover can be used to produce two real valued children members from two real values parent members. Three distinct binary simulated crossover types are defined in the literature (Deb and Agrawal 1995). They are;

- Contracting Crossover
- Expanding Crossover
- Stationary Crossover

In contracting crossovers, two child members having result values between entering parent values are formed. It means that resulting values are closer to each other when compared to entering values.

In expanding crossovers, two child members having result values outside of entering parent values are obtained. It means that resulting values are farther than each other when compared to entering values.

In stationary crossovers, two child members having exact same values of the parents are formed.

To apply simulated binary crossovers, following formula is used. In this formula, x_1 and x_2 are entering parent values, y_1 and y_2 are resulting child values. “ α ” is a random value defining the crossover type.

$$y_1 = \alpha \cdot x_1 + (1 - \alpha) \cdot x_2 \quad (3.4)$$

$$y_2 = \alpha \cdot x_2 + (1 - \alpha) \cdot x_1 \quad (3.5)$$

Various outcomes for different α values are summarized below:

- $\alpha = 1$ $y_1 = x_1$ and $y_2 = x_2$
- $\alpha = 0$ $y_1 = x_2$ and $y_2 = x_1$
- $0 < \alpha < 1$ $y_1 > x_1$ and $y_2 < x_2$
- $\alpha < 0$ $y_1 > x_2$ and $y_2 < x_1$
- $\alpha > 1$ $y_1 < x_1$ and $y_2 > x_2$

Let's assume that, members 5 and 6 from Table 3.1 are chosen for crossover and a contracting binary simulated crossover is applied to them. Assume that random α value came out to be 0.2353 for x values and 0.3364 for y values. Two child members b and b' are formed from the parent members with decision variables b_x , b_y , b'_x and b'_y .

Therefore, the values are calculated as:

$$b_x = 0.2353x0.66 + (1 - 0.2353)x0.83 = 0.79$$

$$b'_x = 0.2353x0.83 + (1 - 0.2353)x0.66 = 0.7$$

$$b_y = 0.3364x1.41 + (1 - 0.3364)x2.51 = 2.14$$

$$b'_y = 0.3364x2.51 + (1 - 0.3364)x1.41 = 1.78$$

Similarly let's take members 1 and 2 from Table 3.1 and apply an expanding binary simulated crossover to them. Assuming a random α value for 1.8333 for x values and 1.6311 for y values two child members a and a' are formed from the parent members with decision variables a_x , a_y , a'_x and a'_y .

Therefore, the values are calculated as:

$$a_x = 1.8333x0.31 + (1 - 1.8333)x0.43 = 0.21$$

$$a'_x = 1.8333x0.43 + (1 - 1.8333)x0.31 = 0.53$$

$$a_y = 1.6311x0.89 + (1 - 1.6311)x1.92 = 0.24$$

$$a'_y = 1.6311x1.92 + (1 - 1.6311)x0.89 = 2.57$$

3.3.6. Mutation

Similar to crossover operations, mutation can be performed for binary real values. The main idea behind mutation is to add a small noise to a small number of elements in the population to avoid members stuck in local optimum regions. Adding small noises help those stuck elements to move another part of the search space and continue searching from there.

3.3.6.1. Binary Mutation

Binary mutation is performed by changing bits from zero to one or from one to zero in a population member. Before application mutation rate should be defined.

Mutation rate is usually a very small number not to affect big portions of the population. Let's assume only one bits of mutation will be performed in the 4-member population given in Table 3.6. That is a mutation rate of 0.0.208 (1/48). Then a random number is generated between 1 and 48 to determine which bit will be mutated. Let's assume the random number came out to be, 22. Then, starting from first bit of first member, 22nd bit corresponds to a one in the second member highlighted in Table 3.8. That bit is turned into zero to form a mutated element named M2. The operation caused the diameter of third pipe in network 2 change from 350 to 150. A totally different member is created with a small change.

Table 3.8. Sample Binary Coded Pipe Diameters

Network	Pipe 1	Pipe 2	Pipe 3	Pipe 4	Chromosome
1	100	200	150	400	000010001110
2	300	250	450	350	100011111 1 01
3	150	350	300	300	001101100100
4	100	450	150	450	000111001111
M2	300	250	450	150	100011111 0 01

3.3.6.2. Real Mutation

In real mutation, similar to simulated binary crossover, decision variable values are changed. The change is performed by using the following formula.

$$y = x + \sigma \cdot r \quad (3.6)$$

In this formula x is the parent member's decision variable value, y is the mutated decision variable value, r is a random value between -1 and 1 defining the strength of the mutation amount and σ is the step size which defines the maximum amount of change permitted for a specific decision variable through mutation. For negative r values entering value will be decreased, and for positive r values entering value will

be increased as a result of mutation. For $r=-1$ and $r=1$, maximum mutation effect will be applied changing the entering value with the maximum permissible amount.

Assuming a σ value of 0.05 and a random number of 0.2837, x value of member a from Table 3.1 will become;

$$y = 0.21 + 0.05 \times 0.2837 = 0.224$$

3.3.7. Pruning and Clustering

After applying crossover and mutation to a parent population of size N , a child population of the same size is obtained. These two populations are combined and sorted by the methods described in Sections 3.3.1 and 3.3.2. Next step is to reduce doubled population size to its original size by pruning. In original formulation of NSGA-II (Deb et al. 2002) pruning is realized by discarding the second half of the population after non-dominated and crowding distance sorting. Above summarized steps can be seen in Figure 3.12. It can be seen in this sample procedure, 3rd front is the front that did not fit into population size, so some of its elements are discarded.

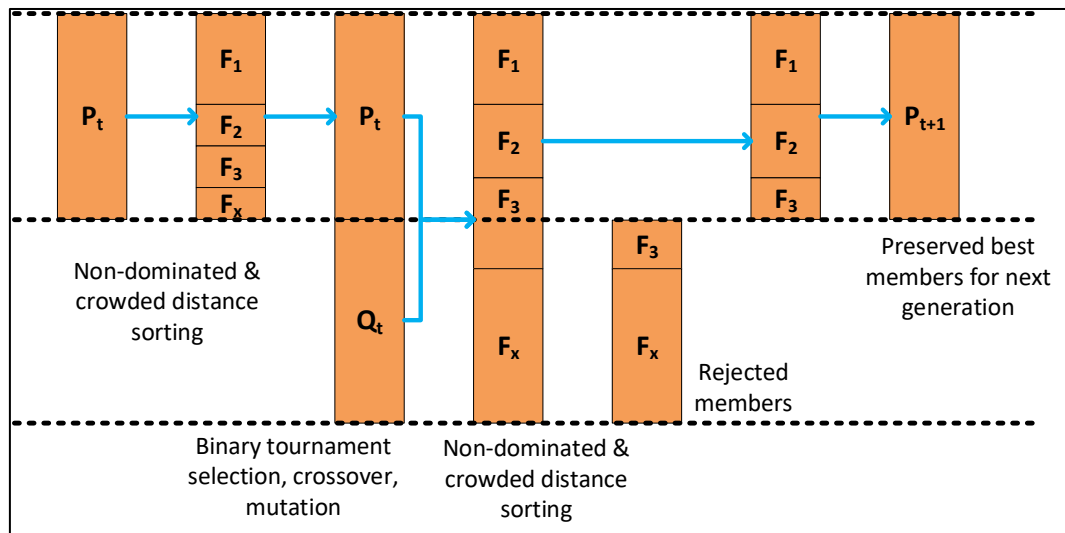


Figure 3.12. NSGA-II Steps

In a variation of NSGA-II mentioned in Section 2.5. (Deb et al. 2003), a clustering concept is introduced to achieve a better diversity among solutions. This concept is used in this thesis study and will be explained in this section.

To improve diversity, clustered NSGA-II is implemented. The implementation only changes the pruning behaviour of original NSGA-II, therefore, it is applied only to the last front that does not fit into the original population size (i.e. front 3 in Figure 3.12).

Let's assume the population in Figure 3.12 is of size N , number of members in fronts 1, 2 and 3 are N_{F1} , N_{F2} and N_{F3} respectively. Also let's assume the number of vacancies for members remaining after first and second fronts in the population is N_V . Following expressions can be written:

$$N = N_{F1} + N_{F2} + N_V \quad (3.7)$$

$$N_{F3} > N_V \quad (3.8)$$

Clustering operations are performed to choose N_V elements out of N_{F3} elements among elements in the 3rd front. At the initial stage, every member in 3rd front is defined as separate clusters. Their in-between normalized Euclidian distances are calculated and archived. Depending on these distances, nearest two members are combined in a new cluster, and the two elements are temporarily removed from the front. In-between distances are updated considering centroids of clusters this time. Clusters having the smallest distances are again combined to form clusters as above. The cluster combining operation is performed until the number of clusters are just enough to fit into the vacant spaces in the population. That is $N_{F3} - N_V$ many times when above definitions are considered. In the final stage, temporarily removed elements are taken into consideration again and this time shortest normalized Euclidian distances among final cluster centroids and original elements are looked for. This is performed to find representative elements for each final cluster. Elements

that are chosen representative for final clusters are moved to the vacant slots in the population and clustered pruning operation is finalized.

The above described procedure is further visually explained by using values from Table 3.1. Considering a last front that did not fit into the population having four members (Figure 3.13). Two members are required to fill in vacant slots from this front. Therefore, clustering mechanism should be run twice to decrease the member count from four to two. Every member is marked to form separate clusters in the initial stage. Members 3 and d are combined to form another cluster (C-3d) as they have the minimum normalized Euclidian distance between them. Then computing in-between distances between C-3d centroid, 1 and b it is found that minimum distance is between C-3d centroid and 1. Therefore, they are combined to form a new cluster (C-3d1, Figure 3.14). As there remain only two clusters, clustering algorithm stops and checks distances between final cluster centroids and the elements they cover. Element b is naturally chosen as it is the only element in its cluster. In cluster C-3d1 (Figure 3.15), nearest element to the centroid is found as element d, so it is chosen as the representative solution. As a result, elements d and b move to the vacant places in the population and they are passed into next generations. Other elements are discarded.

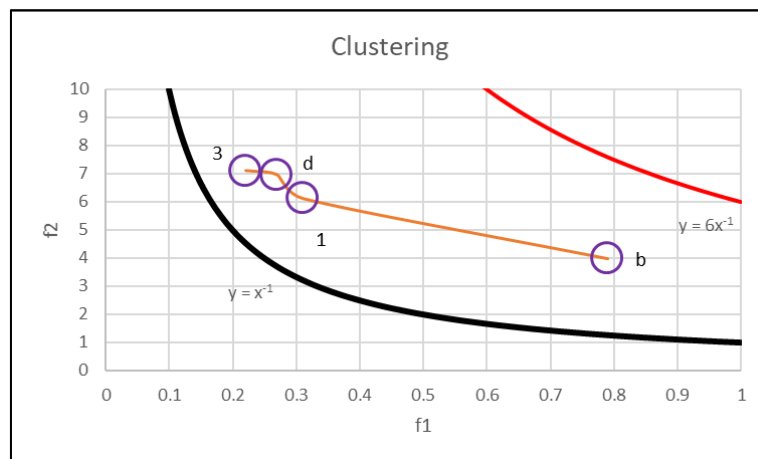


Figure 3.13. Clustering Example Initial Stage

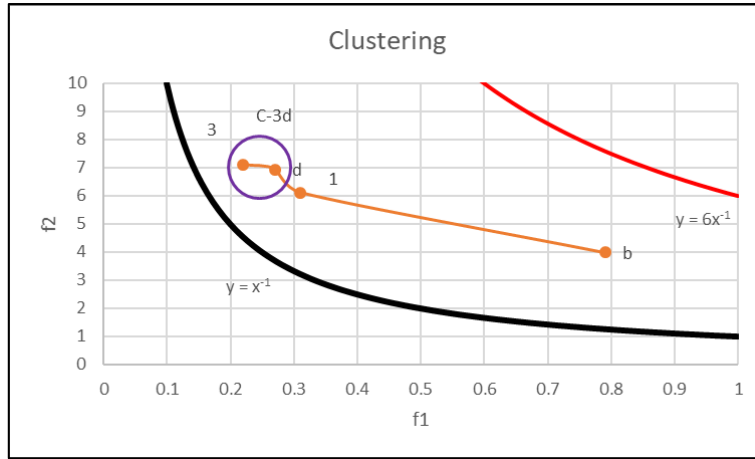


Figure 3.14. Clustering Example Intermediate Stage

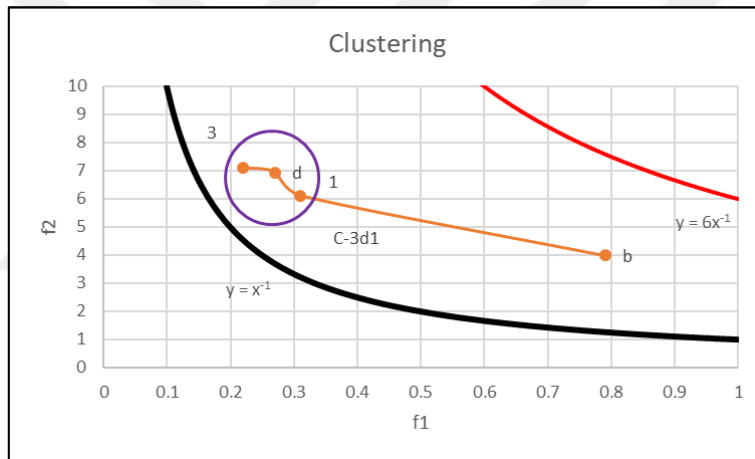


Figure 3.15. Clustering Example Final Stage

3.4. Performance Determination

Some performance measurement methods for multiobjective optimizations are summarized in Section 2.6. Among them hypervolume or S-Metric is chosen for application in this thesis. As described, hypervolumes are said to measure convergence and diversity at the same time.

For application a final pareto optimal curve with its elements and the whole archive of elements throughout the optimization runs are needed. After the optimization runs are completed, the archive is searched to find elements having worst values for every objective. It means that, in a minimization vs minimization problem, the archive is scanned to find two elements, one having maximum value for the first objective, and the other having maximum value for the second objective. These two elements are used to generate a virtual, non-existing worst element for the problem. After generating the worst element, it is orthogonal lines are drawn in every objective direction from this worst element. In addition, orthogonal lines are drawn from the elements constituting pareto optimal curve. These orthogonal lines are intersected to form a hypervolume in the objective function space. The volume of this hypervolume is used to compare different pareto curves obtained by different runs. The bigger hypervolume means a better coverage in the objective function space, therefore making a higher performance specific result.

To illustrate hypervolume, three elements (a, e and 5) that form a pareto curve from Table 3.1 will be used. In addition, an assumption is made to have two elements throughout the optimization runs, w1 and w2. They are the elements having worst objective function values in f1 and f2. w1 and w2 are used to generate the worst element, W. For this, f1 value of w2 and f2 value of w1 are used to generate the non-existent point W. Afterwards W and elements a, e, 5 are used to create the hypervolume by drawing orthogonal lines in each objective direction. Resulting hypervolume is shown with a dotted polygon in Figure 3.16.

To calculate the area of this shape, all four points have to be normalized in both objectives. It should be noted that, normalization of objectives and determining the reference worst point W greatly affects the shape and area of the formed hypervolume.

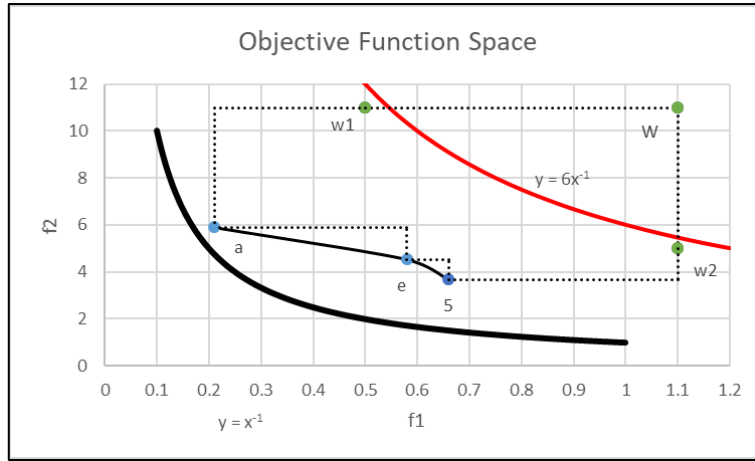


Figure 3.16. Sample Hypervolume Formation

In Table 3.9, normalized values of the four points used to create the hypervolume are given. Normalization is done by using the following formula:

$$Norm_{fi} = \frac{f_i}{f_{max} - f_{min}} \quad (3.9)$$

f_{max} and f_{min} values are taken as 1 and 0.1 for f_1 , 60 and 0 for f_2 . With the normalized values, hypervolume area is 0.1044 units.

Table 3.9. Normalized Values of Points used to Create Hypervolume

Point	f1	f2	Norm f1	Norm f2
a	0.21	5.9	0.2333	0.0983
e	0.58	4.52	0.6444	0.0753
5	0.66	3.65	0.7333	0.0608
W	1.1	11	1.2222	0.1833

3.5. Decision Making

After completing the optimization study, final elements forming a pareto curve are obtained. The next step is to choose among the equal optimal solutions. This procedure is usually named as decision making. Many methods are present in the literature for decision making analysis. They are mostly dependent on the preferences of the decision maker who is the owner of the problem. However, in the absence of decision maker's preferences, a simple method can be used to offer the decision maker as a preferred solution among the pareto optimal solutions. That method is compromise programming (Yu 1973; Zeleny 1973).

To apply compromise programming, a final set of solutions and a "best point" are required. "Best point" is a non-existent, virtual point like the worst point defined in Section 3.4. This time, from the archive of solutions or from the true pareto curve solutions, two members are picked. They are the extreme solutions in the set, each one having the best value in one of the objectives. For example, in a minimization vs minimization problem, one best element is the element having lowest first objective value, another best element is the element having lowest second objective value. These two elements are used to generate the best element. After generating the virtual best element, solutions in the final pareto set are compared with the best element in terms of in between normalized Euclidian distances. Member having the minimum distance means the nearest element to the best point is chosen as the preferred member.

To illustrate, same final solution set used in Section 3.4 is used (Figure 3.17). Original positions and normalized positions are given with in between distances in Table 3.10.

Accordingly, the nearest element to the best point B is element a. With compromise programming, in the absence of any preference of results solution a is selected as the preferred member.

Table 3.10. Normalized Values of Points used for Compromise Programming

Point	f1	f2	Norm f1	Norm f2	Distance
a	0.21	5.9	0.2333	0.0983	0.14694
e	0.58	4.52	0.6444	0.0753	0.53651
5	0.66	3.65	0.7333	0.0608	0.623761
B	0.1	1	0.1111	0.0167	–

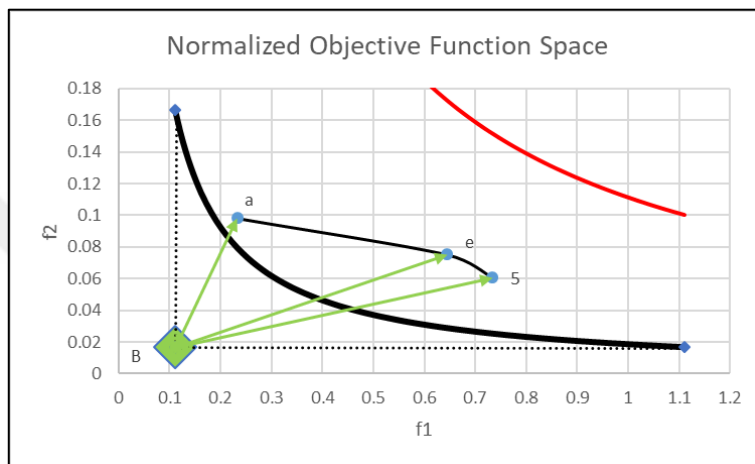


Figure 3.17. Possible Solutions Compared with the Best Point

CHAPTER 4

OBJECTIVES AND CONSTRAINTS

This chapter covers the objectives and constraints used to formulate the multiobjective optimization study done within the scope of this thesis. Objectives and constraints are given below and they will be explained in detail in this Chapter. In Section 4.1 resilience theory is explained with a simple example water distribution network. First objective of this study, network resilience is formulated. In addition, the effect of adding storage tanks to resilience is investigated. In Section 4.2 second objective, cost is explained with sample figures from the literature. Section 4.4 covers the constraints that are used in this thesis. The constraints will be formulated according to concepts defined in Section 3.3.4. Again, example calculations are performed using the simple network defined in this chapter.

$$\min (\text{Cost}) = \min \left(\sum_{i=1}^{np} L_i \times u_{pipe}(D_i) + \sum_{i=1}^{nt} u_{tank}(V_i) \right) \quad (4.1)$$

$$\max(I_n) = \max \left(\frac{\sum_{i=1}^{nj} C_i q_i^* (h_i - h_i^*) + \sum_{t=1}^{nt} C_t q_t^* (h_t - h_t^*)}{\sum_{k=1}^{nr} Q_k H_k + \sum_{j=1}^{npu} P_j + \sum_{t=1}^{nt} Q_t H_t - \sum_{i=1}^{nj} q_i^* h_i^* - \sum_{t=1}^{nt} q_t^* h_t^*} \right) \quad (4.2)$$

Subject to:

- $ev \geq capacity$
- $lv \geq capacity$
- $con \leq 0$
- $osc \leq cyc$
- $ls \leq le$

4.1. Resilience

Resilience is defined as “the capability of a strained body to recover its size and shape after deformation caused especially by compressive stress” (“Resilience” n.d.). In water distribution systems, it is used to define the ability to overcome water demand loads and possible failures by maintaining excess pressure in customers. This excess pressure is more than needed so it remains as a reserve potential energy to be used in peak load hours or failures (pipe breaks, repair works, pump shut downs, etc.). Therefore, water systems having more excess pressures in the demand junctions are more resilient.

As explained in Section 2.2, resilience index and its variations are one of the most widely used indexes to measure water distribution network strength.

4.1.1. Resilience Theory

Despite Todini does not clearly state, resilience theory is formulated on conservation of energy (Todini 2000). However, he formulates resilience on power terms rather than energy terms. As power is energy flow per unit time, in a defined time period power is also conserved.

Considering any system, an energy given as input, is either lost or used. Assuming the system is a water distribution system, input energy is lost by friction in pipes and remaining energy is delivered to the taps of customers for usage (Figure 4.1).

$$\text{Input Energy} = \text{Lost Energy} + \text{Available Energy} \quad (4.3)$$

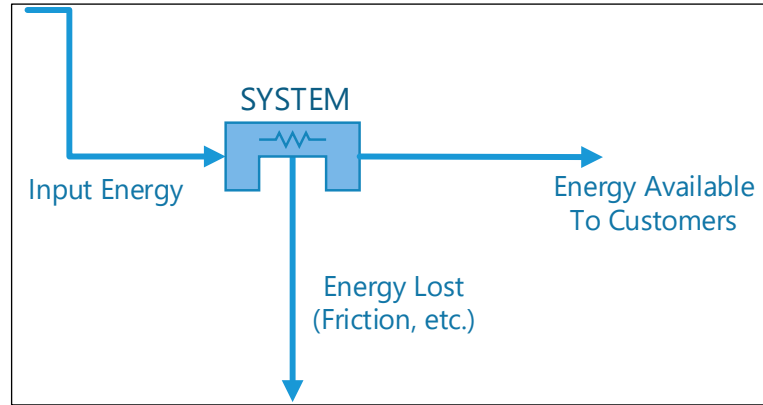


Figure 4.1. Energy Conservation in a Water Distribution System

From this basic energy conservation idea, Todini defines his power terms. These are P_{tot} , P_{int} and P_{ext} . P_{tot} is defined as the total power entering to the system, P_{int} is power dissipated internally in pipes and P_{ext} is defined as power delivered to users. Therefore, the following equation can be written.

$$P_{tot} = P_{int} + P_{ext} \quad (4.4)$$

All power terms are defined in terms of flow (q) and hydraulic grade (h) by Todini (2000). q and h terms are multiplied with each other and multiplied with specific weight of water ($\gamma = \rho \cdot g$) creating power terms. Making a dimensional analysis of Todini's (2000) formulation of power gives a term having a unit of $\text{kg} \cdot \text{m}^2/\text{s}^3$.

$$\gamma \cdot q \cdot h = \rho \cdot g \cdot q \cdot h \quad (4.5)$$

$$\frac{\text{kg}}{\text{m}^3} \cdot \frac{\text{m}}{\text{s}^2} \cdot \frac{\text{m}^3}{\text{s}} \cdot \text{m} = \frac{\text{kg} \cdot \text{m}^2}{\text{s}^3} \quad (4.6)$$

Checking the unit of power, watt gives the same unit, therefore provided power term is correct in terms of its unit.

$$watt = \frac{joule}{s} = \frac{N.m}{s} = \frac{kg.m}{s^2} \cdot \frac{m}{s} = \frac{kg.m^2}{s^3} \quad (4.7)$$

Returning back to P_{ext} and P_{int} , they are defined as follows:

$$P_{ext} = \gamma \sum_{i=1}^{nj} q_i^* h_i \quad (4.8)$$

$$P_{int} = P_{tot} - P_{ext} \quad (4.9)$$

Here, q_i^* term defines the demands of the customer junction, h_i terms define the hydraulic grade the customers have in their taps during water use. nj is the total number of demand junctions in the network, therefore P_{ext} defines the sum of power that customers have while they are consuming water.

P_{tot} is defined again in terms of flow and head. For gravity fed systems it is enough to include only outflow from the source (Q_k) and reservoir water surface height (H_k), but if any pumps are present in the system, that means another energy source for the system and their power are included separately in P_{tot} formulation (P_j). nr and npu are number of reservoir and pumps present in the system in the P_{tot} formulation below:

$$P_{tot} = \gamma \sum_{k=1}^{nr} Q_k H_k + \sum_{j=1}^{npu} P_j \quad (4.10)$$

Next another term P_{max} is defined. To define it, another term is introduced, which is h_i^* . h_i^* terms are not separately defined by Todini (2000) but they are defined as “minimum desired piezometric height value which guarantees full demand satisfaction at the same node” elsewhere (Creaco et al. 2016a). Therefore h_i^* head required at junction i , to satisfy demands at that specific junction, meaning q_i^* amount of water can be drawn without doubt. h_i^* values are predefined values either separately

for each junction or predefining an excess pressure value (p) and adding it to the elevation of junctions (z_i^*).

$$h_i^* = z_i^* + p \quad (4.11)$$

This excess pressure value (p) is a major design variable for water distribution systems. It is usually defined as “minimum operating pressure” and defined as 20 meters for settlements having less than 50000 population and 30 meters for greater populations in Turkey (“İçmesuyu Tesisleri Etüt, Fizibilite ve Projelerinin Hazırlanmasına Ait Teknik Şartname” 2013). A minimum operating pressure of 30 meters is appropriate for a good water service when residential water needs are considered. Similar figures are used by different water authorities in the world.

Therefore, P_{max} is defined as “maximum power that would be dissipated internally in order to satisfy the constraints in terms of demand and head at the nodes” (Todini 2000) and formulated as follows. It is the difference between total power and power delivered to customers satisfying their demands with just enough hydraulic energy of minimum operating pressure.

$$P_{max} = P_{tot} - \gamma \sum_{i=1}^{nj} q_i^* h_i^* \quad (4.12)$$

Finally, Todini (2000) defines the resilience index as I_r and formulates it as below:

$$I_r = 1 - \frac{P_{int}}{P_{max}} = \frac{P_{max} - P_{int}}{P_{max}} \quad (4.13)$$

Making necessary substitutions and cancelling out specific weights, final form of the resilience index is reached:

$$I_r = \frac{\sum_{i=1}^{nj} q_i^* (h_i - h_i^*)}{\sum_{k=1}^{nr} Q_k H_k + \sum_{j=1}^{npu} P_j - \sum_{i=1}^{nj} q_i^* h_i^*} \quad (4.14)$$

To express I_r in words, it is the ratio of total excess power available at nodes to the maximum power that would be internally dissipated to satisfy constraints in the system. It can be seen from the formulation, in a case where all h_i values are equal to the h_i^* values, I_r becomes 0. It means that there is just enough power in the system to satisfy demands. In a case of pipe break, or power shortage, there is no excess power so there will be disruptions in water service somehow. A system having zero resilience index means a weak system in these ways.

In another extreme case, if there occurs no loss in the system, meaning P_{tot} is equal to P_{ext} , then I_r will be equal to 1 meaning all power entering to the system is available for customer use. This relation (given below) can be obtained easily by equating nominator and denominator of I_r .

$$P_{ext} = P_{tot} \quad (4.15)$$

$$P_{ext} = \gamma \sum_{i=1}^{nj} q_i^* h_i = \gamma \sum_{k=1}^{nr} Q_k H_k + \sum_{j=1}^{npu} P_j = P_{tot} \quad (4.16)$$

One of the numerous modifications made to resilience index is done by adding a pipe diameter uniformity coefficient (Prasad et al. 2003; Prasad and Park 2004). This coefficient is calculated separately for every junction and included in the numerator of the original I_r formulation and a new index “network resilience” (I_n) is introduced.

$$I_n = \frac{\sum_{i=1}^{nj} C_i q_i^* (h_i - h_i^*)}{\sum_{k=1}^{nr} Q_k H_k + \sum_{j=1}^{npu} P_j - \sum_{i=1}^{nj} q_i^* h_i^*} \quad (4.17)$$

In Figure 4.2, a sample pipe-node relation is given. C_i values for every node is calculated by adding all pipe diameters (D_j) connected to a node (i) and dividing it by the maximum of those diameters and number of pipes (np_i) connected to that node.

$$C_i = \frac{\sum_{j=1}^{np_i} D_j}{np_i \max\{D_j\}} \quad (4.18)$$

C_i values change between 0 and 1. They are close to zero if there are too much difference between pipe diameters connected to a node. C_i values are equal to one if the pipe diameters connected to a node are all the same. This makes C_i coefficient an operator trying to equalize pipe diameters around every node, therefore creating a diameter uniformity in the network. By this, many impractical pipe diameter selections in the network are tried to be avoided at the design phase.

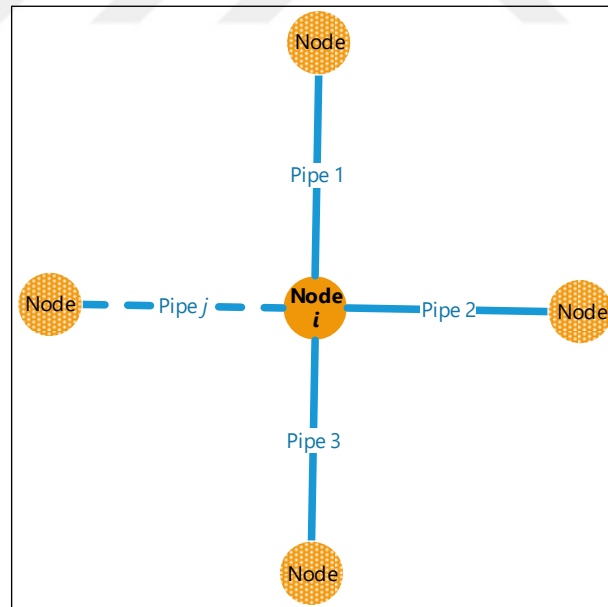


Figure 4.2. Pipe-Node Relations for C Calculations

Terms in I_n formulation are renamed and shortened as follows (Prasad et al. 2003; Prasad and Park 2004):

$$I_n = \frac{X}{X_{max}} \quad (4.19)$$

Where X is weighted surplus power and X_{max} is maximum surplus power.

To calculate I_n , flow and hydraulic grades should be known. To calculate these values, a network solver is needed. EPANET (Rossman 2000) is one of the most widely used water distribution network solvers in the world. To calculate q and h values through EPANET, network data should be entered and simulated. Network data should at least contain:

- Pipe Data
 - Start – End Nodes
 - Diameters
 - Lengths
 - Roughnesses
- Node Data
 - Elevations
 - Demands

After simulation, following results are obtained from the network solver:

- Pipe
 - Velocities & flows
 - Friction headlosses
- Node
 - Actual demands
 - Pressures & hydraulic grades

With these data and results, it becomes possible to calculate C_i and I_r , therefore I_n for any water distribution system.

4.1.2. Steady State Simulations vs Extended Period Simulations

Water distribution systems are designed to provide water to consumers to satisfy their water demands. Customer water demands are not constant through the day, month or year. Therefore, a designer should at least consider hourly demand changes throughout the day to satisfy demands.

In Figure 4.3, a 24-hour typical demand pattern is shown. This typical pattern is most widely seen in residential areas, as it contains two peaks. Peaks occur in the morning and evening, when most consumers satisfy their water needs. The demands decrease during night time as consumption decreases in a residential area. However there remains an amount of consumption during night hours, which may be night use or water leaks in the system. Given such 24-hour pattern coefficients, their average value is taken as average daily consumption and multiplied by the relevant hourly coefficient to obtain consumption specific to that hour.

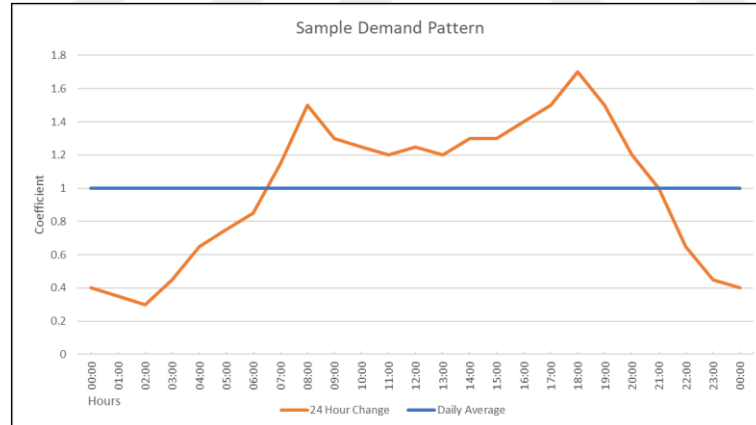


Figure 4.3. A Sample 24-Hour Demand Pattern

At this stage, it is possible to say there are two solution methods in water distribution system calculations. They are steady-state simulations and EPS. In steady state simulations, snapshot solutions for an instance of time are calculated. Therefore, a demand pattern is not needed. However, in EPS there are multiple solutions depending

on the time step of the problem. For example, in a case where a 24-hour demand pattern like Figure 4.3 is provided, it is possible to calculate demands and hydraulic grades for every hour. It makes 24 solutions for a single node.

Looking from network resilience point of view, all the input values would change when an EPS is run. This will give more than one I_n values depending on the EPS length and time step. For example, if a 24-hour EPS is run with a time step of one hour, there will be 24 different I_n values corresponding to different hours of the day. To choose a representative value among these, various statistical measures are suggested (Creaco et al. 2016b). These measures include temporal average, minimum, maximum, median and cumulative Weibull frequency. Median is used in this study.

Median is defined as:

$$median(x) = \frac{1}{2} (x_{[(n+1)/2]} + x_{[(n+1)/2]})$$

Where x is an ordered list of n numbers, $[]$ and $[]$ denote the floor and ceiling functions respectively.

4.1.3. Storage Tank Inclusion

Water storage tanks are storage structures in a water distribution system located at higher elevations than service area elevations. The purpose of including storage tanks is to create reserve volumes in the system to overcome water shortages, to maintain a reliable hydraulic energy on top of the service area and to provide volume to compensate water in peak demand hours. In Section 4.2 water storage tanks will be explained in more detail. In this section, effects of inclusion of storage on the formulation is explained. In Section 4.1.4 effects will be presented on a sample network.

When the I_n formulation is examined, it consists of values (flow and head) from water sources (reservoirs and pumps) and customers. Storage tanks can be included to I_n formulation both in source and customer parts. As water storage tanks provide water

volume for peak hour compensation, they empty in those peak hours (i.e. morning and evening peaks) and refill in low consumption hours (i.e. night hours). Therefore, water storage tanks act as water sources in peak hours and as consumption junctions in low demand hours. This indicates that, flow and head values computed for storage tanks should be included in source terms (Q_k, H_k) while emptying and in consumption terms (q_i^*, h_i) while filling. For h_i^* term, a lower value from demand junctions can be used for storage tanks. Residential customers need a minimum operating pressure (like 30 meters) for a good water service but water storage tanks are structures open to atmosphere so they do not need that much of excess pressure. Rather, a minimum pressure equal to the height of the storage tank may be chosen as h_i^* . As a result, I_n formulation can further be extended to include storage tank terms as follows:

$$I_n = \frac{\sum_{i=1}^{nj} C_i q_i^* (h_i - h_i^*) + \sum_{t=1}^{nt} C_t q_t^* (h_t - h_t^*)}{\sum_{k=1}^{nr} Q_k H_k + \sum_{j=1}^{npu} P_j + \sum_{t=1}^{nt} Q_t H_t - \sum_{i=1}^{nj} q_i^* h_i^* - \sum_{t=1}^{nt} q_t^* h_t^*} \quad (4.20)$$

In this extended formulation, nt represents number of storage tanks, q_i^* represents tank inflow amount (zero when tank is emptying), Q_t is tank outflow amount (zero when tank is filling), h_t and H_t are tank water levels and h_t^* is the head required for storage tanks (which can simply be taken as the height of the storage tank considered). C_t values will be equal to 1 if storage tanks are connected to the water distribution system with a single pipe. If water enters and leaves water tanks from separate pipes it may be necessary to calculate C_t values accordingly. With this extended formulation, it becomes possible to include storage tanks in a water distribution system.

4.1.4. Example Resilience Calculations

Network resilience concept defined in Section 4.1.1 above will be demonstrated on a sample water distribution network Figure 4.4. Sample network is gravity fed with a reservoir and has four pipes forming a loop. There are four consumer junctions. Table 4.1 gives the network data needed.

The water level in the reservoir is assumed to be constant. The junctions consume water according to the pattern given in

Table 4.2 under “Coefficient” column. EPS is performed for 24 hours with 1-hour time step. All pipes have a Hazen-Williams roughness coefficient of 130. Hydraulic simulation is performed by using EPANET (Rossman 2000). For resilience calculations formula 4.20 is used.

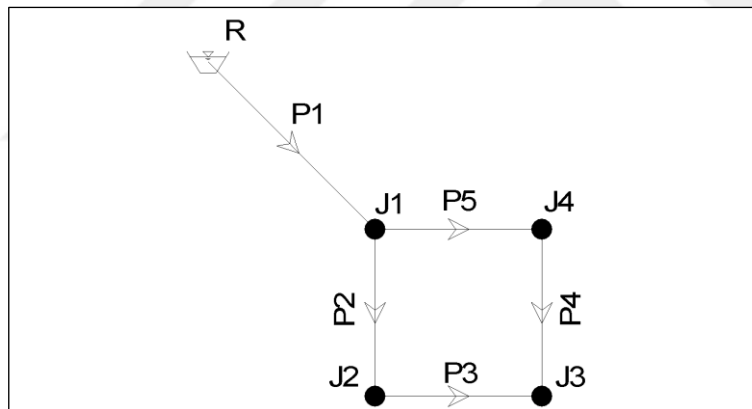


Figure 4.4. A Sample Water Distribution Network Layout

Table 4.1. Node and Pipe Data for the Sample Network

Node ID	Elevation (m)	Average Demand (l/s)	Pipe ID	Diameter (mm)	Length (m)
R	60	–	P1	200	100
J1	0	4	P2	150	100
J2	0	2	P3	100	100
J3	0	8	P4	100	100
J4	0	2	P5	150	100

Table 4.2. Demand Pattern Coefficients

Time (hr)	Coefficient	Time (hr)	Coefficient	Time (hr)	Coefficient	Time (hr)	Coefficient
00:00	0.4	06:00	0.85	12:00	1.25	18:00	1.7
01:00	0.35	07:00	1.15	13:00	1.2	19:00	1.5
02:00	0.3	08:00	1.5	14:00	1.3	20:00	1.2
03:00	0.45	09:00	1.3	15:00	1.3	21:00	1
04:00	0.65	10:00	1.25	16:00	1.4	22:00	0.65
05:00	0.75	11:00	1.2	17:00	1.5	23:00	0.45

In Table 4.3 pipe diameter uniformity coefficients for nodes are given. Results (flow and hydraulic grade) of 24-hour EPS are given in Table 4.4. In Table 4.5 network resilience values with X and $X_{\max}$ values for a 24-hour simulation are given. Minimum pressure required is chosen as 30 meters for every node.

Table 4.3. Node-Pipe Relations and Pipe Diameter Uniformities

Pipe	Diameter (mm)	From Node	To Node	Node	Ci
P1	200	R	J1	R	1.000
P2	150	J1	J2	J1	0.833
P3	100	J2	J3	J2	0.833
P4	100	J3	J4	J3	1.000
P5	150	J4	J1	J4	0.833

Table 4.4. 24-Hour EPS Results of the Sample Network

Time (hr)	Reservoir Outflows & Nodal Demands (l/s)					Calculated Hydraulic Grades (m)			
	R	J1	J2	J3	J4	J1	J2	J3	J4
00:00	16.00	4.00	2.00	8.00	2.00	59.84	59.74	59.39	59.74
01:00	14.00	3.50	1.75	7.00	1.75	59.88	59.80	59.53	59.80
02:00	12.00	3.00	1.50	6.00	1.50	59.91	59.85	59.64	59.85
03:00	18.00	4.50	2.25	9.00	2.25	59.81	59.68	59.24	59.68
04:00	26.00	6.50	3.25	13.00	3.25	59.62	59.37	58.51	59.37
05:00	30.00	7.50	3.75	15.00	3.75	59.50	59.17	58.05	59.17
06:00	34.00	8.50	4.25	17.00	4.25	59.37	58.96	57.55	58.96
07:00	46.00	11.50	5.75	23.00	5.75	58.90	58.17	55.71	58.17
08:00	60.00	15.00	7.50	30.00	7.50	58.20	57.01	52.98	57.01
09:00	52.00	13.00	6.50	26.00	6.50	58.62	57.71	54.61	57.71
10:00	50.00	12.50	6.25	25.00	6.25	58.72	57.87	54.99	57.87
11:00	48.00	12.00	6.00	24.00	6.00	58.81	58.03	55.35	58.03
12:00	50.00	12.50	6.25	25.00	6.25	58.72	57.87	54.99	57.87
13:00	48.00	12.00	6.00	24.00	6.00	58.81	58.03	55.35	58.03
14:00	52.00	13.00	6.50	26.00	6.50	58.62	57.71	54.61	57.71
15:00	52.00	13.00	6.50	26.00	6.50	58.62	57.71	54.61	57.71
16:00	56.00	14.00	7.00	28.00	7.00	58.42	57.37	53.82	57.37
17:00	60.00	15.00	7.50	30.00	7.50	58.20	57.01	52.98	57.01
18:00	68.00	17.00	8.50	34.00	8.50	57.73	56.24	51.14	56.24
19:00	60.00	15.00	7.50	30.00	7.50	58.20	57.01	52.98	57.01
20:00	48.00	12.00	6.00	24.00	6.00	58.81	58.03	55.35	58.03
21:00	40.00	10.00	5.00	20.00	5.00	59.15	58.59	56.69	58.59
22:00	26.00	6.50	3.25	13.00	3.25	59.62	59.37	58.51	59.37
23:00	18.00	4.50	2.25	9.00	2.25	59.81	59.68	59.24	59.68

Table 4.5. 24-Hour Network Resilience Values of the Sample Network

Time (hr)	J1	J2	J3	J4	X	X _{max}	I _n
00:00	99.47	49.57	235.12	49.57	433.72	480.00	0.9036
01:00	87.15	43.46	206.71	43.46	380.78	420.00	0.9066
02:00	74.78	37.31	177.84	37.31	327.24	360.00	0.9090
03:00	111.79	55.65	263.16	55.65	486.25	540.00	0.9005
04:00	160.44	79.54	370.63	79.54	690.16	780.00	0.8848
05:00	184.38	91.16	420.75	91.16	787.44	900.00	0.8749
06:00	208.04	102.57	468.35	102.57	881.52	1020.00	0.8642
07:00	276.96	134.98	591.33	134.98	1138.25	1380.00	0.8248
08:00	352.50	168.81	689.40	168.81	1379.53	1800.00	0.7664
09:00	310.05	150.10	639.86	150.10	1250.10	1560.00	0.8013
10:00	299.17	145.16	624.75	145.16	1214.23	1500.00	0.8095
11:00	288.10	140.15	608.40	140.15	1176.80	1440.00	0.8172
12:00	299.17	145.16	624.75	145.16	1214.23	1500.00	0.8095
13:00	288.10	140.15	608.40	140.15	1176.80	1440.00	0.8172
14:00	310.05	150.10	639.86	150.10	1250.10	1560.00	0.8013
15:00	310.05	150.10	639.86	150.10	1250.10	1560.00	0.8013
16:00	331.57	159.66	666.96	159.66	1317.84	1680.00	0.7844
17:00	352.50	168.81	689.40	168.81	1379.53	1800.00	0.7664
18:00	392.84	185.87	718.76	185.87	1483.34	2040.00	0.7271
19:00	352.50	168.81	689.40	168.81	1379.53	1800.00	0.7664
20:00	288.10	140.15	608.40	140.15	1176.80	1440.00	0.8172
21:00	242.92	119.13	533.80	119.13	1014.97	1200.00	0.8458
22:00	160.44	79.54	370.63	79.54	690.16	780.00	0.8848
23:00	111.79	55.65	263.16	55.65	486.25	540.00	0.9005
Min (I _n)							0.7271
Max (I _n)							0.9090
Average (I _n)							0.8327
Median (I _n)							0.8172

From the results, it can be seen I_n value changes between 0.7271 and 0.9090 during the day. The lowest value is obtained at peak hour of 18:00. Similarly, highest value is obtained at 02:00 where there is lowest consumption in the system. These values indicate that, system is at its weakest state when there is maximum load on it. The system is strong enough to handle all loads in peak hour having 51.14 meters of excess pressure (the minimum required pressure is 30 meters) on J3 which seems the most critical junction.

To investigate storage tank effects the sample network is extended with a tank connected to junction J3 (Figure 4.5). Pipe between tank T and J3 is P6 and it has a diameter of 100 mm with a short length of 1 meter. A fictitious short length is preferred to avoid losses in P6.

Storage tank has a cylindrical shape with 10-meter diameter and 5-meter maximum water height. At 00:00 tank has 1 meter of water and it is allowed to empty fully.

Keeping all other parameters same, again a 24-hour results are given in Table 4.6 and Table 4.7. Nodal demands are the same as Table 4.4 and pipe diameter uniformity coefficients are same as Table 4.3 so they are not repeated.

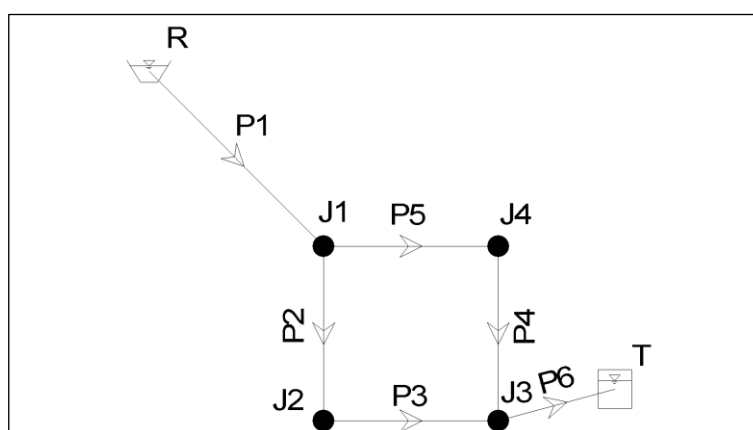


Figure 4.5. A Sample Water Distribution Network Layout with Storage Tank

Table 4.6. 24-Hour Network Resilience Values of the Sample Network with Storage

Time (hr)	X	X _{max}	I _n	Time (hr)	X	X _{max}	I _n
00:00	407.94	1871.40	0.2180	12:00	1243.77	1500.00	0.8292
01:00	388.76	1686.00	0.2306	13:00	1190.28	1440.00	0.8266
02:00	360.27	1500.60	0.2401	14:00	1285.38	1560.00	0.8240
03:00	509.90	1277.40	0.3992	15:00	1279.83	1560.00	0.8204
04:00	697.44	1092.60	0.6383	16:00	1371.39	1680.00	0.8163
05:00	789.29	1002.60	0.7872	17:00	1458.53	1800.00	0.8103
06:00	886.21	1020.00	0.8688	18:00	1634.75	2040.00	0.8013
07:00	1193.78	1380.00	0.8651	19:00	1429.78	1800.00	0.7943
08:00	1536.18	1800.00	0.8534	20:00	1148.45	1664.40	0.6900
09:00	1313.54	1560.00	0.8420	21:00	971.52	1681.80	0.5777
10:00	1253.83	1500.00	0.8359	22:00	659.70	1690.80	0.3902
11:00	1198.20	1440.00	0.8321	23:00	489.76	1596.60	0.3068
							Min (I _n) 0.2180
							Max (I _n) 0.8688
							Average (I _n) 0.6707
							Median (I _n) 0.8058

Table 4.7. 24-Hour EPS Results of the Sample Network with Storage

Time (hr)	Reservoir Outflows, Tank Inflows & Tank Outflows(l/s)			Calculated Hydraulic Grades & Tank Water Levels (m)				
	R	T (in)	T (out)	J1	J2	J3	J4	T
00:00	39.19	23.19	0.00	59.18	58.43	54.09	58.43	1.00
01:00	35.10	21.10	0.00	59.33	58.72	55.14	58.72	2.06
02:00	31.01	19.01	0.00	59.47	58.98	56.09	58.98	3.03
03:00	30.29	12.29	0.00	59.49	59.07	56.93	59.07	3.90
04:00	31.21	5.21	0.00	59.46	59.07	57.47	59.07	4.47
05:00	31.71	1.71	0.00	59.45	59.07	57.70	59.07	4.70
06:00	32.89	0.00	1.11	59.41	59.03	57.78	59.03	4.78
07:00	37.74	0.00	8.26	59.24	58.80	57.72	58.80	4.73
08:00	44.79	0.00	15.21	58.95	58.40	57.31	58.40	4.35
09:00	44.74	0.00	7.26	58.96	58.33	56.65	58.33	3.66
10:00	45.28	0.00	4.72	58.93	58.27	56.32	58.27	3.32
11:00	45.32	0.00	2.68	58.93	58.25	56.10	58.25	3.11
12:00	46.55	0.00	3.45	58.88	58.17	55.98	58.17	2.98
13:00	46.35	0.00	1.65	58.89	58.17	55.82	58.17	2.83
14:00	48.16	0.00	3.84	58.80	58.05	55.75	58.05	2.75
15:00	48.79	0.00	3.21	58.77	58.00	55.57	58.00	2.57
16:00	50.84	0.00	5.16	58.68	57.86	55.42	57.86	2.43
17:00	53.18	0.00	6.82	58.56	57.69	55.18	57.69	2.19
18:00	57.29	0.00	10.71	58.35	57.38	54.86	57.38	1.88
19:00	55.80	0.00	4.20	58.43	57.44	54.38	57.44	1.39
20:00	51.74	3.74	0.00	58.63	57.69	54.20	57.69	1.19
21:00	48.03	8.03	0.00	58.81	57.94	54.38	57.94	1.37
22:00	41.18	15.18	0.00	59.10	58.37	54.78	58.37	1.73
23:00	35.61	17.61	0.00	59.32	58.72	55.48	58.72	2.43

In Figure 4.6 it can be seen that, I_n pattern follows the demand pattern. When there are peaks in the demand pattern I_n has crests (at 08:00 and at 18:00). On the other hand, I_n pattern of the same network with storage tank does not follow demand pattern. Instead, it follows the tank water level which can be seen in Figure 4.7. When storage tank reaches maximum fullness (at 06:00), I_n also reaches its maximum value. When tank is at its lowest level (at 00:00) I_n has also its lowest value. I_n starts to decrease or increase sharply when water flow changes direction at storage tank entrance (i.e. starts filling at an empty state or starts emptying at a full state).

Another comment that should be made is that, at peak demand hour of 18:00, network with storage has bigger I_n value than network without storage. This may lead to a conclusion of adding storage tanks to networks makes water distribution networks resilient to peak hour loads. In addition, network without storage reaches its initial I_n value at the end of the day, if same daily demand pattern is present. However, network with storage reaches to a higher I_n value at the end of the day as storage tank final level is higher than initial level.

A final remark is on the statistical measures. It is seen that, network with storage has lower I_n values on all of the four statistical measures (minimum, maximum, average, median). However median of the two networks are very close to each other. This may indicate that using median I_n values for representative values is appropriate. Therefore, first objective of this study is selected as maximizing the median of network resilience values obtained from EPS solutions of the studied water distribution system.

It should be noted that an optimization study is not performed in this example calculation.

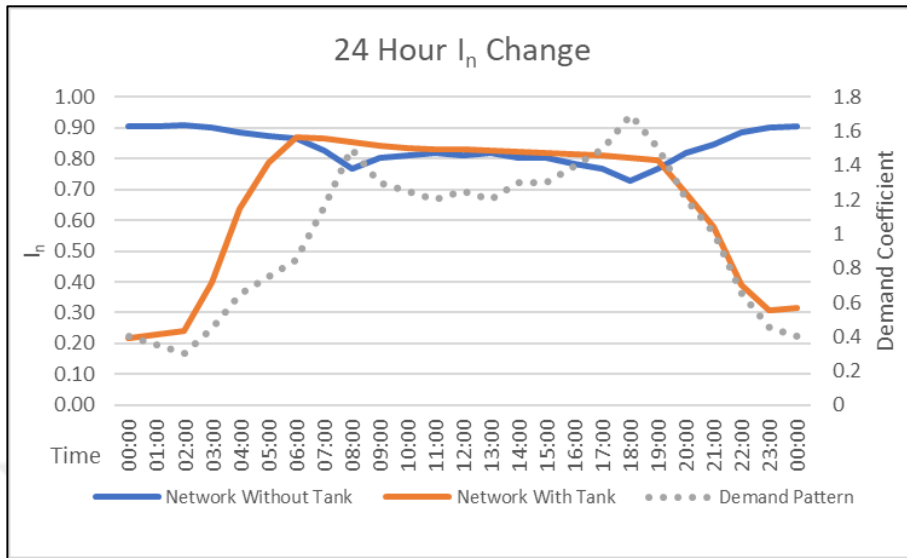


Figure 4.6. 24-hour Change of Sample I_n values

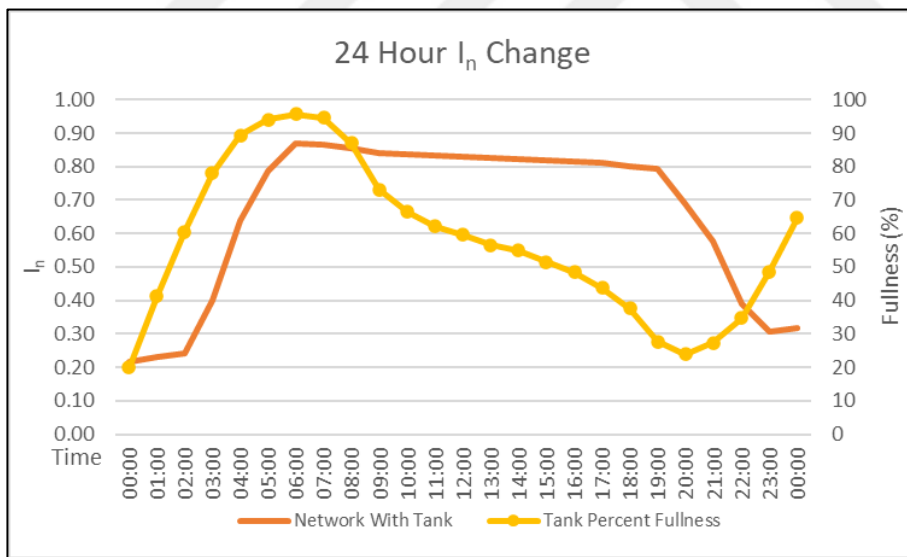


Figure 4.7. 24-hour Change of Sample I_n values and Percent Tank Fullness

4.2. Cost

Many single or multiobjective optimization studies dealing with water distribution systems consider only pipe diameter optimizations. When considering only pipe diameters, each study is provided some unit pipe costs with various diameters so it becomes easy to calculate unit construction cost for the designs.

However, when other elements in the network such as pumps and tanks are involved in the optimum design studies, number of studies become very limited. Almost all of tank and pump optimization studies are made on the network provided by Walski et.al (Walski et al. 1987). Some others used real world networks and did not provide relevant cost information. A comparative pipe unit cost table is provided from three different works from the literature in Table 4.8. 2018 Turkey market prices for ductile iron pipes are included in the same table as well. The market prices were obtained from Bank of Provinces (İller Bankası).

Among the works in the literature only Walski et.al provided a tank cost (Table 4.9). Therefore, to have consistent pipe and tank costs, figures from Walski et.al are used for the network case studies performed on other two networks (Alperovits and Shamir 1977; Fujiwara and Khang 1990, 1991). It should be noted that, Alperovits and Shamir provided some predefined unit cost values while Fujiwara and Khang (1990) provided a formula to generate unit costs for different diameters. Their formula is given below. In their work only six diameters are defined. They are included in Table 4.8. Other diameters and their costs are calculated when needed.

$$Unit\ Cost\ \left(\frac{\$}{m}\right) = b_1 \times D^\beta \quad (4.21)$$

In the study b_1 coefficient is defined as 1.1 and β is given as 1.5. D represents the diameter of the relevant pipe in inches.

Table 4.8. Unit Pipe Cost Comparison

Pipe Diameter (in)	Walski et al. (1987) (\$/m)			Alperovits and Shamir (1977) (1/m)	Fujiwara and Khang (1990) (\$/m)	Ductile Iron Pipes Market Price (2018) (TL/m)
	Urban	Residential	Development			
1	—	—	—	2.00	—	—
2	—	—	—	5.00	—	—
3	—	—	—	8.00	—	—
4	—	—	—	11.00	—	91.96
6	85.96	46.59	41.99	16.00	—	240.57
8	91.21	64.96	58.40	23.00	—	290.06
10	111.88	82.35	73.82	32.00	—	339.23
12	135.83	106.30	95.80	50.00	45.73	371.29
14	164.70	131.89	118.77	60.00	—	445.83
16	191.93	159.12	143.04	90.00	70.40	503.26
18	217.19	187.66	168.96	130.00	—	572.74
20	251.97	219.16	197.18	170.00	98.39	641.49
22	—	—	—	300.00	—	715.26
24	305.12	280.51	252.62	550.00	129.33	789.35
30	358.27	380.91	346.13	—	180.75	—
32	—	—	—	—	—	1177.14
40	—	—	—	—	278.28	1599.30

Table 4.9. Tank Costs

Volume of Tank (gal)	Cost (\$)
50,000	115,000
100,000	145,000
250,000	325,000
500,000	425,000
1,000,000	600,000

Above unit costs can be named as u_{pipe} . They are multiplied by pipe lengths L_i to calculate total pipe costs.

AWWA Manual M42 of Steel Water-Storage Tanks (*Steel Water-Storage Tanks M42 Manual of Water Supply Practices* 2013) state three different type water storage facilities. They are reservoirs, standpipes and elevated tanks (Figure 4.8). The costs given in Table 4.9 are for elevated tanks. Tank shapes for all types are assumed to be cylindrical.

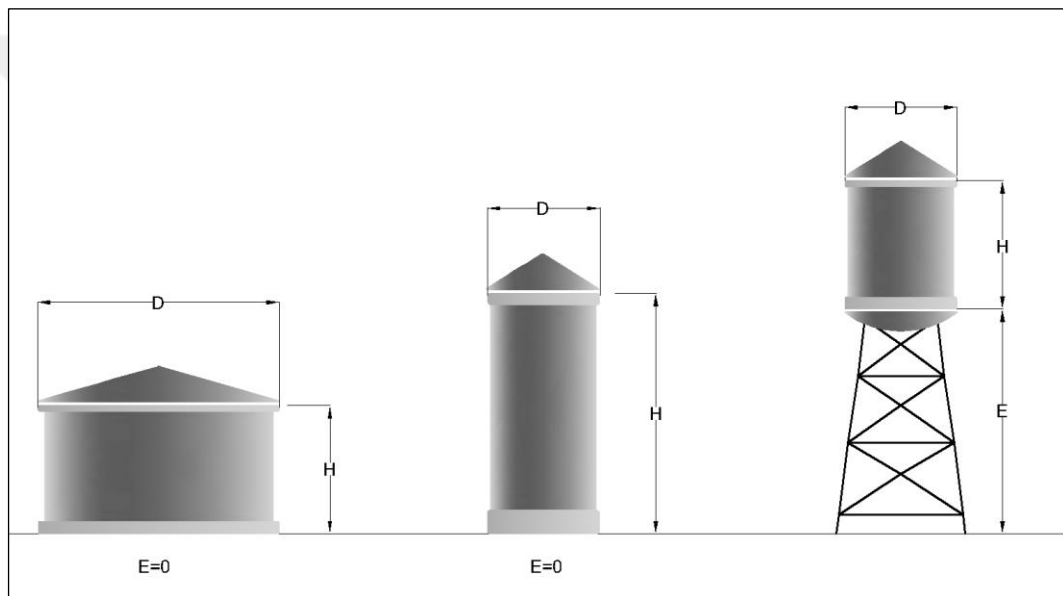


Figure 4.8. Three Different Types of Water Storage: Reservoirs, Standpipes and Elevated Tanks

In AWWA M42 manual, various typical tank sizes (height (H), diameter (D) and elevation (E)) are defined. According to those typical tanks, a range for tank sizes can be summarized in Table 4.10. These values are used throughout this study for tank sizing operations.

Table 4.10. Minimum and Maximum Tank Sizes Used in the Study

Storage Tank Type	Min H (m)	Max H (m)	Min D (m)	Max D (m)	Min E (m)	Max E (m)
Reservoir	5	15	5	65	0	0
Standpipe	10	30	4	30	0	0
Elevated Tank	3	15	5	40	10	50

Approximate relative costs for three different types of tanks are given in Figure 4.9 (*Steel Water-Storage Tanks M42 Manual of Water Supply Practices* 2013).

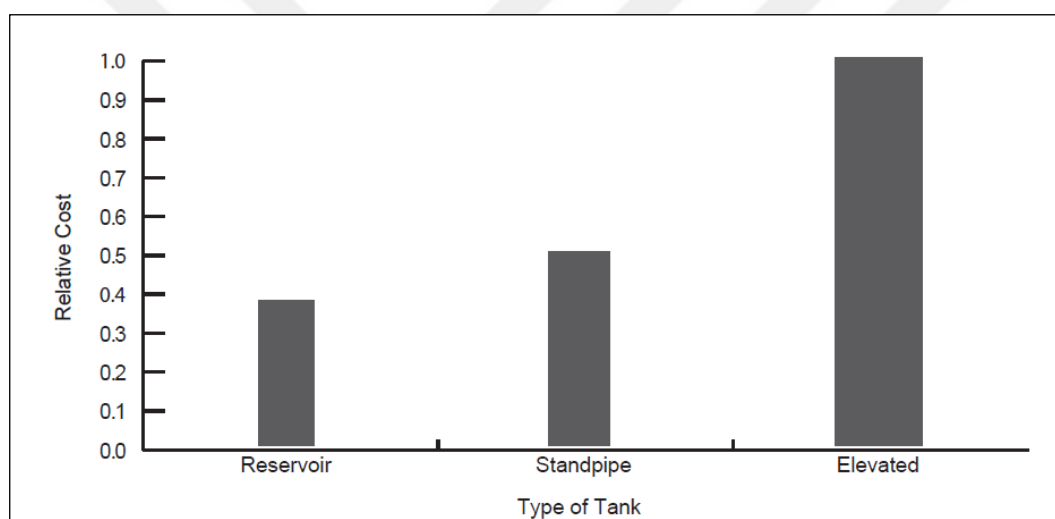


Figure 4.9. Relative Costs for Different Types of Tanks

(*Steel Water-Storage Tanks M42 Manual of Water Supply Practices* 2013)

These relative costs and costs in Table 4.9 are used to interpolate intermediate values for the three types of tanks. 3rd degree polynomials are fitted and their equations are used to calculate tank costs in the optimization study. The fitted curves are given in Figure 4.10, Figure 4.11 and Figure 4.12.

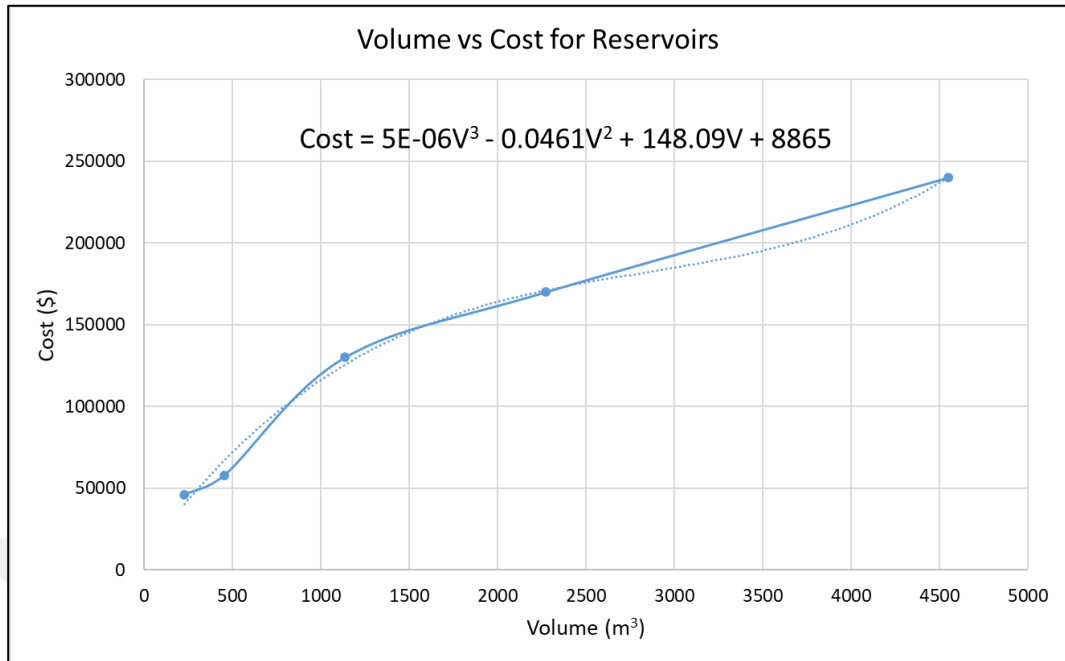


Figure 4.10. Interpolated Costs for Reservoirs

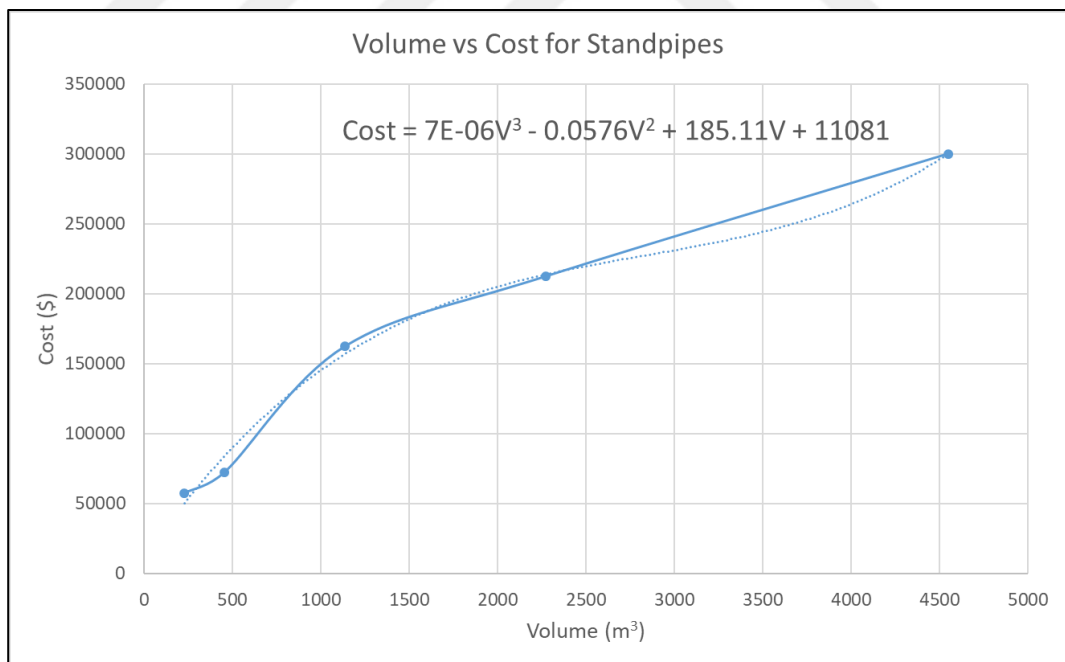


Figure 4.11. Interpolated Costs for Standpipes

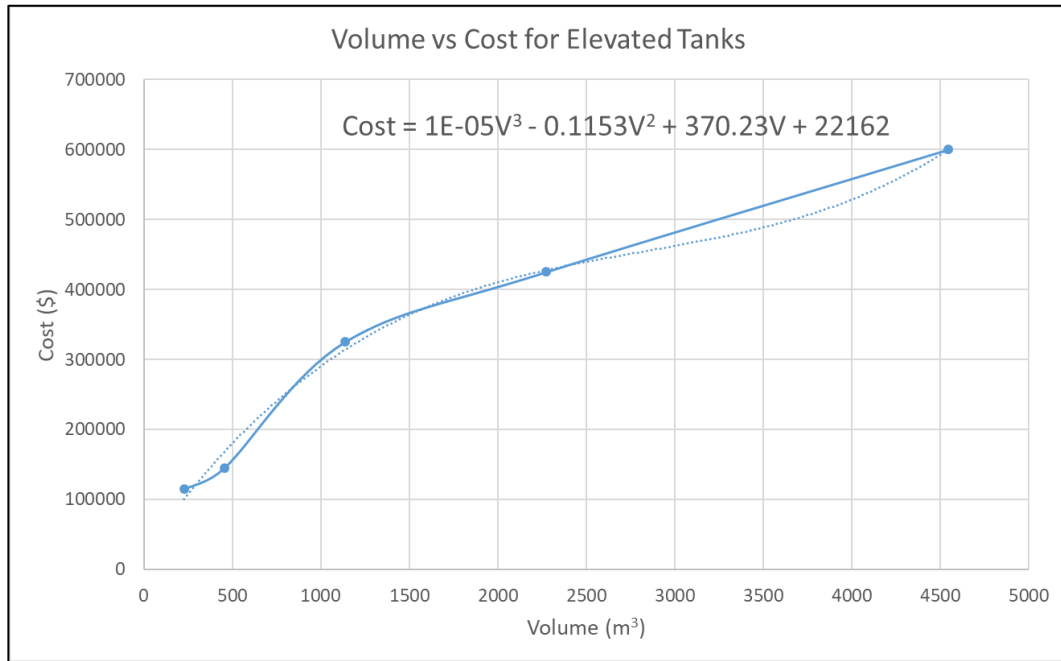


Figure 4.12. Interpolated Costs for Elevated Tanks

Similar to market price costs obtained from Bank of Provinces, storage tank construction costs are also obtained. The prices are given in Table 4.11. A 4th degree polynomial is fitted and tank costs are formulated in terms of volume. Recall that the tanks mentioned here are prismatic shaped reinforced concrete structures that are different than tanks mentioned above Figure 4.13. Tank cost functions can be named as u_{tank} .

Table 4.11. Reinforced Concrete Storage Tank Costs for 2018

Volume (m ³)	Average Construction Costs (TL)	Average Mechanical Costs (TL)	Average Electrical Costs (TL)	Average Total Costs (TL)
50	116697	17066	22441	156204
100	167259	41209	26588	235056
300	256813	63811	53108	373732
500	317950	69725	54091	441766
1000	436587	71518	57365	565470
2000	681880	103974	62826	848680
3000	899812	148917	63824	1112553
5000	1373549	210434	67326	1651309
7500	1683133	270159	67399	2020691
10000	2270838	315035	75490	2661363
15000	3250197	583911	78662	3912770
20000	4927727	1015390	81864	6024981

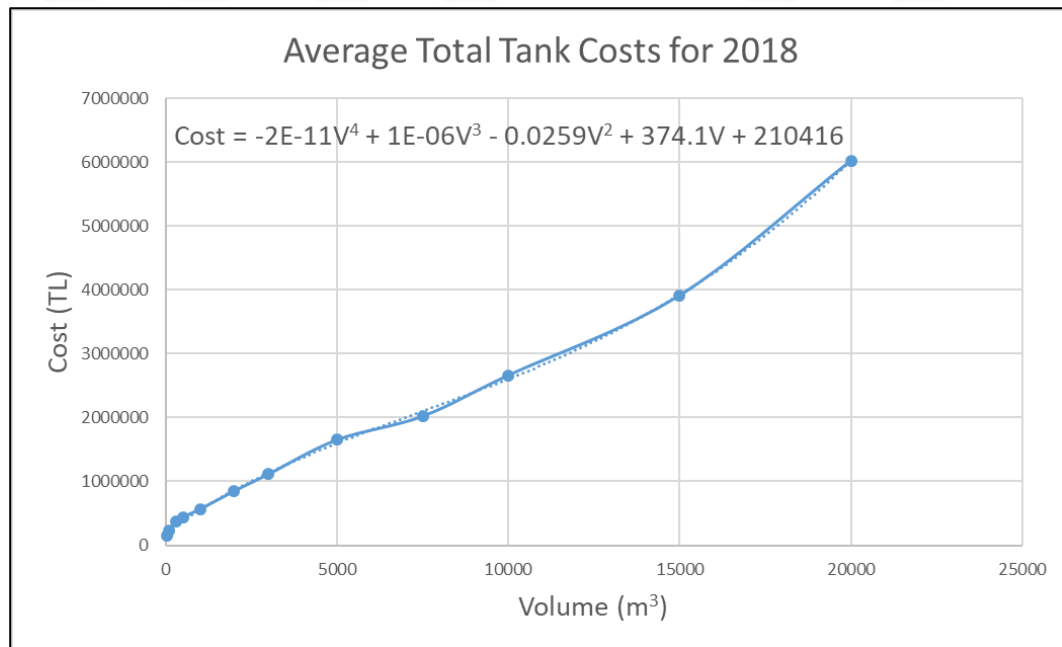


Figure 4.13. Interpolated 2018 Costs Reinforced Concrete Tanks

Although it is not a direct interest in this thesis work, pump equipment and energy costs may be present in pumped networks. Pump equipment costs may be relatively easy to compute but energy costs should be considered for some amortization years.

Sample figures are provided by Walski et.al with pipe and tank costs so those figures can be adopted for other problems (Walski et al. 1987).

Total costs are calculated simply by adding pipe costs, tank costs and pump – energy costs (if there are any) for a network. This total cost becomes the second objective which is minimization for this thesis study.

4.3. Decision Variables

5 groups of decision variables are considered in this study. They are:

- Pipe diameters (D_i), (mm)
- Pipe diameter choices (n_{pd}), (unitless)
- Storage tank sites in the network, (unitless)
- Storage tank dimensions
 - Height (H_i), (m)
 - Diameter (D_i), (m)
 - Elevation (E_i), (m)

To discretize tank dimensions tank size limitations in Table 4.10 are used. When centimetre accuracy is preferred, 1200 choices for H_i , 3500 choices for D_i , and 4000 choices for E_i are possible for selection.

It is assumed that only one tank can be added to a junction. Therefore, tank site contribution to selection set is calculated by permutation.

According to these assumptions, search space size is calculated by the following formula:

$$SS = n_{pd}^{np} \times P(nj, nt) \times H_i^* \times D_i^* \times E_i^* \quad (4.22)$$

$$SS = n_{pd}^{np} \times P(nj, nt) \times 1200 \times 3500 \times 4000$$

After simplifications;

$$SS = 1.68 \times 10^{10} \times n_{pd}^{np} \times \frac{nj!}{(nj - nt)!} \quad (4.23)$$

4.4. Constraints

A real-world multiobjective optimization study is most likely a constrained problem. Without proper constraints, obtained solutions will contain designs that are impossible or impractical to realize. Therefore, constraints are needed and they are formulated as methods described in Section 3.3.4 for this thesis study.

Recalling the constraint handling method, the first requirement is to express all constraints in an inequality form like:

$$g(x) \geq h$$

Then the inequality is transformed into the following form.

$$\frac{g(x)}{h} - 1 \geq 0$$

Moreover, constraint violations are computed by using the left size of the above inequality:

$$CV = \frac{g(x)}{h} - 1 \quad (4.24)$$

Constraints are investigated under three sections here. They are node-based, pipe-based and storage tank-based constraints. The constraints will be explained by making some modifications to the sample network with storage tank given in Section 4.1.4. These modifications are performed to illustrate violation of constraints in order to calculate sample constraint violation terms.

First modification to the network is to change diameter of pipe P1 from 200 mm to 110 mm. Results of EPS of this modification is given in Table 4.12 and Table 4.13.

Second modification to the original network is done by keeping diameter of pipe P1 as 200 mm but changing the diameter of cylindrical tank T from 10 m to 3 m. Results of EPS of this modification is given in Table 4.14.

It can be seen from the resulting tables that, there are intermediate time steps smaller than 1 hour. This is because there are intermediate solutions at those hours when storage tanks empty or fill.

Table 4.12. EPS Results of Pipe Parameters for First Modification

Time (hr)	Pipe Headloss Gradients (m/km)						Pipe Velocities (m/s)					
	P1	P2	P3	P4	P5	P6	P1	P2	P3	P4	P5	P6
00:00	48.87	2.04	9.03	9.03	2.04	6.02	2.25	0.49	0.85	0.85	0.49	0.68
01:00	46.02	2.00	9.44	9.44	2.00	9.04	2.18	0.49	0.87	0.87	0.49	0.85
02:00	42.74	1.93	9.68	9.68	1.93	12.24	2.09	0.48	0.88	0.88	0.48	1.00
03:00	42.99	1.64	6.23	6.23	1.64	0.91	2.10	0.44	0.70	0.70	0.44	0.25
04:00	45.52	1.38	3.15	3.15	1.38	6.17	2.16	0.40	0.48	0.48	0.40	0.69
05:00	48.97	1.35	2.30	2.30	1.35	14.45	2.25	0.39	0.41	0.41	0.39	1.10
06:00	53.59	1.36	1.72	1.72	1.36	24.86	2.36	0.39	0.35	0.35	0.39	1.47
07:00	61.29	1.13	0.10	0.10	1.13	81.06	2.54	0.36	0.07	0.07	0.36	2.78
07:49	202.22	7.26	24.68	24.68	7.26	0.00	4.84	0.98	1.46	1.46	0.98	0.00
08:00	330.77	11.87	40.38	40.38	11.87	0.00	6.31	1.27	1.91	1.91	1.27	0.00
09:00	253.76	9.11	30.98	30.98	9.11	0.00	5.47	1.10	1.66	1.66	1.10	0.00
10:00	235.98	8.47	28.81	28.81	8.47	0.00	5.26	1.06	1.59	1.59	1.06	0.00
11:00	218.80	7.85	26.71	26.71	7.85	0.00	5.05	1.02	1.53	1.53	1.02	0.00
12:00	235.98	8.47	28.81	28.81	8.47	0.00	5.26	1.06	1.59	1.59	1.06	0.00
13:00	218.80	7.85	26.71	26.71	7.85	0.00	5.05	1.02	1.53	1.53	1.02	0.00
14:00	253.76	9.11	30.98	30.98	9.11	0.00	5.47	1.10	1.66	1.66	1.10	0.00
15:00	253.76	9.11	30.98	30.98	9.11	0.00	5.47	1.10	1.66	1.66	1.10	0.00
16:00	291.09	10.45	35.53	35.53	10.45	0.00	5.89	1.19	1.78	1.78	1.19	0.00
17:00	330.77	11.87	40.38	40.38	11.87	0.00	6.31	1.27	1.91	1.91	1.27	0.00
18:00	417.06	14.97	50.91	50.91	14.97	0.00	7.16	1.44	2.16	2.16	1.44	0.00
19:00	330.77	11.87	40.38	40.38	11.87	0.00	6.31	1.27	1.91	1.91	1.27	0.00
20:00	218.80	7.85	26.71	26.71	7.85	0.00	5.05	1.02	1.53	1.53	1.02	0.00
21:00	156.10	5.60	19.06	19.06	5.60	0.00	4.21	0.85	1.27	1.27	0.85	0.00
22:00	70.29	2.52	8.58	8.58	2.52	0.00	2.74	0.55	0.83	0.83	0.55	0.00
23:00	57.35	2.36	10.23	10.23	2.36	5.87	2.45	0.53	0.91	0.91	0.53	0.67

Table 4.13. EPS Results of Node and Tank Parameters for First Modification

Time (hr)	Nodal Pressures (m)				Tank Levels (m)	Tank Water Volumes (m ³)
	J1	J2	J3	J4	T	T
00:00	55.11	54.91	54.01	54.91	1.00	78.54
01:00	55.40	55.20	54.26	55.20	1.25	97.86
02:00	55.73	55.53	54.56	55.53	1.55	121.92
03:00	55.70	55.54	54.91	55.54	1.91	150.27
04:00	55.45	55.31	55.00	55.31	2.00	157.24
05:00	55.10	54.97	54.74	54.97	1.75	137.67
06:00	54.64	54.51	54.33	54.51	1.36	106.67
07:00	53.87	53.76	53.75	53.76	0.83	65.11
07:49	39.78	39.05	36.58	39.05	0.00	0
08:00	26.92	25.74	21.70	25.74	0.00	0
09:00	34.62	33.71	30.62	33.71	0.00	0
10:00	36.40	35.55	32.67	35.55	0.00	0
11:00	38.12	37.33	34.66	37.33	0.00	0
12:00	36.40	35.55	32.67	35.55	0.00	0
13:00	38.12	37.33	34.66	37.33	0.00	0
14:00	34.62	33.71	30.62	33.71	0.00	0
15:00	34.62	33.71	30.62	33.71	0.00	0
16:00	30.89	29.85	26.29	29.85	0.00	0
17:00	26.92	25.74	21.70	25.74	0.00	0
18:00	18.29	16.80	11.71	16.80	0.00	0
19:00	26.92	25.74	21.70	25.74	0.00	0
20:00	38.12	37.33	34.66	37.33	0.00	0
21:00	44.39	43.83	41.92	43.83	0.00	0
22:00	52.97	52.72	51.86	52.72	0.00	0
23:00	54.26	54.03	53.01	54.03	0.00	0

Table 4.14. EPS Results of Tank Parameters for Second Modification

Time (hr)	Tank Levels (m)	Tank Water Volumes (m ³)	Time (hr)	Tank Levels (m)	Tank Water Volumes (m ³)
	T	T		T	T
00:00	1.00	7.07	13:00	1.36	9.60
00:20	5.00	35.34	14:00	3.01	21.25
01:00	5.00	35.34	15:00	0.58	4.09
02:00	5.00	35.34	16:00	2.18	15.44
03:00	5.00	35.34	16:59	0.00	0.00
04:00	5.00	35.34	17:00	0.01	0.09
05:00	5.00	35.34	17:14	0.00	0.00
06:00	5.00	35.34	18:00	0.00	0.00
07:00	3.89	27.50	19:00	0.00	0.00
08:00	1.59	11.23	20:00	0.00	0.00
08:39	0.00	0.00	21:00	3.68	25.98
09:00	0.00	0.00	22:00	3.70	26.12
10:00	2.45	17.30	22:18	5.00	35.34
11:00	1.66	11.73	23:00	5.00	35.34
12:00	2.83	19.98	00:00	5.00	35.34

4.4.1. Node-Based Constraints

A basic constraint that can be present in a water distribution network is a constraint to satisfy pressure requirements in the network. The satisfaction should consider all 24-hour results with intermediate time steps. The constraint should ensure a minimum predefined pressure (i.e. 30 meters) at every node at every time step. If there are violations in any node at any time step, a measure for this violation should be formulated.

If 30 meters is selected as minimum pressure, there are lower pressures than 30 meters for every node between 08:00 and 20:00 in Table 4.13. Constraint violation formulation is based on number of pressure values lower than 30 meters. This number

should be equal to zero. To express it in inequality form, it may be said that this number should be less than or equal to zero. If this number is indicated with “*non*”, number of junctions is “*nj*” and number of time steps is “*nts*” then;

$$non \leq 0$$

$$-non \geq 0$$

Then the constraint violation term is defined as:

$$CVPRES = -\frac{non}{nj \times nts} \quad (4.25)$$

As the constraint violation is structured on a table of values, normalization from number of nodes and number of solutions should be performed. Recall that every solution violating the constraint has equal weight regardless of its degree of violation. Other forms of constraint violation formulations giving more weight for bigger violations can also be produced.

To make an example from Table 4.13, value of *non* is 19, *nj* is 4 and *nts* is 25. As a result;

$$CVPRES = -\frac{19}{4 \times 25} = -0.19$$

4.4.2. Pipe-Based Constraints

Although not used in this thesis study, two possible constraint violations will be formulated by using pipe results of an EPS given in Table 4.12. These constraints are velocity and headloss gradient constraints.

A commonly used design consideration for pipe velocities is to keep them between 0.5 m/s and 2 m/s (“İçmesuyu Tesisleri Etüt, Fizibilite ve Projelerinin Hazırlanmasına Ait Teknik Şartname” 2013).

Similar to CVPRES formulation if number of solutions violating 0.5 m/s are “ nlv ” (number of low velocities), number of solutions violating 2 m/s are “ nhv ” (number of high velocities) and number of pipes are “ np ” then;

$$\begin{aligned}nlv &\leq 0 & nhv &\leq 0 \\-nlv &\geq 0 & -nhv &\geq 0\end{aligned}$$

Then the constraint violation terms are defined as:

$$CVLV = -\frac{nlv}{np \times nts} \quad (4.26)$$

$$CVHV = -\frac{nhv}{np \times nts} \quad (4.27)$$

A numerical example by using Table 4.12 can be given as:

$$CVLV = -\frac{41}{6 \times 24} = -0.2847$$

$$CVHV = -\frac{28}{6 \times 24} = -0.1944$$

Another pipe-based constraint can be implemented by using headloss gradients. Although there are not strict limitations for headloss gradients a maximum value of 100 m/km can be used. Again, number of solutions violating this constraint is counted and normalized just like constraints above. Let number of headloss gradients higher than 100 m/km is “ nhl ” then;

$$nhl \leq 0$$

$$-nhl \geq 0$$

Then the constraint violation term is defined as:

$$CVHLG = -\frac{nhl}{np \times nts} \quad (4.28)$$

A numerical example by using Table 4.12 can be given as:

$$CVHLG = -\frac{15}{6 \times 24} = -0.1042$$

4.4.3. Storage Tank-Based Constraints

A number of constraints based on storage tanks can be implemented. The main concern of these constraints is to help selection of such a design for tanks that;

- Tanks empty and fill completely in 24 hours
- Tank levels should not stay constant for too long (to maintain good water quality)
- Tank empty and fill cycles should not be too many
- Tank levels at the end of the day should be equal to or greater than levels at the beginning of the day
- Tanks should provide balancing volumes preferably to satisfy legislation requirements given in Section 2.3.

Above mentioned concerns are expressed as constraints below.

To ensure tanks fill and empty completely through the day, water volumes entering to the tanks and leaving from tanks should be taken into account. Sample values are given in Table 4.14. Consecutive differences are taken from these instantaneous water volume values to find total entering and leaving volumes. Both of these values should be equal to the actual volume of the tank. This equality can be turned into inequality of greater than or equal to. If total entering volume is represented with “*ev*”, total leaving volume is represented with “*lv*” and capacity of storage tank is represented with “*capacity*” then;

$$ev \geq capacity$$

$$\frac{ev}{capacity} - 1 \geq 0$$

Then the constraint violation term is defined as:

$$CVEV = \frac{ev}{capacity} - 1 \quad (4.29)$$

A numerical example by using Table 4.14 can be given as:

$$CVEV = \frac{112.25}{35.34} - 1 = 2.1763$$

As the constraint violation value is positive it is not violating the defined constraint. Therefore, it is not considered as a constraint violation value. However, if similar results from Table 4.13 are used, the corresponding constraint violation value would be:

$$CVEV = \frac{97.76}{392.70} - 1 = -0.7510$$

As the result is a negative value, it is considered as a constraint violation and included in calculations where necessary.

In a similar way leaving volumes can be expressed as constraint violations:

$$lv \geq capacity$$

$$\frac{lv}{capacity} - 1 \geq 0$$

Then the constraint violation term is defined as:

$$CVLV = \frac{lv}{capacity} - 1 \quad (4.30)$$

A numerical example by using Table 4.13 can be given as:

$$CVLV = \frac{157.24}{392.70} - 1 = -0.5996$$

To avoid too long constant water levels in tanks, a constraint can be implemented. This constraint will help to avoid too long full and empty tank water levels. Volume changes considered above for CVEV and CVLV are again used for this constraint. A limit for maximum constant water level duration is predefined. 3 hours is an average value for this purpose. Longer or shorter durations are also possible. Checking the water volume changes, consecutive three-hour constant water levels are counted. Let this number be “con”.

$$con \leq 0$$

$$-con \geq 0$$

Then the constraint violation term is defined as:

$$CVCON = -\frac{con}{nts} \quad (4.31)$$

A numerical example by using Table 4.13 can be given as:

$$CVCON = -\frac{15}{24} = -0.625$$

Having too many empty and fill cycles in a storage tank is an indication of too small tank design. To avoid it a constraint can be implemented depending on the water volume changes. Water volume changes are examined and a count is performed on fill to empty and empty to fill changes. Let this count be represented by “osc”

(oscillation). And mostly depending on the peaks and crests of the water demand curve (Figure 4.3) a maximum number for fill-empty cycles (“cyc”) is predefined (i.e. 5). “osc” value is compared against this maximum value and a constraint is formed as;

$$osc \leq cyc$$

$$\frac{cyc}{osc} - 1 \geq 0$$

Then the constraint violation term is defined as:

$$CVOSC = \frac{cyc}{osc} - 1 \quad (4.32)$$

A numerical example by using Table 4.14 can be given as (assuming cyc=5):

$$CVOSC = \frac{5}{7} - 1 = -0.2857$$

The value 7 is found by counting the crests of volume change.

An important constraint for storage tanks is to have water levels at the end of the day greater than or equal to water levels at the beginning of the day. Satisfying this constraint will ensure storage tanks will not deplete but preserve enough water to serve the next day. Let “ls” represent level of a tank at the beginning of the day, and “le” represent the water level at the end of the day. Then;

$$ls \leq le$$

$$\frac{le}{ls} - 1 \geq 0$$

Then the constraint violation term is defined as:

$$CVLSE = \frac{le}{ls} - 1 \quad (4.33)$$

A numerical example by using Table 4.13 can be given as:

$$CVLSE = \frac{0}{1} - 1 = -1$$

Lastly, to control total volume of storage tanks, another constraint can be implemented. However, this constraint may be conflicting with other constraints defined above so it may not help to achieve feasible elements when used with above constraints.

This final constraint is based on the difference of actual tank volumes (“*av*”) and figures predefined (“*pdv*”) (i.e. from legislations). As most legislation defines a minimum volume for storage tanks, it is possible to write:

$$av \geq pdv$$

$$\frac{av}{pdv} - 1 \geq 0$$

Then the constraint violation term for volume differences is defined as:

$$CVVD = \frac{av}{pdv} - 1 \quad (4.34)$$

Assuming total actual volume should be greater than or equal to %33 of average daily consumption (“*pdv*”) (see Section 2.3) a numerical example by using Table 4.4 and Table 4.14 can be given as:

$$CVVD = \frac{35.34}{1200} - 1 = -0.9706$$

After defining and computing different constraint violations, it is now possible to sum them up to get a single value under the name overall constraint violation (OCV). As all constraint violation values are normalized, they are in the same order.

To sum up this chapter, objectives and constraints can be written as follows:

$$\min (\text{Cost}) = \min \left(\sum_{i=1}^{np} L_i \times u_{pipe}(D_i) + \sum_{i=1}^{nt} u_{tank}(V_i) \right)$$

$$\max(I_n) = \max \left(\frac{\sum_{i=1}^{nj} C_i q_i^* (h_i - h_i^*) + \sum_{t=1}^{nt} C_t q_t^* (h_t - h_t^*)}{\sum_{k=1}^{nr} Q_k H_k + \sum_{j=1}^{npu} P_j + \sum_{t=1}^{nt} Q_t H_t - \sum_{i=1}^{nj} q_i^* h_i^* - \sum_{t=1}^{nt} q_t^* h_t^*} \right)$$

Subject to:

- $ev \geq capacity$
- $lv \geq capacity$
- $con \leq 0$
- $osc \leq cyc$
- $ls \leq le$

CHAPTER 5

METHODOLOGY

5.1. Coding

Developed methodology is coded on MATLAB 2019b update 3 (9.7.0.1261785) environment, using open source EPANET codes. EPANET is developed on C language. These open source C codes are compiled by Visual Studio 2019 (16.4.3). “Desktop development with C++” workload should be present in Visual Studio configuration for compilation. As a result, dll files are produced. With these dll and header files EPANET hydraulic solver and its functions became available under MATLAB environment.

EPANET was originally developed by USEPA until version 2.0.12. (Rossman 2000). Since 2008, EPANET is continuously being modernized and improved by “Open Water Analytics” team (“The Water Distribution System Hydraulic and Water Quality Analysis Toolkit” 2019). In 2019 developed version reached 2.2. Compiling and coding process is schematized in Figure 5.1.

The developed code consists of a main code with 21 additional functions. Main code consists of three major parts: initialization, generation of first random population and main optimization process. The process starts with a random population and as iterations go by one or more feasible members are found that satisfy all constraints. Then these feasible members are increased in number by the NSGA-II procedure and feasible search for the optimization is performed until the generation limit is reached.

The computer used in this study has an Intel Core i7 4702MQ 2.20 GHz CPU, 32 GB DDR3 L memory, NVIDIA GeForce GTX 850M graphics card and 240 GB Corsair Force LS SSD with Windows 10 64-bit operating system.

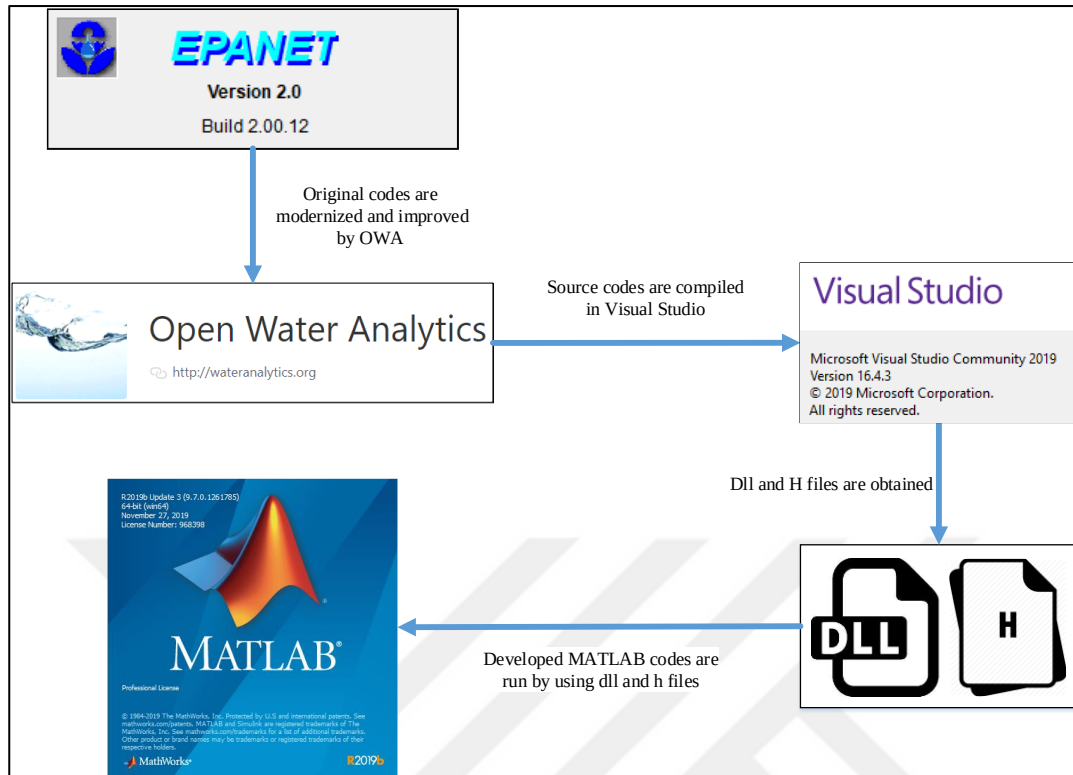


Figure 5.1. Compiling and Coding Process

5.2. Variables

Objective functions given in formulas 4.1 and 4.2 with constraints given in formulas from 4.25 to 4.34 will be explained in the following sections.

Complete process starting from initial population generation to the end of iterations goes over a variable named “pop”. Contents of “pop” is summarized in Table 5.1. The structure of the chromosomes being recorded in variable pop consists of two main parts: pipe parameters and tank parameters (Figure 5.2). A sample of the variable pop is given below (Figure 5.4 and Figure 5.5). A simplified flowchart of the whole process is provided in Figure 5.3.

Position field in main variable pop includes integer pipe diameters, tank counts, tank sites, tank types, tank height, diameter and elevation values. It should be noted that size of position field depends on tank counts. Although the methodology is capable

of dealing any number of tanks, it is usually used as a fixed variable decided at the beginning of the procedure. This makes every population member's position field to be of same length. Another thing to note is, tank types are again decided at the beginning. In this thesis study elevated tank types are used however the methodology is again capable of dealing population members with different tank types.

Depending on the decision variables set in position field, costs, constraints and NSGA-II parameters are calculated within the procedure and stored in necessary fields in the variable pop. Costs are calculated separately for pipes and tanks. They are stored accordingly in TPi Cost field (Table 5.1).

It can also be noted that with two binary encoding-decoding functions, pipe diameters are converted from integer to binary and back whenever a binary operation takes place.

	Pipe Parameters	Tank Parameters					
	Integer Diameters	Number	Site	Type	Height	Diameter	Elevation
Bit length	<i>np</i>	1	<i>nt</i>	<i>nt</i>	<i>nt</i>	<i>nt</i>	<i>nt</i>
Note: np=Number of pipes, nt=Number of tanks							

Figure 5.2. Two Main Parts of MATLAB Variable “pop”

Table 5.1. Summary of variable “pop”

Field Name	Content	Size
Position	Integer and binary pipe diameters, all tank parameters	$np+5nt+1$
Cost	I_n and Cost values	2
TPi Cost	Tank and Pipe Costs	4
Rank	Front number of the member	1
Domination Set	Set of elements the member dominates	Variable
Dominated Count	Number of elements dominating the member	1
Crowding Distance	Normalized crowding distance of the member	1
Generation	Number of current generation (for archiving purposes)	1
CVLV	Overall constraint violations of outflow volumes from tanks	nt
CVEV	Overall constraint violations of inflow volumes to tanks	nt
CVLSE	Overall constraint violations tank end-of-simulation water levels	nt
CVVD	Overall constraint violations of tanks' required minimum volumes	1
CVCON	Overall constraint violations of tank level constancy	nt
CVOSC	Overall constraint violations of tank level oscillations	nt
CVPRES	Overall constraint violations of negative pressure junctions	1
OCV	Total overall constraint violation value	1
Capacity	Tank total volumes (m^3)	nt
FI	Binary code indicating whether the member is feasible or not	1

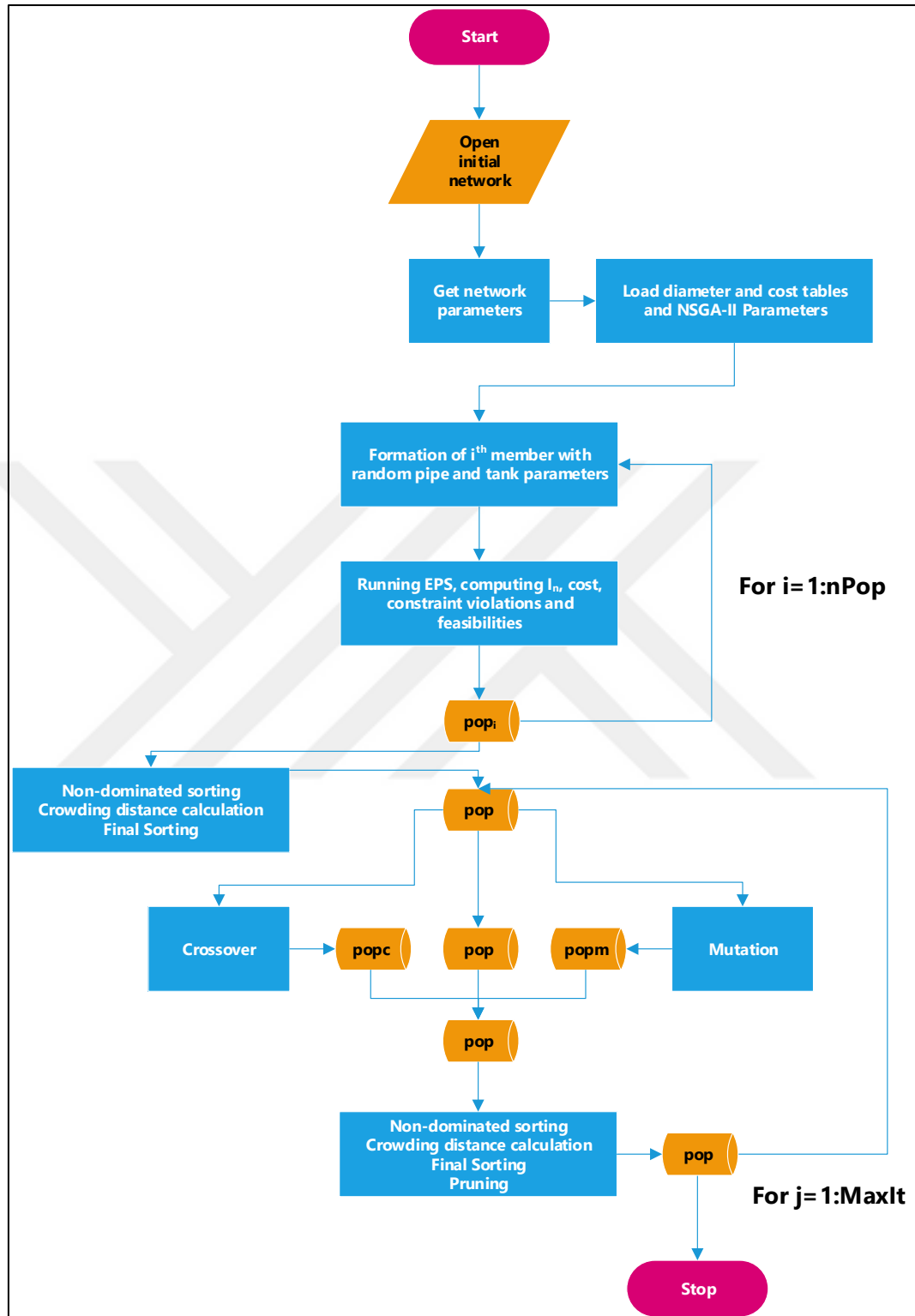


Figure 5.3. Flowchart for the Complete Methodology

Pipe Parameters				
	Integer Diameters			
Solution No	Diameter1	Diameter2	Diameter3	Other Diameters
1	600	350	400	...
2	300	100	450	...
3	350	100	750	...
4	350	600	500	...
<i>i</i>	...	...	...	...
	<i>Bit length = np (number of pipes)</i>			

Figure 5.4. Pipe Parameter Part of the MATLAB Variable “pop”

Solution No.	nt_i	Tank Parameters / Place - Type - Height - Diameter - Elevation																			
1	3	45	21	52	1	3	2	8.84	8.84	29.10	13.18	12.94	16.90	0	15.58	0					
2	5	27	52	55	16	7	3	1	1	3	2	5.69	13.01	13.06	7.82	14.90	5.10	38.05	7.91	27.72	10.31
3	2	1	21	2	3	23.73	4.20	28.37	5.46	0	14.00										
4	4	20	48	21	53	1	3	3	1	13.79	11.16	7.24	13.52	59.29	30.44	12.13	38.05	0	33.38	15.21	0
i	...	Bit Length Depending on the Number of Tanks ($=5 \times nt_i$)																			
Colour Legend																					
	Tank Places (Junction Ids)		Tank Types 1-Reservoir 2-Standpipe 3-Elevated		Tank Heights (m)		Tank Diameters (m)		Tank Elevations (m)												

Figure 5.5. Tank Parameter Part of the MATLAB Variable “pop”

5.3. Workflow of the Code

Developed code runs three major parts. First part initializes the variables, second and third parts consist of iterations. Second part iterations generate a predefined number of initial population members. Finally, third part runs iterative NSGA-II procedure until the stopping criteria are met.

5.3.1. Initialization

Initialization part begins with loading an initial EPANET *.inp* file, which should have a network layout to be optimized. The layout should at least contain one source reservoir, one consumption junction and one pipe. After opening the initial network, initial parameters (number of nodes, number of junctions, number of pipes, length of pipes, nodal elevations, etc.) are read and stored as MATLAB variables. Afterwards, data tables defining pipe diameters, pipe unit costs and tank sizes are loaded as MATLAB variables. Then, NSGA-II parameters such as number of generations and population size are defined. These defined data tables and variables are kept constant throughout the optimization process and called by various functions whenever needed.

5.3.2. Initial Population Formation

Generation of first random population includes random selection of pipe diameters and tank parameters. Tank parameters consist of number of tanks in the network, places of tanks in the network (as junction ids that the tanks are connected to), tank heights (H), tank diameters (D) and tank elevations (E). Water tanks are assumed as cylindrical tanks having three different types: reservoirs, standpipes and elevated tanks. Assuming cylindrical shapes makes coding of tank dimensions easier in EPANET as it accepts directly height and diameter values as tank dimensions. H, D and E values for reservoir, standpipe and elevated tank are shown in Figure 4.8. A detailed flowchart for this stage is given in Figure 5.6.

To start with good initial members a method named as seeding is used (Haupt and Haupt 2004; Sivanandam and Deepa 2008). This method helps the algorithm to reach optimum results quicker.

While forming initial population pipe diameters, minimum and maximum pipe diameter selections for every pipe is determined rather than a complete random selection. Then random pipe diameter selections are made within these minimum and maximum values. The application is performed by assigning biggest diameters for once to all pipes and an EPS run is computed. Then discharges for each pipe at minimum and maximum peak hours are recorded. Depending on these flows, minimum pipe areas are defined by using 2 m/s flow velocity, and maximum pipe areas are defined by using 0.5 m/s flow velocity. Then these minimum and maximum areas are used to determine minimum and maximum pipe diameter limits for every pipe.

Calculations on any pipe are done by using the formulas below (A: pipe area, m², F flows, m³/s):

$$F_{max} = 0.5 \times A_{max} \quad (5.1)$$

$$F_{min} = 2.0 \times A_{min} \quad (5.2)$$

$$A_{max} = \frac{\pi \times D_{max}^2}{4} \quad (5.3)$$

$$A_{min} = \frac{\pi \times D_{min}^2}{4} \quad (5.4)$$

Random pipe diameter selections are then made for each pipe's separate minimum and maximum diameter intervals. It should be noted that the methodology is developed to choose from a list of 16 distinct integer diameters. If there are less choices in the problem, some diameters can be duplicated to fill the list of 16.

A numerical example is given in Table 5.2 which is derived from the sample network without tank given in Section 4.1.4. Diameters calculated at the end are rounded to bottom for minimum diameters and rounded to top for maximum diameters to help increase diversity in pipe diameter selection. Pipe diameter list given in Table 4.8 are used.

Table 5.2. Sample Minimum and Maximum Pipe Diameter Limits

Time	P1	P2	P3	P4	P5
Flows (l/s)					
03:00	12.00	4.50	3.00	3.00	4.50
19:00	68.00	25.50	17.00	17.00	25.50
Pipe Areas (m ²)					
03:00	0.0060	0.0023	0.0015	0.0015	0.0023
19:00	0.1360	0.0510	0.0340	0.0340	0.0510
Calculated Pipe Diameters (mm)					
03:00	87.40	53.52	43.70	43.70	53.52
19:00	416.13	254.82	208.06	208.06	254.82
Selected Pipe Diameters (mm)					
Minimum	75	50	25	25	50
Maximum	450	300	250	250	300

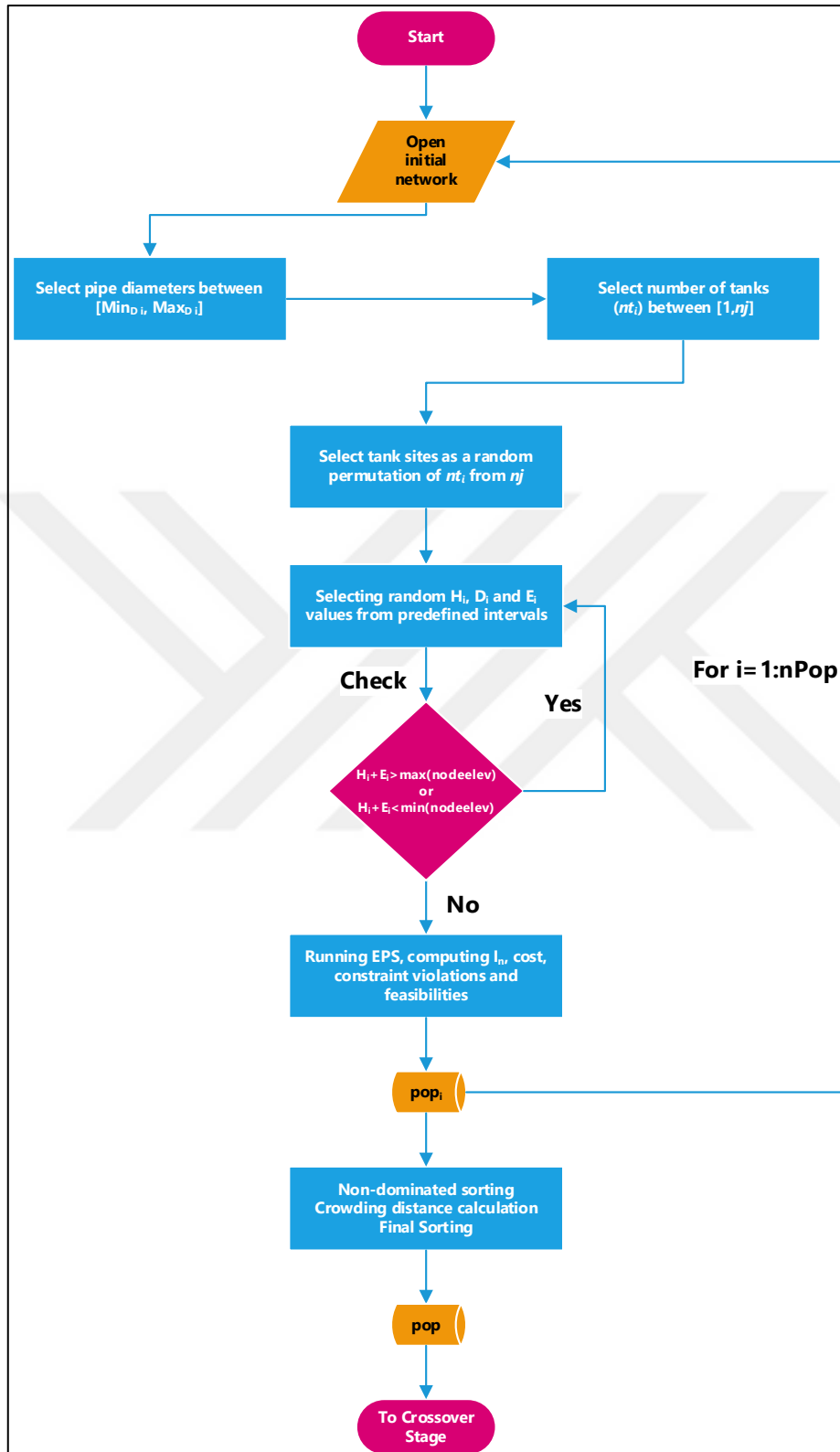


Figure 5.6. Detailed Flowchart for Initial Population Formation

After selecting random diameters for all of the pipes, random values for number of tanks, tank sites and tank H, D, E values are determined.

Number of tanks for a member is determined by a random integer between one and number of consumer junctions that are not connected to another tank. For a network having 10 junctions, can have a random tank number any integer between 1 and 10. All other tank parameters depend on this value as it determines the size of the chromosome by changing numbers of tanks.

When number of tanks for a member is defined, that number of junctions are randomly selected again from consumption junctions that are not affiliated with another tank. It is formulated by taking one random permutation of number of tanks from number of junctions.

Placement of tanks to a selected junction is made by a short pipe of 1 m length and a big diameter of Ø1000 mm to keep the headloss in this pipe negligible. Positioning of the tank is made to a coordinate 10-meter east and 10-meter north of the junction it is being connected. This is done for visual purposes; any other placement is also possible.

Tank H, D and E values are chosen between minimum and maximum values given in Table 4.10.

A do-while loop present in the code ensures selected H and E values not being higher than top reservoir elevations and not being lower than the lowest junction elevations. This makes the iterations reach feasible solutions faster.

After defining all the pipe and tank parameters as MATLAB variables, they are transferred into the initially opened EPANET file. As the EPANET file is realized in the memory, this time EPS runs are performed providing a base data for several objective function and constraint calculations. As soon as these calculations are completed, the initial population formation process ends with marking the current member as feasible or infeasible by checking its constraint violations. Formation of

initial population is performed one-by-one in a for-loop of length defined on previous stage.

Last phase of the procedure starts with recently formed initial population and makes several crossover and mutation calculations on randomly selected elements of the initial population and forms child and mutation populations respectively. These two populations are merged with the initial population, sorted by NSGA-II methods and reduced (pruned) into original population size according to the rules defined in Chapter 3. Afterwards the resulting population is treated as an initial population and crossover and mutation operations are repeated until the end of stopping criteria.

5.3.3. Iterative NSGA-II Procedure

This part selects and produces child population members within the NSGA-II framework. To perform selection and child member formation a core mechanism composed of three functions are used.

5.3.3.1. Core Mechanism

Figure 5.7 gives a detailed flowchart for constraint-domination processed included in crossover and mutation phases. Three major functions that are formulated according to the rules given above are used. These functions are F1-Feasibility, F2-Constraint Domination and F3-Selection.

F1-Feasibility Function:

This function simply checks every member in terms of feasibility and gives the binary code 1 or 0 according to their constraint violation values. If a member's all constraint violation values are greater than or equal to zero then the member is marked as feasible.

This means, violating any constraint makes a member infeasible. To be a feasible member, it should satisfy all the constraint rules defined.

Using the feasibility values given by this function, following functions F2 and F3 checks members according to a three-step check defined above.

F2-Constraint Domination Function:

This function checks if solution i constrained-dominates solution j . It works according to Deb's rules as follows:

- If i and j are both feasible and i dominates j in the usual manner, i also constrained-dominates j .
- If i is feasible and j is infeasible, i constrained-dominates j .
- If i and j are both infeasible and if;
 - $all \left(sum \left(abs \left(i.CVLSE \right) \right) < sum \left(abs \left(j.CVLSE \right) \right) \right) and$
 - $all \left(sum \left(abs \left(i.CVEV \right) \right) < sum \left(abs \left(j.CVEV \right) \right) \right) and$
 - $all \left(sum \left(abs \left(i.CVLV \right) \right) < sum \left(abs \left(j.CVLV \right) \right) \right) and$
 - $all \left(sum \left(abs \left(i.CVOSC \right) \right) < sum \left(abs \left(j.CVOSC \right) \right) \right) and$
 - $all \left(sum \left(abs \left(i.CVCON \right) \right) < sum \left(abs \left(j.CVCON \right) \right) \right)$

are true, then i constraint-dominates j .

Here summation of absolute value of violations are taken into consideration and compared among different members to deal with both negative and positive constraint violations.

F3-Selection Function:

Selection is done by modifying the usual binary tournament selection for the needs of constraint violations. Again, a similar three-step check is made along two randomly selected members considering feasibilities. The winner of this three-step check is also the winner of the tournament selection and becomes one of the parents for crossover phase. The three-step check is performed as follows:

- If i and j are both feasible the one having lower rank (front) is selected. If they have equal ranks, then the one having higher crowding distance wins the tournament and is selected.
- If either i or j is feasible and the other one is infeasible, the feasible one wins the tournament and is selected.
- If i and j are both infeasible, then the one having lower overall constraint violation is selected. The constraint violations are calculated as:

$$\begin{aligned} i.OCV &= \text{sum} (abs (i.OCVLSE)) & j.OCV &= \text{sum} (abs (j.OCVLSE)) \\ &+ \text{sum} (abs(i.CVEV)) & &+ \text{sum} (abs(j.CVEV)) \\ &+ \text{sum} (abs(i.OCVLV)) & &+ \text{sum} (abs(j.OCVLV)) \\ &+ \text{sum} (abs(i.CVOSC)) & &+ \text{sum} (abs(j.CVOSC)) \\ &+ \text{sum} (abs(i.CVCON)) & &+ \text{sum} (abs(j.CVCON)) \end{aligned}$$

This ensures selecting lower constraint violating members and forces a search in the infeasible space towards the feasible region.

These three functions provide the main search mechanism for this study. Any additions or subtractions to constraints (i.e. adding another constraint or removing one of the constraints), should be reflected to above three functions to ensure a good search in the objective function space.

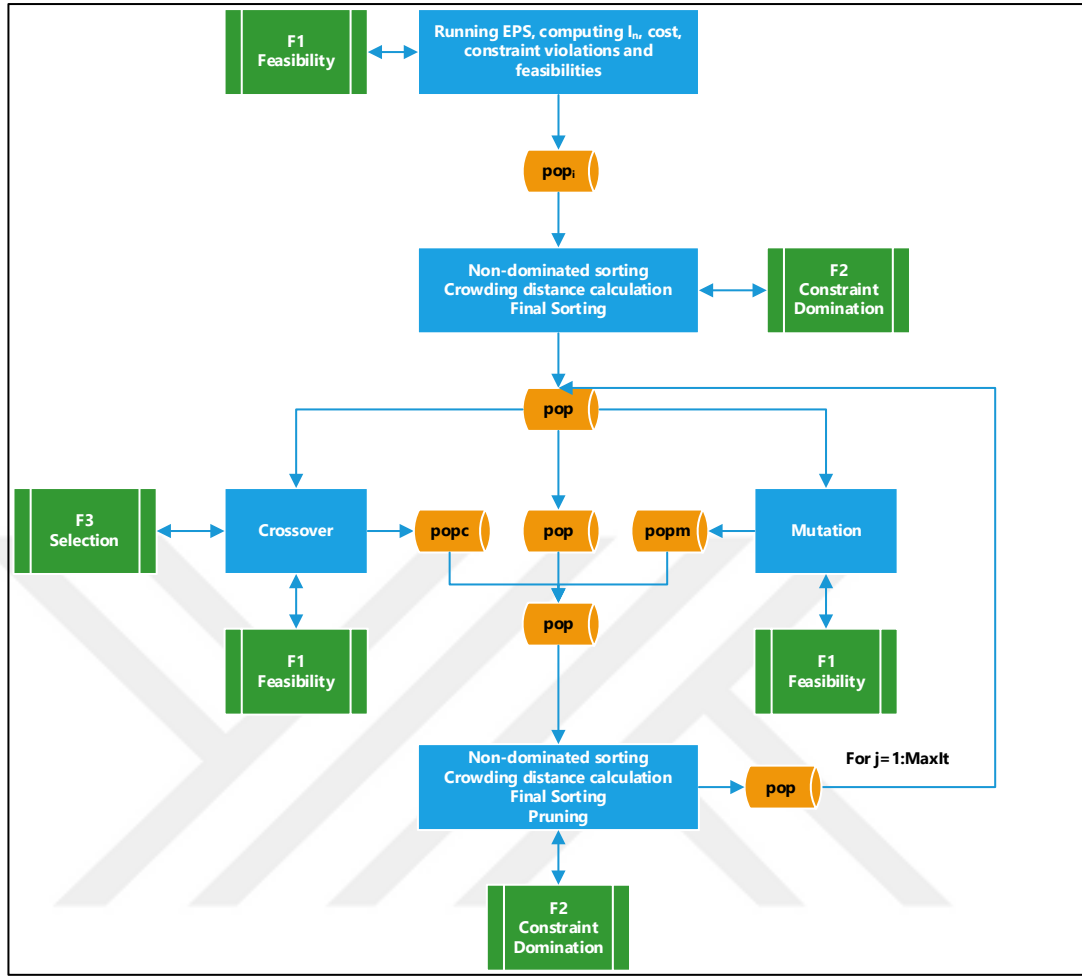


Figure 5.7. Detailed Flowchart of Constraint Handling Process

Considering initialization, crossover, mutation and constraint handling processes, the optimization runs are observed to show a two-phase search behavior. In the first phase an infeasible search based on constraint violations are done. This phase can be named as “Lowest Overall Constraint Violation Sorting”. It is the phase of promoting members having lower overall constraint violations to the next generations. As soon as a feasible solution is found, usual NSGA-II rules are followed and a partial feasible-infeasible search is done. This stage is quickly passed within some generations leaving its place to the final phase of feasible search. Explained phases are shown in Figure 5.8.

In the flowchart, three iteration counts are given. These iteration counts add up to the maximum generation count. Usually, it_0 is observed to be less than 50 iterations. When it exceeds 100 iterations, the process is advised to be restarted with different random seeds.

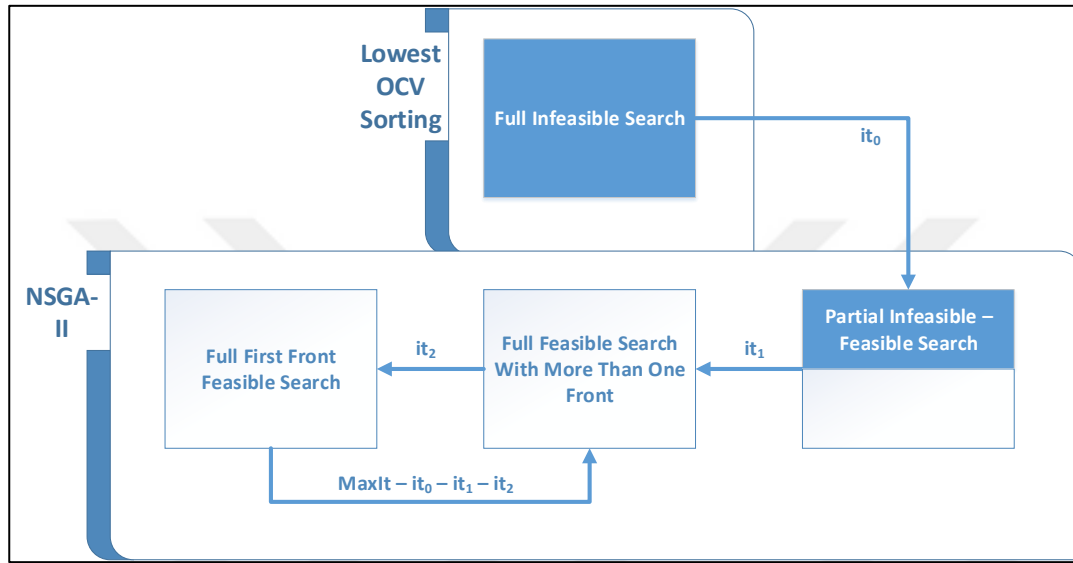


Figure 5.8. Flowchart Describing Infeasible and Feasible Search Phases

5.3.3.2. Crossover

Crossover phase can be detailed as follows. It takes place in two main parts. First part is single point binary crossover of pipe diameters, second part is simulated binary crossover of tank numbers. Binary encoder and decoder functions take place in binary crossover.

Two random members from the initial population are chosen for binary crossover. A random cutting point is defined within the two parent members, and according to that cutting point, binary codes of the members' diameters are cut and exchanged with each other (Figure 5.9).

0	1	1	1	0	1	0	0
0	1	0	0	1	1	0	1
0	1	1	0	1	1	0	1
0	1	0	1	0	1	0	0

Figure 5.9. Sample One Point Binary Crossover for Pipe Diameters

Simulated binary crossover is applied for tank count crossovers. This formulation takes the number of tanks from the two previously selected members, and generates two new tank counts according to the formulation below. Sum of entering tank counts does not change after the crossover operation. For example, two members having five and eight tanks may result in six and seven tanks respectively. Resulting number of tanks are in between the entering numbers.

$\alpha = \text{random value between } (0,1)$ $y_1, y_2: \text{parent tank counts}$
 $nt_1, nt_2: \text{child tank counts}$

$$y_1 = \alpha \times nt_1 + (1 - \alpha) \times nt_2 \quad (5.5)$$

$$y_2 = \alpha \times nt_2 + (1 - \alpha) \times nt_1 \quad (5.6)$$

According to the new tank counts generated, rest of the tank property chromosome needs to be resized. To fit into new chromosome sizes, necessary random cuts are made (same cutting schema is applied to all remaining tank parameters) (Figure 5.10).

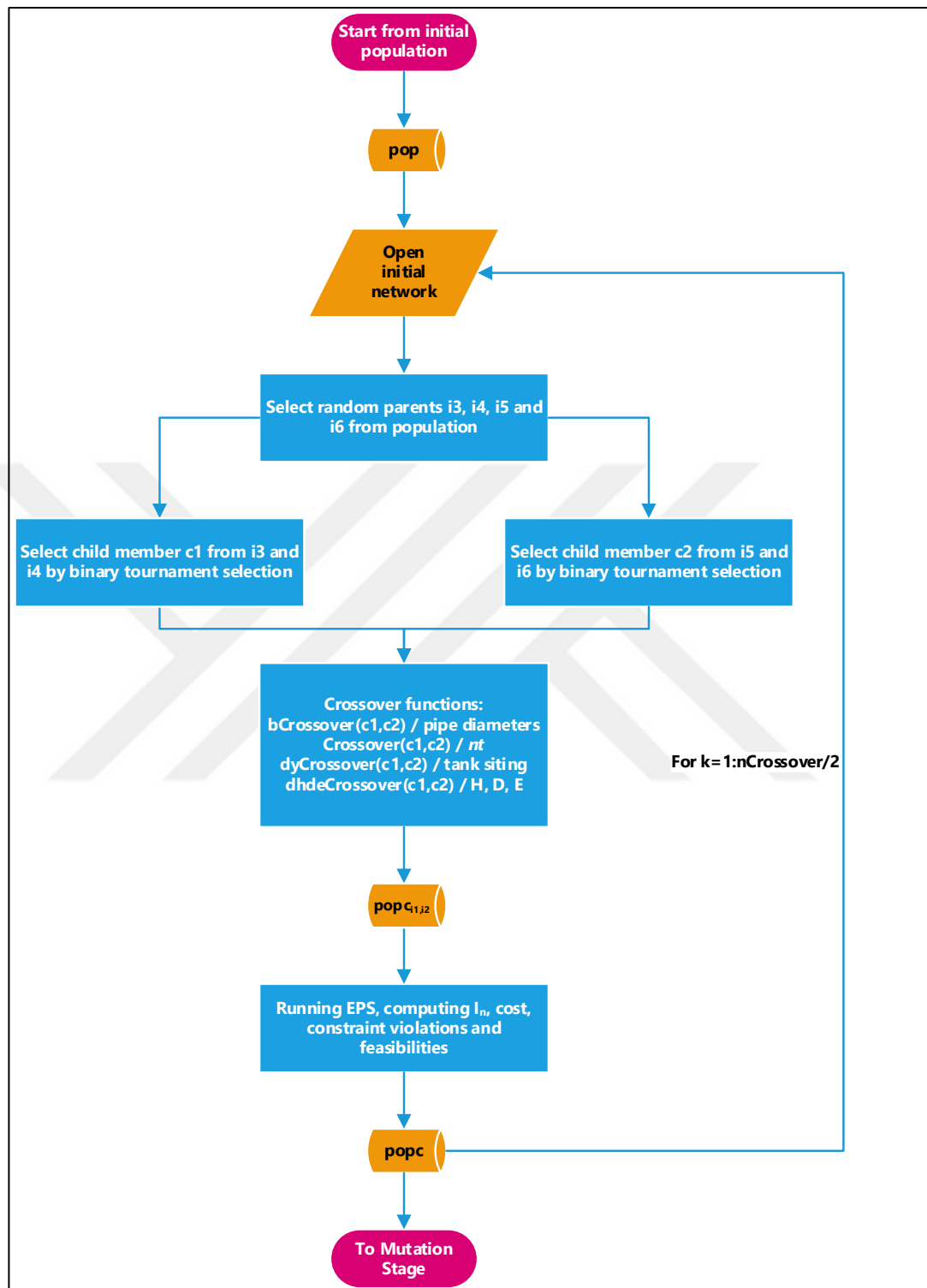


Figure 5.11. Detailed Flowchart for Crossover

5.3.3.3. Mutation

Similarly, mutation phase consists of two main parts, binary and simulated binary parts. In binary crossover, at a predefined low rate, one bit of randomly selected member's chromosome is mutated from either 0 to 1 or vice versa (Figure 5.12). Again, binary encoder and decoder functions take place in binary mutation.

0	1	1	1	0	1	0	0
0	1	1	0	0	1	0	0

Figure 5.12. Sample One Point Mutation for Pipe Diameters

Simulated binary mutation takes place in tank numbers again at a low rate. The selected member's tank number is slightly increased or decreased randomly in this phase. According to the new tank number, either some deletions are made in the chromosome, or some new random property tanks are added to the chromosome.

Mutation in tank places takes place of changing a random tank's siting to another available junction randomly. All the original tank parameters are preserved and moved to the new site, while the tank at the old site is removed.

Mutation of H, D and E parameters takes place by applying an addition or subtraction of a random noise on the original parameters according to the formulation below:

$$\begin{aligned}
 H_{new} &= H_{old} \pm \sigma_H * rand \\
 D_{new} &= D_{old} \pm \sigma_D * rand \\
 E_{new} &= E_{old} \pm \sigma_E * rand
 \end{aligned}
 \tag{5.7}$$

σ_H , σ_D and σ_E are parameters the maximum amount of random noise considering available minimum and maximum tank sizes. They are defined by using values from Table 4.10 as:

$$\begin{aligned}\sigma_H &= (H_{max} - H_{min})/\rho_H \\ \sigma_D &= (D_{max} - D_{min})/\rho_D \\ \sigma_E &= (E_{max} - E_{min})/\rho_E\end{aligned}\tag{5.8}$$

ρ_H , ρ_D and ρ_E are defined such that, a maximum of 0.02 meters of mutation for every dimension can be obtained.

Finally, *rand* is a random number between 0 and 1. With this formulation, a small addition or subtraction (a random noise) to the desired parameter is made acting as a mutation operator.

As a result, some of the six parameters undergo mutation with an equal probability at every generation.

Detailed flowchart for mutation stage is given in Figure 5.13.

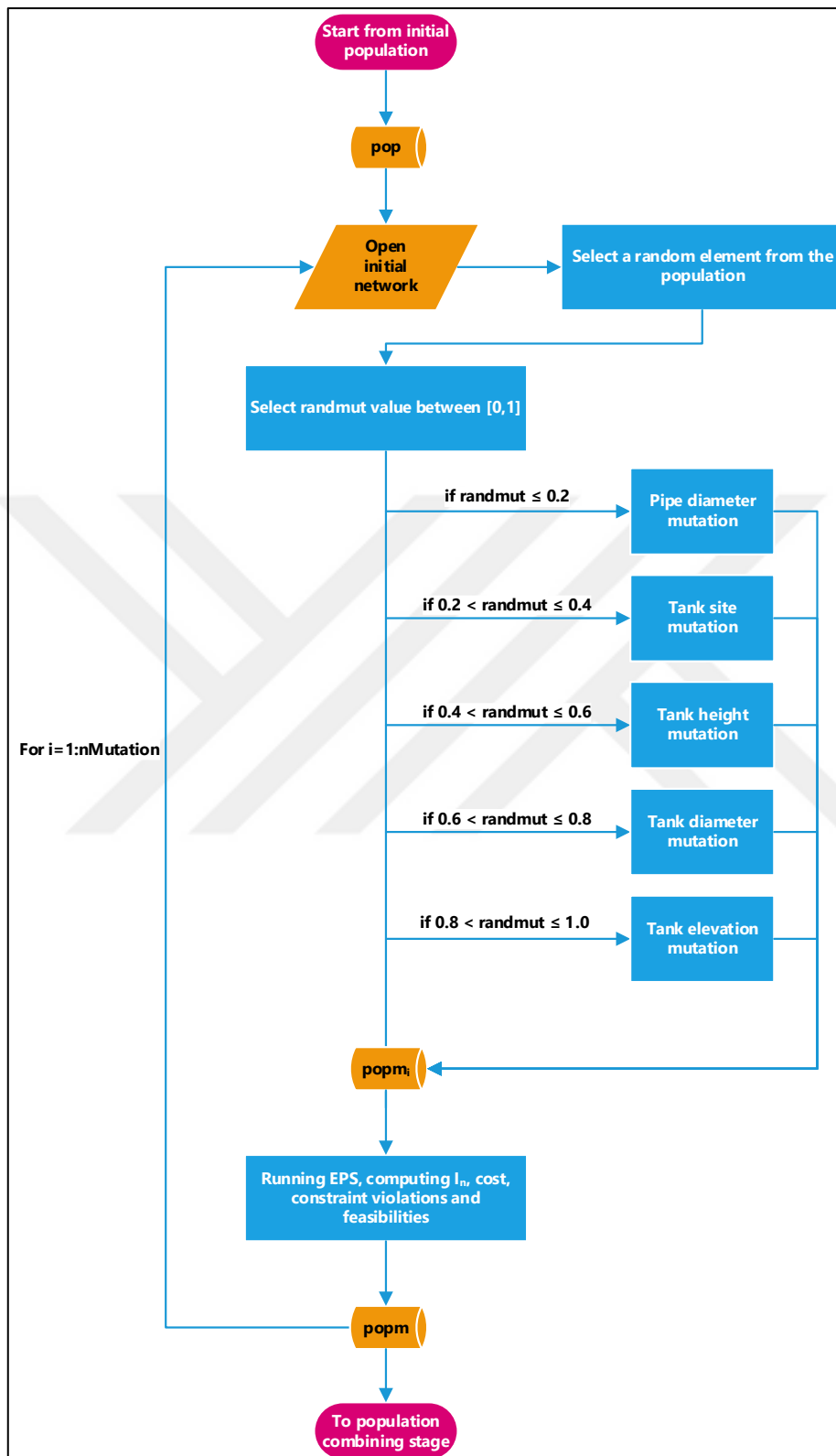


Figure 5.13. Detailed Flowchart for Mutation

5.4. Constraint Handling

Constraints can be summarized as:

- Minimum pressures at junctions (CVPRES)
- Maintaining an increase or same tank levels as the beginning level at the end of simulation (CVLSE)
- Filling and emptying at least half volumes of the tank throughout the simulation (CVEV and CVLV)
- Not having too many (more than 5) tank filling and emptying cycles in 24 hours (CVOSC)
- Not having too long (more than 3 hours) constant water levels in 24 hours (CVCON)

Above constraints are satisfied according to the constraint handling definitions given in Section 4.4. They are coded as MATLAB variable names given in parenthesis above.

The constraints are formulated below. In each constraint violation, values for tanks are computed separately. For example, in a network having two water tanks, two CVLSE values are computed and stored separately under this variable. And if constraints satisfy the predefined inequality, they are considered as zero violation and their computed values are not considered in OCV.

CVLSE

This constraint is implemented to have final water levels in every tank higher than or equal to the starting water levels. The aim is to have sustainable water levels at the end of the day, which will provide the starting level for the next day. Therefore, an overall decrease in levels are avoided with this constraint.

The formulation depends on two variables, le and ls indicating water level in a tank at 24th and 0th hour respectively.

$$le \geq ls$$

$$CVLSE = \frac{le}{ls} - 1 \geq 0$$

(5.9)

CVEV

This constraint along with CVEV, forces the tanks to completely empty and fill at least one times during 24 hours. It is essential both to use the provided active volume of the tank and to supply fresh water to the consumers. If too small volumes enter to and leave out the tanks, big volumes of water will get aged and quality problems will arise.

The formulation depends on two variables, *ev* and *capacity* indicating total volume entering to the tank in 24 hours and capacity of the tank respectively.

$$ev \geq 0.5 \times capacity$$

$$CVEV = \frac{ev}{0.5 \times capacity} - 1 \geq 0$$

(5.10)

CVLV

Similar formulation as CVEV is valid for CVLV to ensure at least one capacity of volume leaves throughout the day. This time variables are *lv* and *capacity*.

$$lv \geq 0.5 \times capacity$$

$$CVLV = \frac{lv}{capacity} - 1 \geq 0$$

(5.11)

CVOSC

This constraint is implemented to avoid too many tank level fill-empty change occurrences. It highly depends on the shape of the daily water use pattern. In this study, the water consumption pattern defined by Walski (Walski et al. 1987) (Figure 5.14). This pattern has a decrease two hours after start, then a big increase until 11th hour then a gradual decrease until end of the day. It has three changes in the fill-empty actions. An upper limit of five is defined for tanks according to this pattern (cyc value) and the constraint violation is formulated below. The only variable used is *osc*, that is calculated by counting number of top and bottom peaks in tank volume.

$$osc \leq 5$$

$$CVOSC = 1 - \frac{osc}{5} \geq 0$$

(5.12)

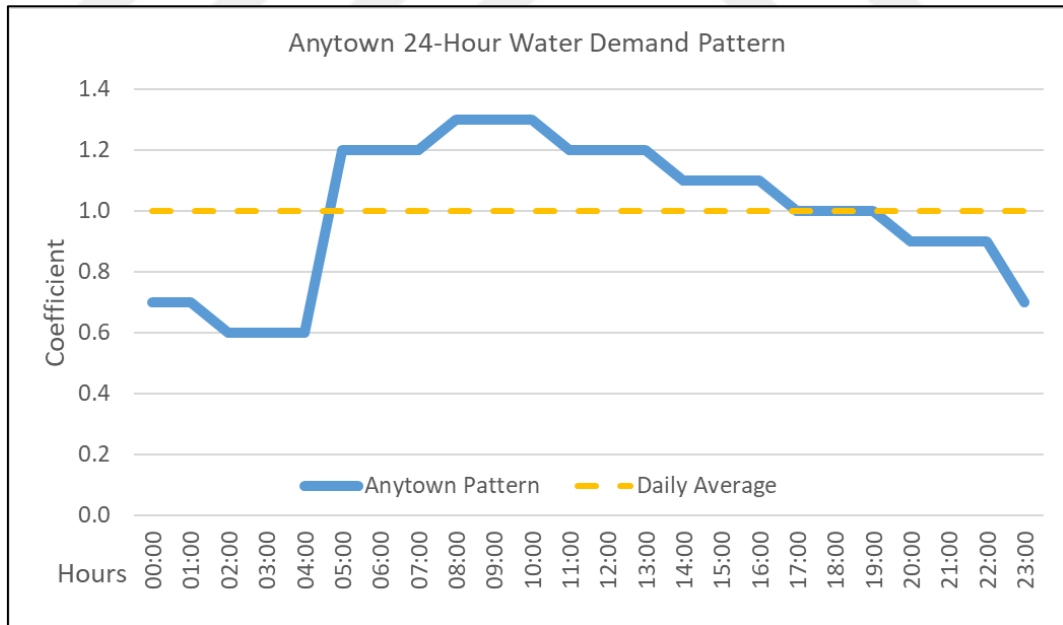


Figure 5.14. Anytown Demand Pattern for EPS Studies

CVCON

This constraint is implemented to avoid too long consecutive still water levels in the tanks. Again, in consistence with water use pattern, a three hour still water level is allowed with a constraint violation value of zero and longer unchanged water levels are computed by the below formulation. The variable used here is named as *con*. The calculation is done in an if clause for only values of $con \geq 3$. This variable is calculated by counting still water level durations in tank levels.

$$con \leq 3$$

$$CVCON = 1 - \frac{con}{3} \geq 0 \quad (5.13)$$

Constraint violations are computed in the function where I_n is computed. This is done to reduce the number of EPANET runs and to improve performance. Whenever a network is formed with diameters and tanks in the memory, it is run once and all the necessary results are stored in a MATLAB variable that is passed to other functions.

5.5. Clustering – Hypervolume – Compromise Programming

Clustering, hypervolume calculations and compromise programming operations are performed as they are explained in Sections 3.3.7, 3.4 and 3.5.

5.6. Stopping Criteria

Rather than using a simple stopping criterion like fixed number of iterations, a more complex criteria is preferred. In this criterion a counter starts counting up to a predefined number of iterations whenever all members reach the first front. Referring to Figure 5.8, full feasible first front search is exploited to a predefined number of iterations up to all members are distinct and in the first front.

CHAPTER 6

CASE STUDIES

Three different networks are tested using the developed methodology. These networks are 2-loop network (Alperovits and Shamir 1977), Hanoi network (Fujiwara and Khang 1990) and Ankara N8.3 Network (Ar 2011; Çağlar 2015; Doruk 2001; Eker 1998; Gençoğlu 2007, 2015; Güç 2006; Keleş 2005; Nadiroğlu 2014; Nohutçu 2002; Özkan 2001, 2008; Şendil 2013; Yağız 1999; Yıldız 2011).

First two networks are fictitious networks from the literature. N8.3 network is a real water distribution network from Ankara.

These networks are tested with various population sizes; then their hypervolume values are compared to find an optimal population size. After having determined the optimum population size for each network, compromise programming is applied to select one solution among the final pareto optimal solutions.

6.1. 2-Loop Network

6.1.1. Network Data

Proposed sample network is a gravity fed network with a single source, R1 (Alperovits and Shamir 1977). It feeds six consumer junctions with elevations and average nodal demand values given in Table 6.1. The network has eight pipes, each having a length of 1000 m with a Hazen Williams roughness coefficient of 130. The layout of the network is presented in Figure 6.1.

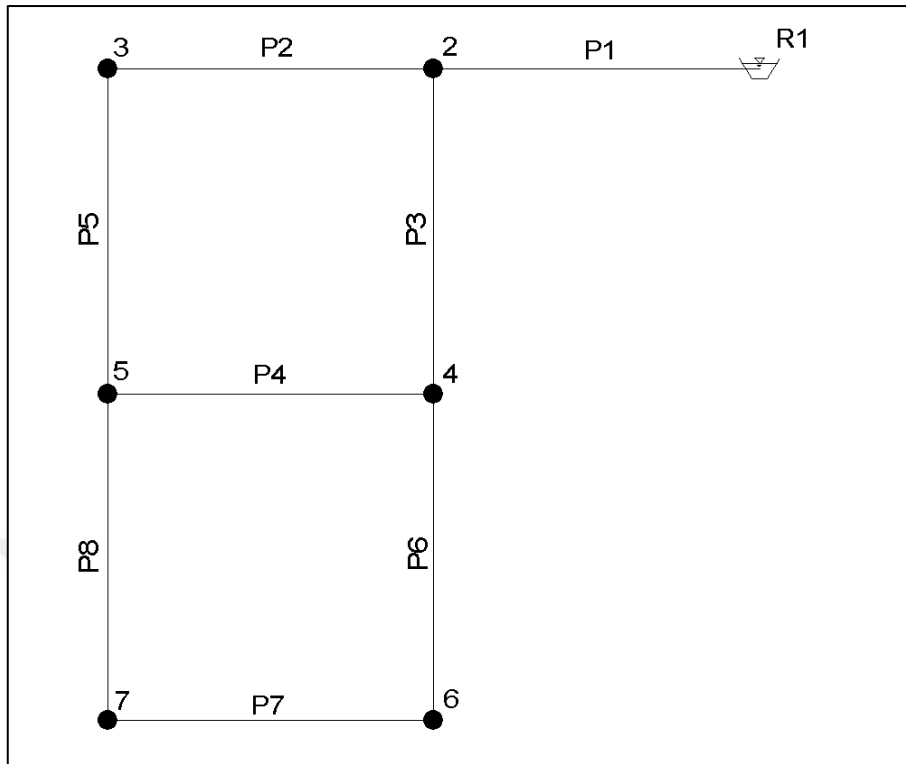


Figure 6.1. Layout of 2-loop Network

Table 6.1. 2-loop Network Characteristics

Node ID	Elevation (m)	Demand (m ³ /h)
R1	210	–
2	150	100
3	160	100
4	155	120
5	150	270
6	165	330
7	160	200

In the original work, this sample network is used to find minimum cost designs by optimizing pipe diameters. A minimum operating pressure of 30 meters is required and steady-state simulations are considered. Also, the original study does not contain

cost data for storage tanks. Therefore, several modifications to the original problem are made. These modifications are, changing original unit pipe costs, increasing number of pipe diameter choices, adding a tank cost function based on volume and adding a daily demand curve for conducting a 24-hour EPS.

Abovementioned modifications are adopted from Anytown network and presented in Table 4.8 (unit pipe costs), Figure 4.12 (tank costs) and Figure 5.14 (daily demand pattern). Diameter choices are completed to a total of 16 by adding 6 new diameters (1, 2, 3, 4, 22 and 32 inches). Unit costs of these six new diameters are interpolated and extrapolated linearly. Unit cost for “development” areas from Table 4.8 are taken into account because they are the unit prices defining pipes to be laid to a new settlement area. As a result, EPS and tank design capabilities are added to the original 2-loop network.

With the formula given in Section 4.3, size of search space of 2-loop network for 1, 2 and 3 tanks are given below:

- 1 Tank / ($n_j=6, n_p=8, n_t=1, n_{pd}=16$) **SS = 4.329×10^{20}**
- 2 Tanks / ($n_j=6, n_p=8, n_t=2, n_{pd}=16$) **SS = 2.165×10^{21}**
- 3 Tanks / ($n_j=6, n_p=8, n_t=3, n_{pd}=16$) **SS = 8.659×10^{21}**

Daily water demand of 2-loop network is 27664 m³ and balancing flow amount is 784 m³.

6.1.2. Implementation

Methodology defined in Chapter 5 is applied to the network. Objective functions are given in formulas 4.1 and 4.2. Constraints are defined in formulas 5.9 to 5.13. Crossover is applied 98% of the elements in the population, while 2% of the population is mutated to keep the child population size equal to the parent population size.

While initial population is being formed, number of tanks are externally set. Tank counts are not treated as variables. Different solutions for 1, 2 and 3 tanks are obtained.

For every tank count, five population sizes are set. Population sizes are 50, 100, 150, 200 and 250.

For each tank count and population size configuration, a complete run is performed.

Elements are allowed to evolve through generations with a minimum of 500 and a maximum of 2000. The iterations are performed until every element is feasible and on the first non-dominated front. When these criteria are met, iterations are stopped.

After optimization runs, hypervolumes and compromise programming codes are run. A normalization by 10^6 is applied to the cost values to make a better comparison between different solutions. For every configuration, a hypervolume value is obtained. Design alternatives that have higher hypervolume values are considered as better than others.

After measuring performance of the 15 solutions in groups of 5 (for every tank count), optimum population size for a tank count is determined as the run having maximum hypervolume. After selecting population size, the trial having maximum hypervolume is determined as the best run. And the preferred element in that best run is selected as the preferred solution for a specific tank count.

6.1.3. Results

The results are summarized in Table 6.2. Also, summary statistics for every run is given in Table 6.3. Following the tables, a figure for pareto front and hypervolume is provided for the selected population size of one tank solutions. Rest of the similar figures are given in Appendix A from Figure A1 to Figure A14. In the plots hypervolumes are given with black lines, every member created through the optimization run are given with blue dots, final pareto approximate elements are

given by green dots, imaginary corner points of the hypervolumes are given with red dots, imaginary best solution is given with a bigger magenta dot, preferred solution is given with a red circle and a red line indicating the normalized shortest distance between the ideal point and the preferred solution. Given definitions are visually represented in Figure 6.2.

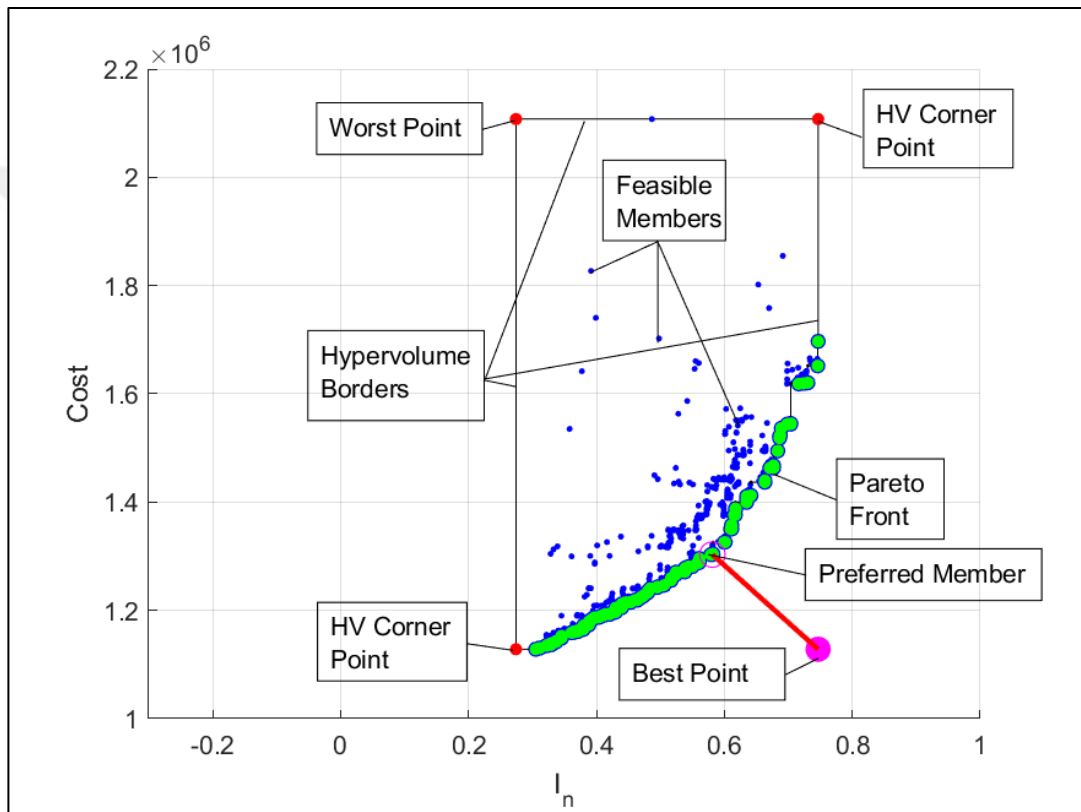


Figure 6.2. Representations for Final Pareto Curve and Hypervolume Plots

The solutions are coded as 2LXXXPYT in which XXX denotes the population count and Y indicates tank count. For example, 2L50P1T means trials performed with 50 population and 1 tank.

Table 6.2. Summary Results of 2-loop Network

Tank Count	Population Size	HV	Selected Member	I_n	Cost	Average Pipe Cost	Average Pipe Diameter (mm)	Average Tank Cost	Average Tank Volume (m ³)
1	50	0.333	19	0.4983	1414414	1270775	417	251440	800
	100	0.262	37	0.5531	1322384	1049229	358	251483	800
	150	0.381	29	0.5812	1303061	1023823	354	250576	796
	200	0.190	25	0.5395	1287088	1066361	367	252317	804
	250	0.353	191	0.5860	1333967	1109710	378	255809	821
2	50	0.484	43	0.3065	1447921	1138387	369	300537	784
	100	0.544	58	0.4738	1695208	1310311	437	477834	1784
	150	0.661	96	0.4780	1539460	1184992	402	404228	1229
	200	0.747	175	0.4300	1294428	979540	338	308120	811
	250	0.321	241	0.4400	1383930	1056483	367	410166	1241
3	50	0.086	25	0.3334	1381709	1009020	344	376515	931
	100	0.255	99	0.4322	1524507	1042960	363	661222	2336
	150	0.258	99	0.3675	1318954	1050227	350	350992	857
	200	0.287	157	0.3438	1319845	894569	318	543381	1687
	250	0.548	235	0.3532	1421579	1035741	353	525844	1551

Table 6.3. Summary Statistics of 2-loop Network

Tank Count	Population Size	Total Iteration	Total Run Time (s)	Average Iteration Time (s)	Final Percent of F1 Elements
1	50	500	577	1.15	100.0
	100	500	1342	2.68	100.0
	150	500	2146	4.29	100.0
	200	500	3188	6.38	100.0
	250	643	6084	9.46	100.0
2	50	500	490	0.98	100.0
	100	500	1218	2.44	100.0
	150	500	2184	4.37	100.0
	200	967	6641	6.87	100.0
	250	500	4844	9.69	100.0
3	50	500	635	1.27	100.0
	100	1129	3222	2.85	100.0
	150	903	4275	4.73	100.0
	200	992	6846	6.90	100.0
	250	1089	10474	9.62	100.0

In Figure 6.3, final pareto curve for 150 population 1 tank solutions. Remaining 14 pareto curves are given in Appendix A from Figure A1 to Figure A14.

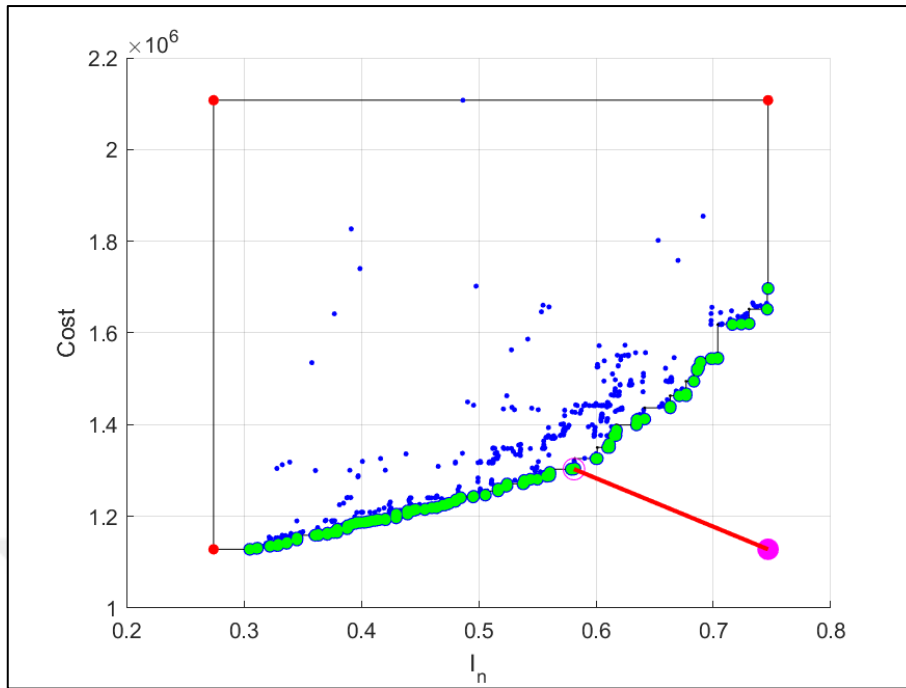


Figure 6.3. Pareto Curve for 2L150P1T

Three designs are selected as the preferred solution, namely:

- Solution 29 from 150 population 1 tank configuration
 - $I_n = 0.5812$ Cost= 1303061
- Solution 175 from 200 population 2 tank configuration
 - $I_n = 0.4300$ Cost= 1294428
- Solution 235 from 250 population 3 tank configuration
 - $I_n = 0.3532$ Cost= 1421579

The design values (decision variables) for the preferred solutions are given in Table 6.4 and Table 6.5.

Table 6.4. Pipe Design Values for the Preferred Solutions

Pipe ID	Diameter (mm)		
	2L150P1T/29	2L200P2T/175	2L250P3T/235
P1	600	500	750
P2	300	350	250
P3	500	550	400
P4	250	250	450
P5	250	250	250
P6	450	450	400
P7	350	350	75
P8	250	75	250
Total Pipe Cost	1054790	993610	1040020

Table 6.5. Tank Design Values for the Preferred Solutions

Solution ID	Tank Sites	Tank Heights (m)	Tank Diameters (m)	Tank Elevations (m)	Total Tank Volumes (m ³)	Tank Cost
2L150P1T/29	6	8.55	10.81	33.09	785	248271
	7	8.83	7.46	35.60	785	300818
2L200P2T/175	3	8.80	7.59	37.55		
	7	3.14	11.36	32.42	963	381559
2L250P3T/235	6	3.29	14.03	32.67		
	3	3.10	7.47	40.97		

Two graphs are given for 29th solution of trial 2L150P1T. First one is storage tank fullness percent versus time (Figure 6.4). Tank is placed at fifth node (with respect to drawing order) with a label “6” as seen in Figure 6.1 and Table 6.5. It is seen that water level at the end of the day is higher than water level at the beginning of the day, water levels do not stay still more than three hours, tank fill and empty cycles

do not occur more than five times and tank is filled and emptied almost fully through the day. When these conditions are met and all constraints are satisfied. With the satisfaction of constraints, good operation of a water tank is performed for the defined network.

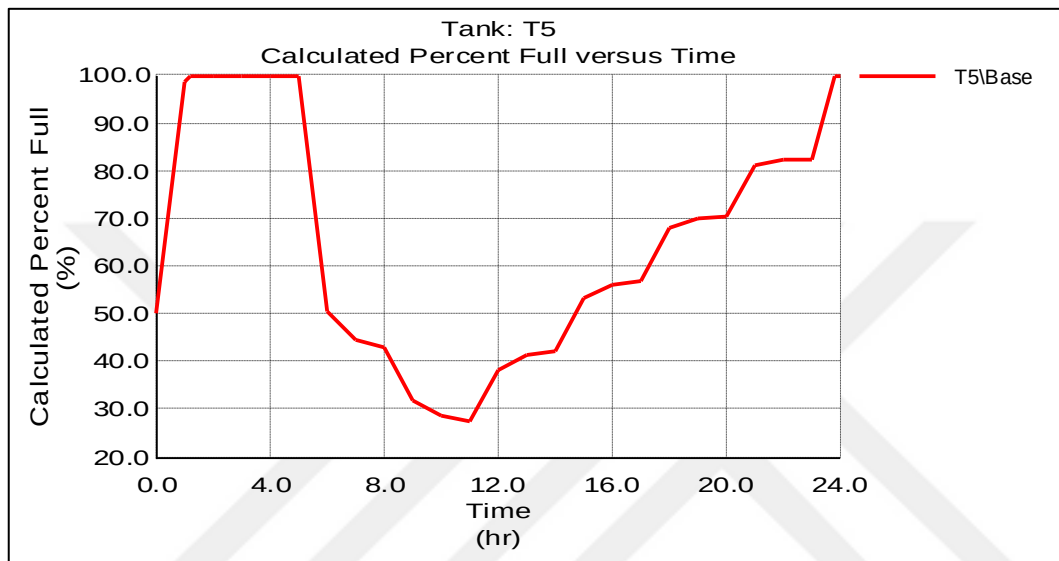


Figure 6.4. Tank Oscillations for Solution 2L150P1T/29

In the second graph, nodal pressures versus time are plotted (Figure 6.5). This graph shows that pressures below 30 meters do not occur in the network through 24 hours meaning that all consumer nodes are supplied with enough pressurized water. This condition satisfies another constraint.

Other tank oscillation and pressure variation graphs are provided in Appendix A for solutions 2L200P2T/175 and 2L250P3T/235 from Figure A15 to Figure A18. In those graphs it is again clear that all constraints are satisfied and good operations are performed with even more tank counts.

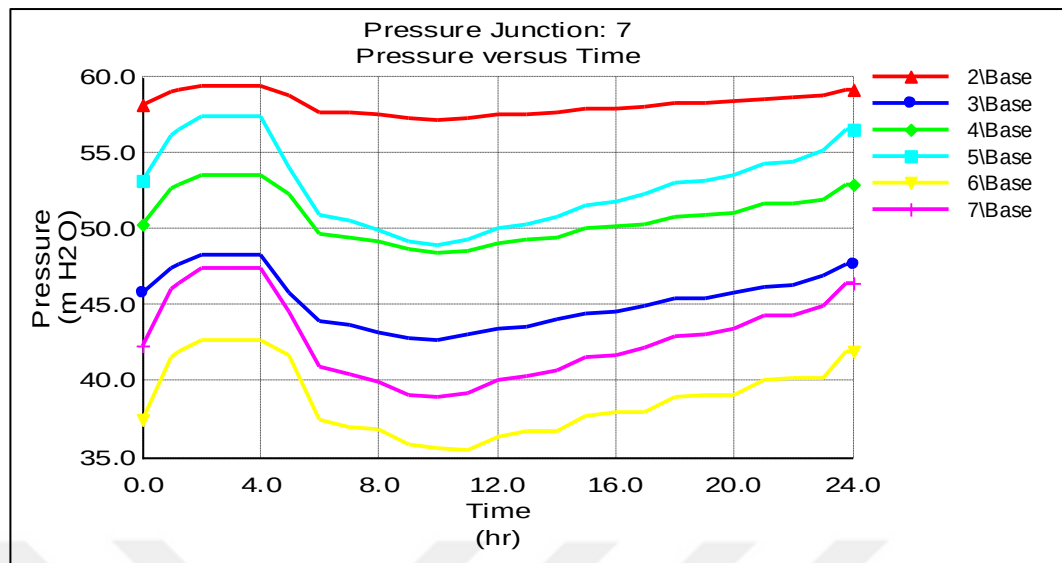


Figure 6.5. Pressure Variations for Solution 2L150P1T/29

A heat map of tank placements is given in Figure 6.6. In this map, a colour ramp from red to green is defined and darker green indicates most preferred tank sites whereas yellow and orange colours indicate a lesser preferred site. Red colour indicates a site that is not preferred. Accordingly, node 7 is the most preferred tank site, and it is followed by nodes 6 and 3. Nodes 4 and 5 are less preferred but node 2 is not preferred at all. This can be interpreted as tank sites far from the source are more preferred, as tanks act as a second source point and provide water nearby from shorter paths. This is a good practice to increase strength of a system far from the source point in order to provide good service to consumers.

In addition, most preferred sites are not the most water demanding nodes, which are nodes 4 and 5. More than 50 % water is demanded at these nodes but tanks are not placed on them. So, water demand amount does not seem to affect tank sites directly.

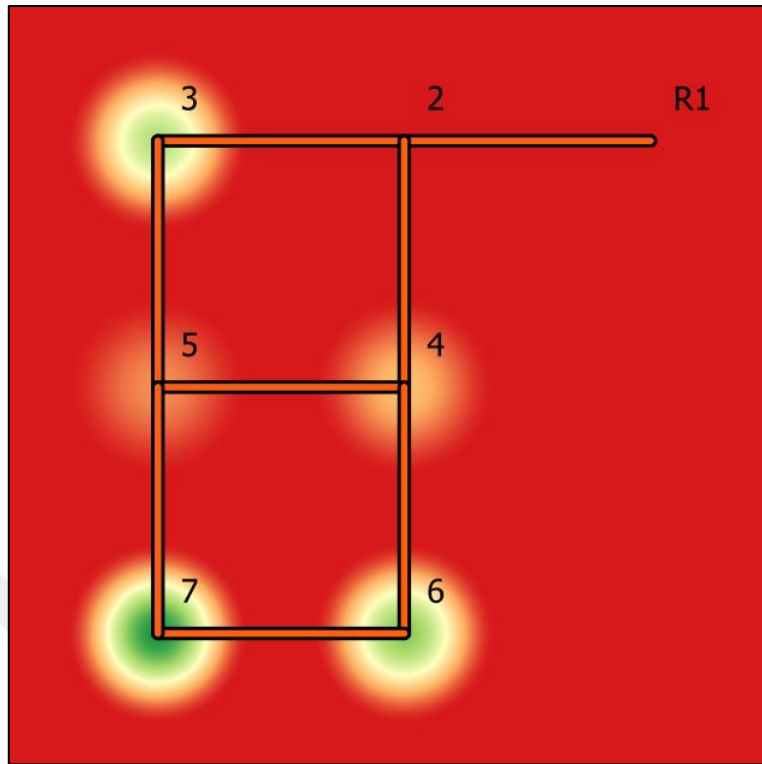


Figure 6.6. Heat Map of Tank Placements for 2-Loop Network

In a final graph given Figure 6.7 in, behaviour of the members of the population through generations are presented. Horizontal thick lines represent number of feasible elements in an iteration, while thinner lines represent number of first front elements for that generation. All population sizes are plotted on the same graph. It is seen that except for 250 population, all solutions ended in 500 generations and all solutions reached quickly to full feasible search defined in Figure 5.8. It is also seen that elements reach full first front search phase later when higher population sized are selected. This is because gene pool gets larger it slows down the search speed while diversity is increased and there is a higher probability to find better elements. However, with smaller population sized this time there are lesser choices for gene selection, search speed increases but diversity is negatively affected which leads to a lower probability of finding better elements. That is why an optimum population size for a specific network design is looked for through this chapter.

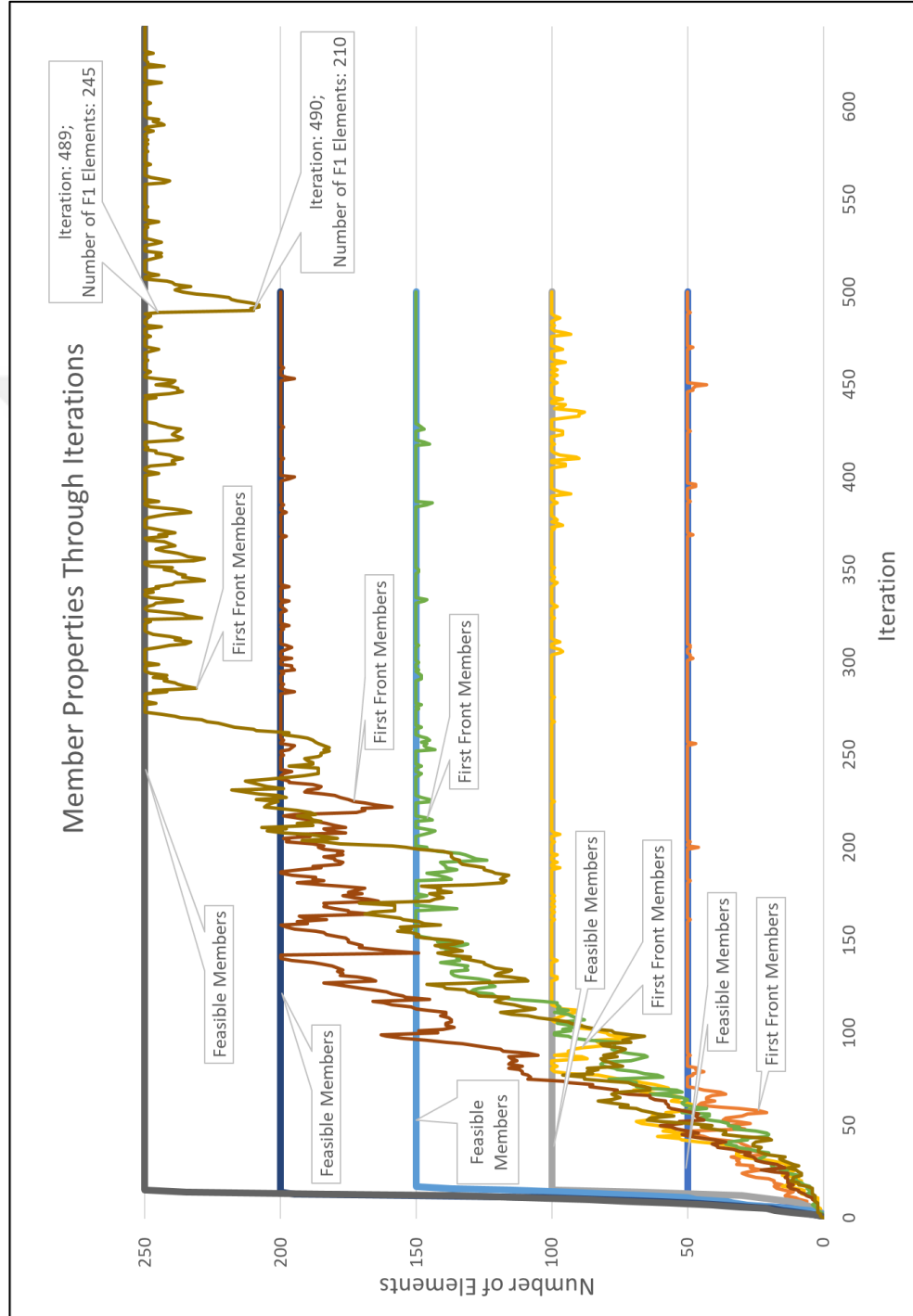


Figure 6.7. Member Properties for 2L1T

Similar graphs to Figure 6.7 for two and three tank solutions are given in Appendix A in Figure A19 and Figure A20. With the use of Figure 6.7, it becomes easier to track the effects of crossover and mutation operators through generations. By observing their effects, it is possible to fine tune relevant parameters of the two operators. In addition, it is possible to check the NSGA-II mechanism for possible improvements.

An example for tracking the changes between two generations, 2L250P1T solution set is selected and a change in first front member amount is marked in Figure 6.8. At the marked iterations, namely 489 and 490, number of first front elements drop from 245 to 210. This means at a result of some crossover and/or mutation operations new child members are formed and these new members dominated some old parent members pushing them to lower fronts and taking their place in the first front. This situation is traced down in the population archive data and it is seen that

- 245 first front and 5 second front elements are present at iteration 489,
- 210 first front and 40 second front elements are presents at iteration 490,
- 18 members from iteration 489 dropped down from the best 250 list and 6 first front, 12 second front new members are formed in iteration 490,
- 24 first front members in iteration 489 became second front members and 4 second front members are still second front members in iteration 490,

In Figure 6.8, from a narrow part of the I_n vs Cost diagram of first front elements between iterations 489 (blue line) and 490 (red line) can be observed. In the figure it is possible to observe the newly formed first front elements (yellow circles), preserved elements due to elitism (green circles), and dominated members that are not present in the new first front (orange squares). When the population data is investigated in detail, it is seen that presented new first front elements are results of various crossover operations. It is seen that 3 new members took place of 20 old members while making the search for a pareto front.

Examples for mutation-formed elements are also found but their changes are too small to represent on a figure so they are not included here.

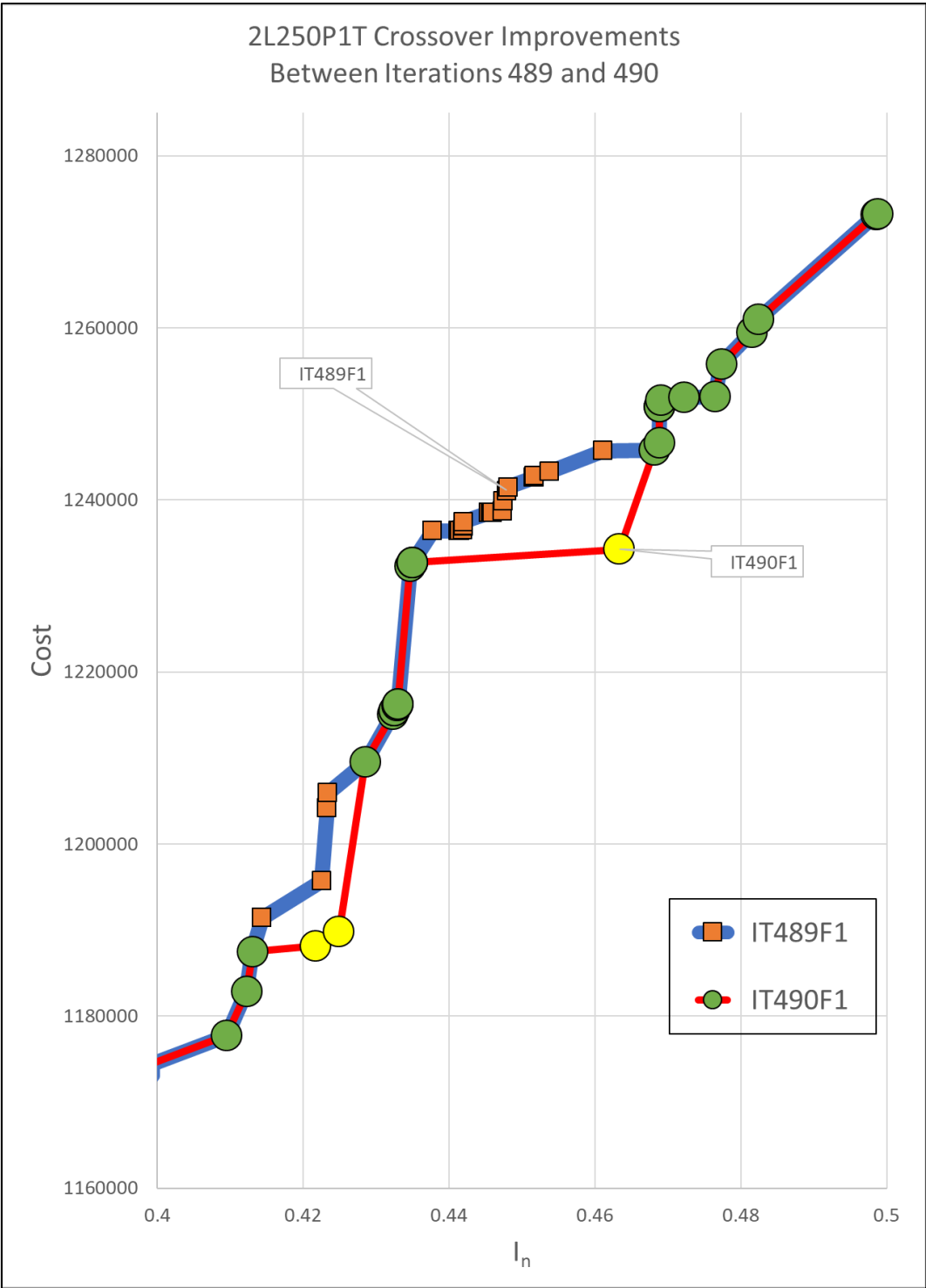


Figure 6.8. Crossover Improvement Example Between Two Iterations

6.1.4. Comparison with Published Works

It is not possible to directly compare the obtained results with similar studies from the literature because unit cost values are adopted from another study to be able to include storage tank optimization. To overcome this difficulty, several trials are performed and three comparisons are made possible.

To make the first comparison, developed methodology is tried by omitting tanks, optimizing only pipe diameters and using original unit pipe costs. By removing tank optimization, all tank related constraints are also omitted and only original constraint which is the minimum pressure requirement of 30 meters is applied. Also, the EPS nature of the developed methodology is converted to a single steady state solution. A population of 150 is tried and compared with the true pareto curve obtained for 2-loop network (Wang et al. 2015). Results of the comparison is given in Figure 6.9. With this comparison, it seems that the developed methodology works correctly and provides similar solutions with the ones published. The compared curves separate from each other on the high-cost region. This may be improved by fine tuning mutation and constraint parameters as the true pareto curve is obtained from a much different algorithm with different parameters.

Additionally, results of four single-objective optimization (cost minimization) results from two different studies are plotted on the same graph. One of these studies is the original paper that defined 2-loop network (Alperovits and Shamir 1977). Second one is a cost minimization study that uses genetic algorithms (Gençoğlu 2007). Final designs of these two studies are used to calculate I_n values and plotted as points on Figure 6.9. As both of the studies focus on cost minimization, their results appear on the lower left part of the true pareto front having low cost and low network resilience values. The original study's final solution is dominated by the true pareto front and Gençoğlu's (2007) solution. On the other hand, Gençoğlu (2007) managed to find a point on the low-cost region of the true pareto curve. This

is the only possible thing when single objective optimization is considered. It is not possible to find higher cost solutions on the curve.

Hydraulic performances over minimum pressures are compared for solutions given in Figure 6.9. To compare the findings of the developed methodology, two end solutions from the pareto curve are included as comparison (2L150P0T/200 low cost – low strength, 2L150P0T/199 high cost – high strength). Results are tabulated in Table 6.6. As predicted, four designs from the literature has minimum pressures around 30 meters. However, it is possible to produce many different designs having minimum pressures between 30.45 and 47.52 meters.

Table 6.6. Minimum Pressure Comparison

Solution ID	Minimum Pressure (m)
2L150P0T/199	47.52
2L150P0T/200	30.45
Alperovits & Shamir (1977)	30.045
Gençoğlu (2007)	30.803

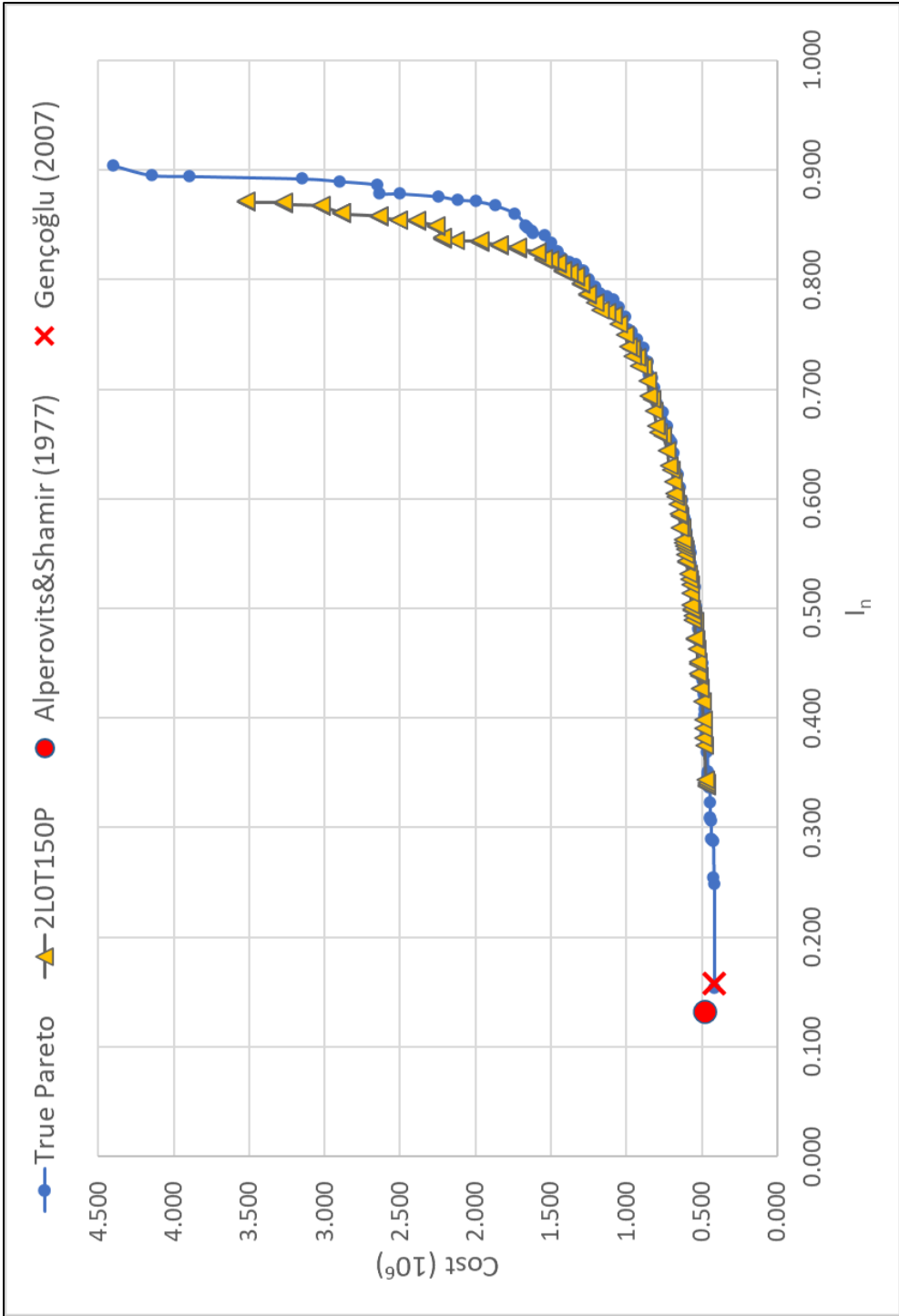


Figure 6.9. Comparison with True Pareto Curve

Second comparison is performed by using the developed methodology with modified pipe unit costs and different constraint combinations. With these trials, the effects of constraints are clearly observed, and it is concluded that by placing constraints, pareto curves within an envelope of solutions are obtained.

In this comparison, constraints including tank parameters are removed one by one and resulting pareto curves are plotted (Figure 6.11). Removed constraints are summarized for each step in Table 6.7. Only one constraint is left at the end of trials. This last constraint, CVPRES is the minimum pressure constraint defined as 30 meters in the original study. Releasing the problem from constraints moves the final pareto front towards south-east direction of the graph. Also, in Figure 6.10 tank fullnesses of the seven trials can be observed.

Table 6.7. Removed Constraints for Comparison

Solution Name	Removed Constraints	Remaining Constraints
7 Constraints	–	CVLV, CVEV, CVLSE, CVVD, CVCON, CVOSC, CVPRES
6 Constraints	CVLV	CVEV, CVLSE, CVVD, CVCON, CVOSC, CVPRES
5 Constraints	CVLV, CVEV	CVLSE, CVVD, CVCON, CVOSC, CVPRES
4 Constraints	CVLV, CVEV, CVLSE	CVVD, CVCON, CVOSC, CVPRES
3 Constraints	CVLV, CVEV, CVLSE, CVVD	CVCON, CVOSC, CVPRES
2 Constraints	CVLV, CVEV, CVLSE, CVVD, CVOSC	CVCON, CVPRES
1 Constraint	CVLV, CVEV, CVLSE, CVVD, CVOSC, CVCON	CVPRES

Within the removal of first three constraints, a distinct difference between solutions is not observed (Solutions “7 Constraints” to “4 Constraints”). The designed tank volumes in first four trials are between 785 m³ and 790 m³. But when Figure 6.10 is analysed, it can be seen that after removal of the CVLV and CVEV constraints,

storage tank do not fill up to its full capacity. After removing third constraint, CVLSE, water levels lower than initial levels at the end of day becomes feasible. In Figure 6.10, 6.53 m starting level decreases to 6.42 m at the end of 24-hours.

Removal of fourth constraint, namely CVVD, which forces the volume of the designed tank to be greater than the balancing volume of the system results in a much different pareto front than the others. Although the tank seems to fill and empty during the day, design volume in this case decreases almost by 40% to 496 m³. This means a smaller reserve volume and balancing volume is created which may not solve water shortage problems during a failure event. When CVVD is released, smaller storage tanks become feasible and preferred in the designs.

When the constraint CVOSC, which prevents excessive fill & empty oscillations is released, tank volumes decrease approximately by 50% down to 257 m³. From Figure 6.10, it is possible to observe 21 filling and 20 emptying cycles which make a storage tank unsuitable for operation. This many cycles also indicate the insufficient volume of the design. In the pareto curve comparison graph, releasing CVOSC generates a distinct set of solutions than the rest of the trials.

Lastly, the constraint CVCON which forces water levels not to stay still more than three hours is removed, resulting pareto curve makes the biggest move in the plot. With this last release, storage tank volume decreases to the minimum allowable value of 58.9 m³ defined in Table 4.10. This means that a storage tank is not recognized as a useful element within the absence of constraints and to minimize its cost, the minimum allowable tank is chosen. With such a small volume and no constraints, designed tank empties immediately at the beginning of the day and stays at that level for 24 hours. With an element that does not store and balance flows in the system it is again an inoperable structure.

CVCON is released after CVOSC because within the absence of CVCON, CVOSC is automatically satisfied as very long (even 24-hour) still water levels are allowed in the tanks, having excessive water level oscillations is not very possible.

It can be said that the developed constraint handling methodology is most sensitive to constraint CVCON. Then the methodology is sensitive to CVOSC and CVVD. From the provided solutions, it is not possible to determine which constraint the methodology is sensitive after these three constraints.

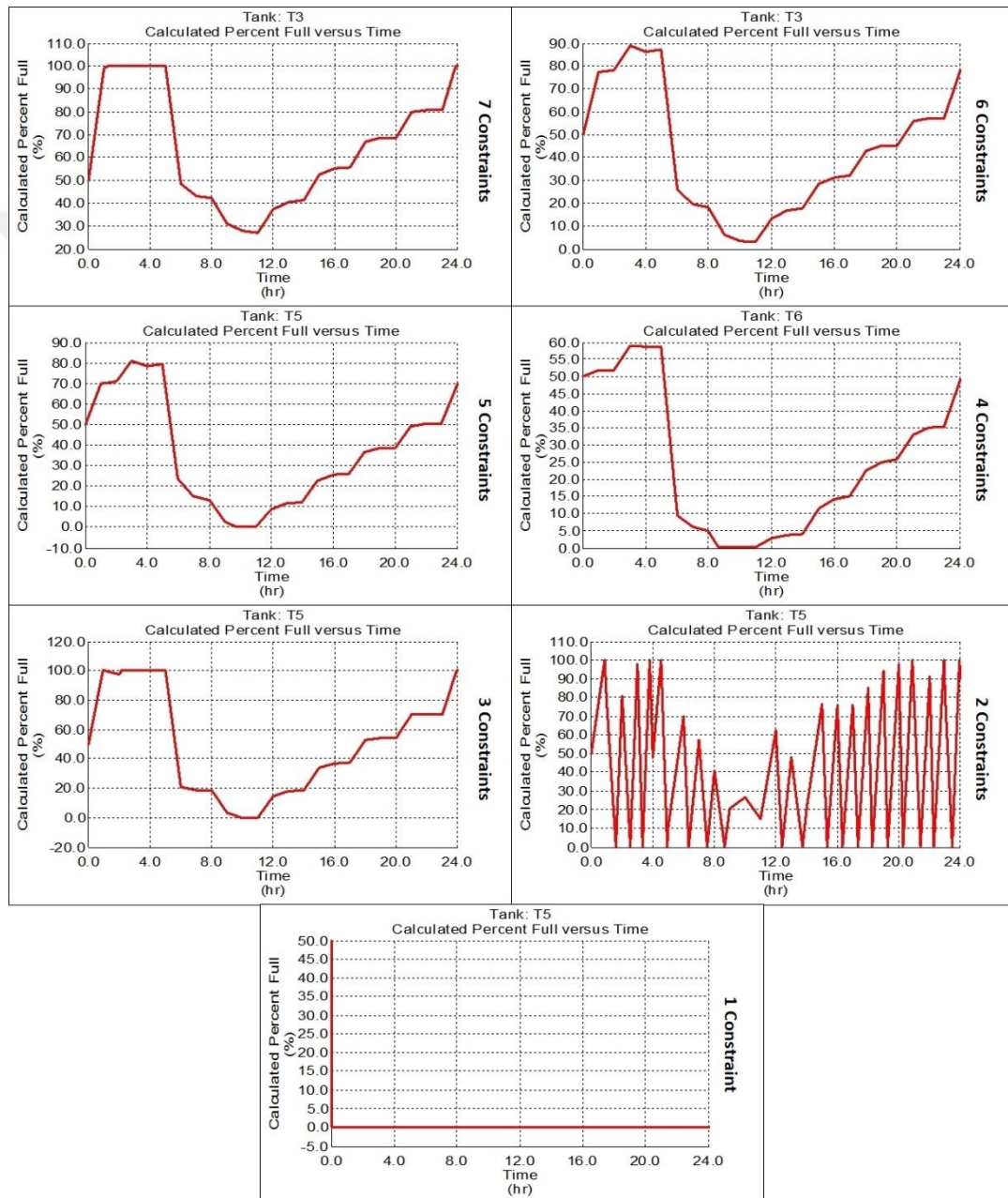


Figure 6.10. Tank Fullness Graph of Seven Different Constrained Solutions

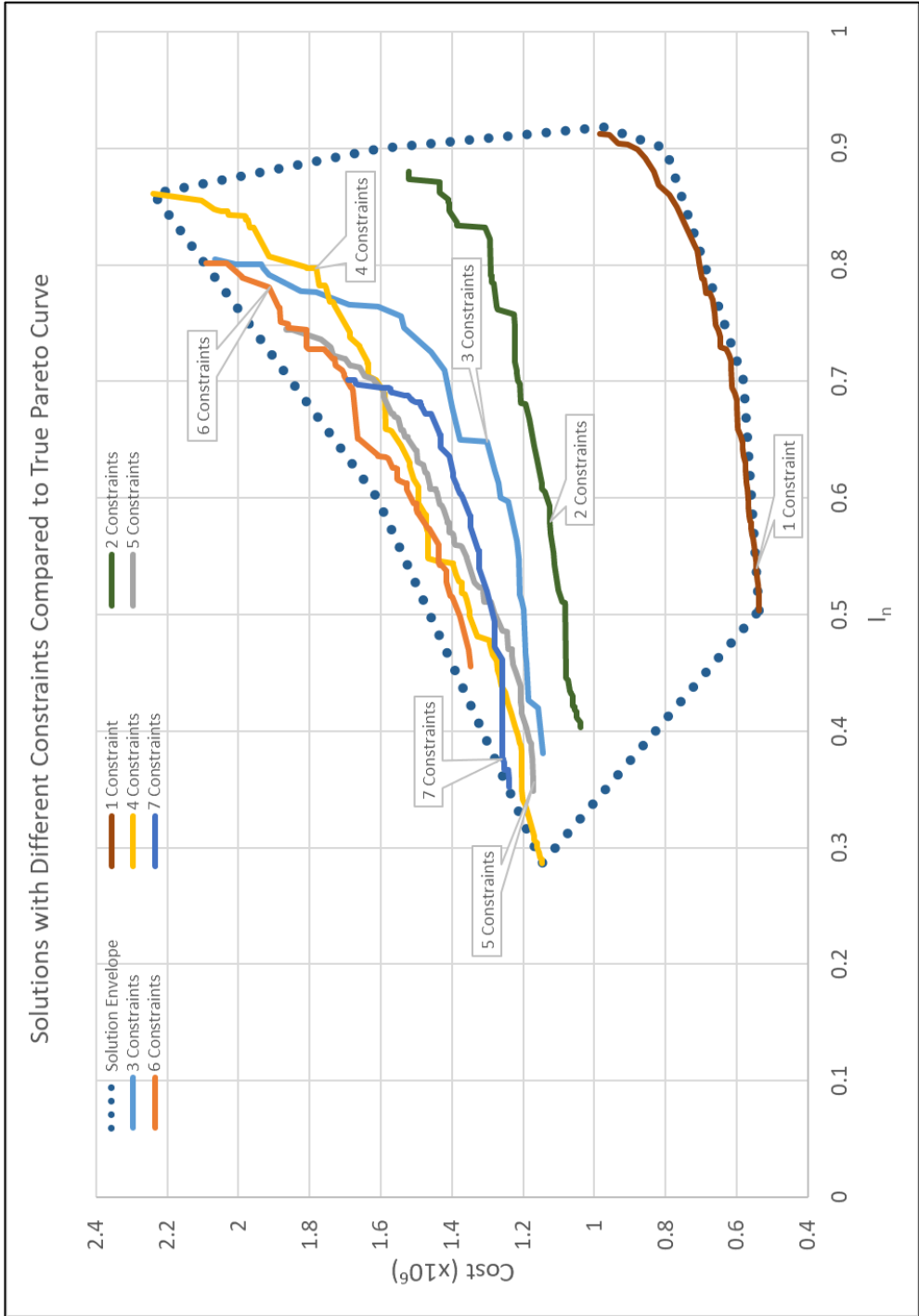


Figure 6.11. Comparison of Envelope of Solutions with the True Pareto Curve

Third and final comparison is made by again constraint removing actions but with original pipe unit costs and zero tank costs to make results comparable to the true pareto front. This time, different tank counts are also included, and for each tank count optimum population sizes obtained in Section 6.1.3 are used. Solution codes ending with “C” means all the constraints are present in those solutions (constrained cases). Solutions ending with “U” means only one constraint, CVPRES is present in the solutions (unconstrained cases). Figure 6.12 presents the comparison results.

It should be noted that true pareto front is obtained without a storage tank inclusion and under steady-state solutions. On the other hand, trial solutions are performed with a single tank inclusion and an EPS solution. True pareto front is calculated from a population having 114 members, while the solutions compared to it are performed with their optimum population sizes.

It is clearly observed that, by including constraints and increasing their number moves the final solutions far from the true pareto curve. On the other hand, when only original constraints are included, final solutions overpass the true pareto curve and seem to provide much better designs. However, these unconstrained cases provide operationally bad designs in which storage tanks do not actually have any role in the system. This issue is covered in more detail above.

Investigating Figure 6.12, it can be said that, by fine-tuning constraint parameters or releasing some of the constraints, it can be possible to achieve solutions on the true pareto front. In addition, it can be said that it becomes harder to find diverse solutions when tank counts are increased. Although tank count increase greatly widens the search space, it is not possible to say that feasible search space is enlarged with similar scales. Therefore, shorter pareto curves are obtained when tank counts are increased.

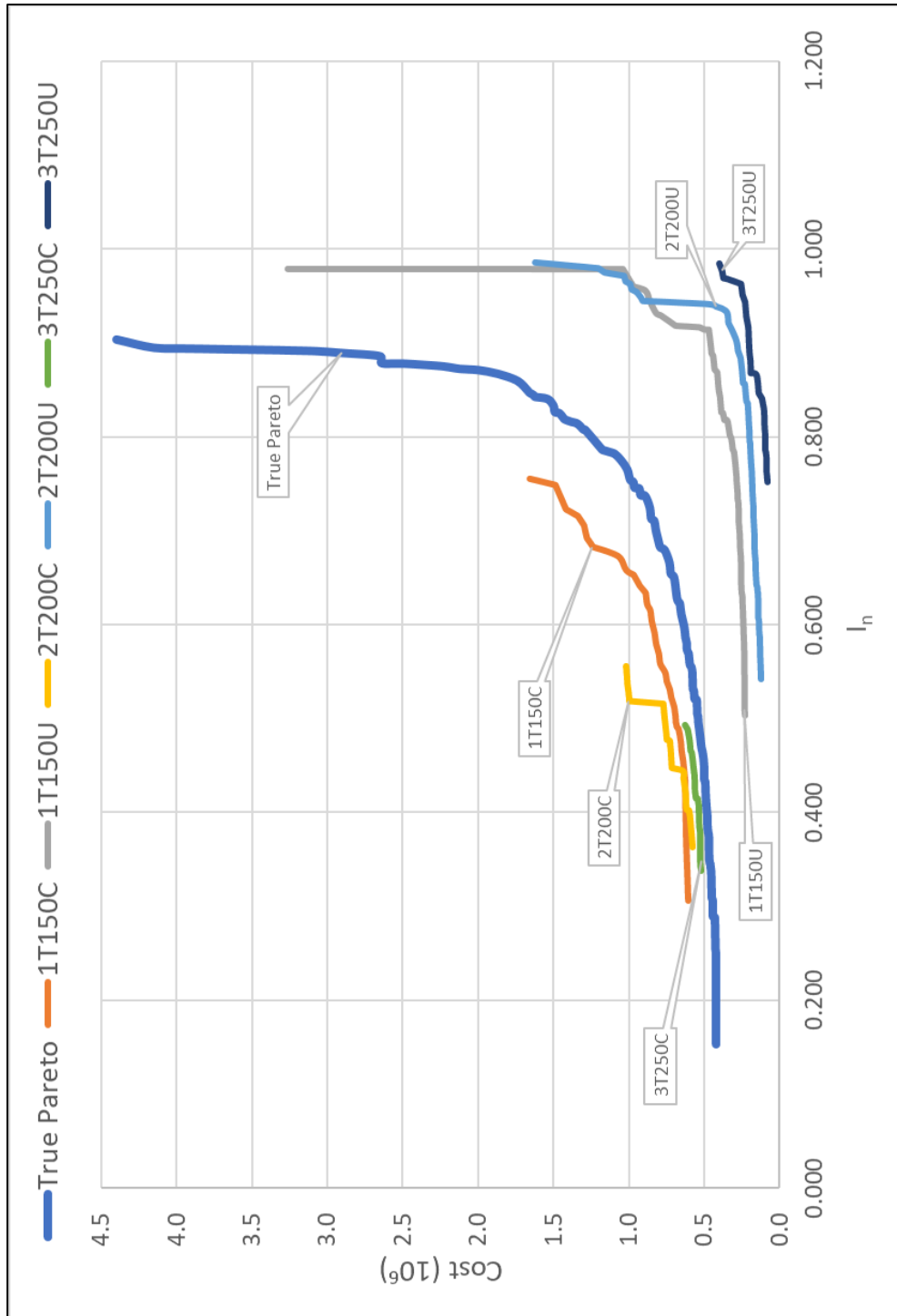


Figure 6.12. Comparison of Constraint Effects to the True Pareto Curve

Table 6.8. Hanoi Network Characteristics

Junction ID	Demand (m ³ /hr)	Junction ID	Demand (m ³ /hr)	Pipe ID	Length (m)	Pipe ID	Length (m)
2	890	23	1045	P1	100	P22	500
3	850	24	820	P2	1350	P23	2650
4	130	25	170	P3	900	P24	1230
5	725	26	900	P4	1150	P25	1300
6	1005	27	370	P5	1450	P26	850
7	1350	28	290	P6	450	P27	300
8	550	29	360	P7	850	P28	750
9	525	30	360	P8	850	P29	1500
10	525	31	105	P9	800	P30	2000
11	500	32	805	P10	950	P31	1600
12	560			P11	1200	P32	150
13	940			P12	3500	P33	860
14	615			P13	800	P34	950
15	280			P14	500		
16	310			P15	550		
17	865			P16	2730		
18	1345			P17	1750		
19	60			P18	800		
20	1275			P19	400		
21	930			P20	2200		
22	485			P21	1500		

Similar modifications are applied to Hanoi network as 2-loop network defined in Section 6.1.1. Diameter choices are completed to a total of 16 by adding 10 new diameters (14, 18, 22, 26, 28, 32, 34, 36, 38 and 48 inches). Unit costs of these 10 new diameters are interpolated and extrapolated linearly. When compared to 2-loop network bigger diameter choices are preferred to use consistent diameters with the original study. Also, the water consumption in Hanoi network is much bigger than

2-loop network, therefore providing water to Hanoi consumers is only possible by using bigger diameters.

As a result, EPS and tank design capabilities are added to the original Hanoi network.

With the formula given in Section 4.3, size of search space of Hanoi network for 1 to 5 tanks are given below:

- 1 Tank / ($n_j=31, n_p=34, n_t=1, n_{pd}=16$) – SS = 5.537×10^{52}
- 2 Tanks / ($n_j=31, n_p=34, n_t=2, n_{pd}=16$) – SS = 1.361×10^{54}
- 3 Tanks / ($n_j=31, n_p=34, n_t=3, n_{pd}=16$) – SS = 3.947×10^{55}
- 4 Tanks / ($n_j=31, n_p=34, n_t=4, n_{pd}=16$) – SS = 1.105×10^{57}
- 5 Tanks / ($n_j=31, n_p=34, n_t=5, n_{pd}=16$) – SS = 2.984×10^{58}

Daily water demand of Hanoi network is 492517 m^3 and. balancing flow amount is 13958 m^3 .

6.2.2. Implementation

Same implementation details given in Section 6.1.2 are applied. The only difference is that instead of trying up to three tank configurations, Hanoi network is optimized by trying up to five tank configurations. As there are much more available nodes in Hanoi network more tanks can also be tried. This makes a total number of solutions of 25 (5 tanks x 5 population sizes). This time a normalization of 10^7 to cost values are applied.

6.2.3. Results

The results are summarized in Table 6.9. Also, summary statistics for every run is given in Table 6.10. Following the tables, a figure for pareto front and hypervolume (Figure 6.14) is provided for the selected population size of one tank solutions. Rest of the similar figures are given in Appendix B.

The solutions are coded as HANXXXPYT in which XXX denotes the population count and Y indicates tank count. For example, HAN50P1T means trials performed with 50 population and 1 tank.

In Figure 6.14, it can be seen that there are three major groups of elements during the optimization process. Careful examinations of pop variable between generations reveal that, there are great increments in I_n values due to evolving mechanisms. A similar example is shown in Section 6.1.3 and Figure 6.8.

It is also noticeable that in Figure 6.14, three groups of evolving solutions form pareto curves of approximately 0.1 units of width in I_n axis ($\sim 0.3 - \sim 0.4$ in the first group, $\sim 0.4 - \sim 0.5$ in the second group and $\sim 0.5 - \sim 0.6$ in the third group). This seems reasonable when best estimate pareto curve from Wang et al. (2015) is taken into consideration. The curve obtained by Wang et al. (2015) has an I_n variation approximately between 0.2 and 0.35 having a width of 0.15 units.

Table 6.9. Summary Results of Hanoi Network

Tank Count	Population Size	HV	Selected Member	I _n	Cost	Average Pipe Cost	Average Pipe Diameter (mm)	Average Tank Cost	Average Tank Volume (m ³)
1	50	0.786	50	0.4382	38276009	19115284	898	36803057	16471
	100	0.267	100	0.2628	39011379	19410338	946	40634691	16928
	150	1.279	150	0.5101	34728562	17367053	881	33386958	15981
	200	0.447	199	0.5203	33761499	15202649	813	19969817	14011
	250	0.814	249	0.5497	36171568	14558842	794	20069279	14032
2	50	0.189	50	0.1849	78163233	18548814	899	89243920	35023
	100	1.665	44	0.4590	18822440	16198457	847	11409883	19089
	150	1.205	148	0.4656	17776090	16676768	825	7240769	15352
	200	2.590	75	0.5123	17797838	16097342	846	5383093	15082
	250	1.886	248	0.4272	18456955	17246976	846	18796345	21201
3	50	1.028	49	0.3624	17271907	15531317	808	4818004	19847
	100	2.810	23	0.3575	17105436	16306350	817	8584365	21772
	150	1.579	149	0.4296	18014608	17251575	849	13203995	22413
	200	2.258	200	0.4524	16151941	15689714	821	3401423	17748
	250	1.570	194	0.4654	15690462	14549318	753	2075597	14420
4	50	0.186	50	0.3166	20717661	16455018	820	10932080	29242
	100	0.562	100	0.3164	17145872	15503511	793	3415687	21122
	150	1.774	95	0.4556	15898146	14707491	769	2164753	15193
	200	1.823	28	0.4622	16735302	15360886	801	2741770	16710
	250	3.466	250	0.5318	16643608	15367082	809	2063209	14914
5	50	0.353	49	0.2884	21722758	15679341	811	15282572	38973
	100	2.435	100	0.3912	17467691	15556029	809	2562633	17970
	150	2.236	25	0.3582	17395700	15971561	826	5565407	24711
	200	1.924	197	0.3951	17128133	14962947	776	2911353	19676
	250	0.836	193	0.3626	15561510	13650464	715	2462621	16581

Table 6.10. Summary Statistics of Hanoi Network

Tank Count	Population Size	Total Iteration	Total Run Time (s)	Average Iteration Time (s)	Final Percent of F1 Elements
1	50	500	695	1.39	100.0
	100	500	1258	2.52	100.0
	150	500	2872	5.74	100.0
	200	500	3834	7.67	100.0
	250	500	5178	10.36	100.0
2	50	507	453	0.89	100.0
	100	533	1829	3.43	100.0
	150	534	2797	5.24	98.7
	200	500	3871	7.74	100.0
	250	500	5244	10.49	100.0
3	50	636	1241	1.95	100.0
	100	1258	4795	3.81	100.0
	150	647	3823	5.91	100.0
	200	810	7478	9.23	100.0
	250	636	6524	10.26	100.0
4	50	603	956	1.59	100.0
	100	847	3506	4.14	100.0
	150	1285	8150	6.34	100.0
	200	1195	10983	9.19	100.0
	250	1876	21132	11.26	100.0
5	50	737	1229	1.67	100.0
	100	1658	6729	4.06	100.0
	150	1078	7137	6.62	100.0
	200	1357	11645	8.58	100.0
	250	1414	15608	11.04	100.0

In Figure 6.14, final pareto curve for 150 population 1 tank solutions. Remaining 24 pareto curves are given in Appendix B.

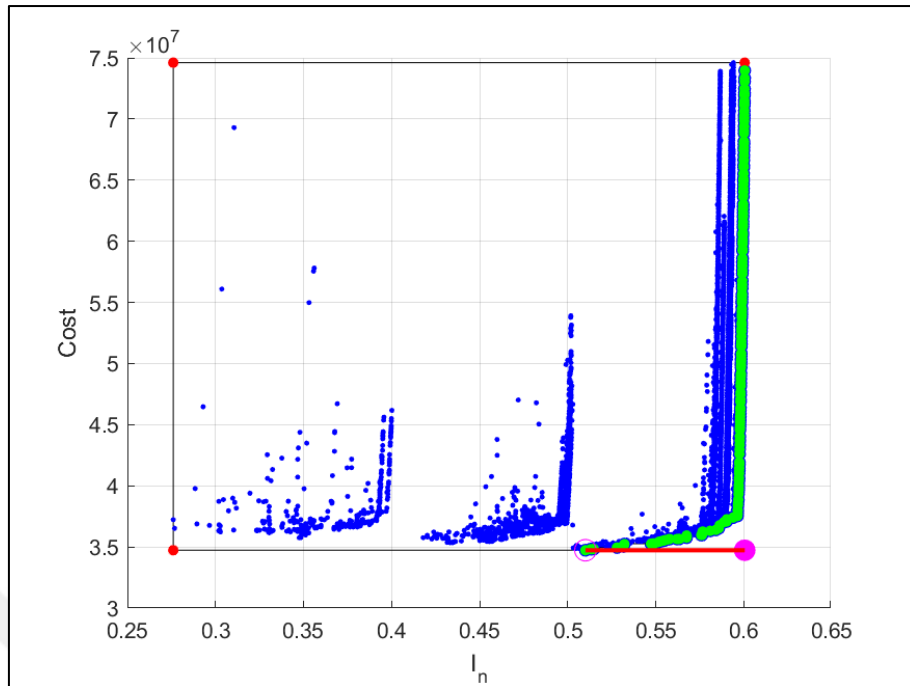


Figure 6.14. Pareto Curve for HAN150P1T

Five designs are selected as the preferred solution, namely:

- Solution 150 from 150 population 1 tank configuration
 - $I_n = 0.5101$ Cost= 34728562
- Solution 75 from 200 population 2 tank configuration
 - $I_n = 0.5123$ Cost= 17797838
- Solution 23 from 100 population 3 tank configuration
 - $I_n = 0.3575$ Cost= 17105436
- Solution 250 from 250 population 4 tank configuration
 - $I_n = 0.5318$ Cost= 16643608
- Solution 100 from 100 population 5 tank configuration
 - $I_n = 0.3912$ Cost= 17467691

The design values (decision variables) for the preferred solutions are given in Table 6.11 and Table 6.12.

Table 6.11. Pipe Design Values for the Preferred Solutions

Pipe ID	Diameter (mm)				
	HAN150P1T/150	HAN200P2T/75	HAN100P3T/23	HAN250P4T/250	HAN100P5T/100
1	1200	1200	1200	1200	850
2	1200	1200	1200	1200	1200
3	1200	1200	650	1200	1200
4	1200	1200	850	1200	1000
5	1200	1200	650	1000	1000
6	1000	1200	550	850	900
7	950	1000	800	850	950
8	950	900	550	900	950
9	650	700	500	850	700
10	650	900	800	700	800
11	800	550	850	700	800
12	700	550	1000	700	550
13	600	700	700	700	500
14	600	750	750	850	450
15	300	450	650	550	750
16	300	550	650	750	900
17	750	650	900	1000	1200
18	800	1000	850	1000	1200
19	850	1200	1200	1200	1200
20	1200	1200	1200	1200	850
21	650	600	450	800	800
22	500	600	650	850	850
23	950	1200	1200	1000	1200
24	650	500	850	550	550
25	600	350	900	450	350

Table 6.11. Pipe Design Values for the Preferred Solutions Continued

26	850	500	750	350	300
27	850	700	450	300	700
28	650	400	450	500	700
29	950	1000	850	600	650
30	850	700	450	600	700
31	750	650	400	450	700
32	650	650	300	500	700
33	600	350	550	400	450
34	700	300	900	400	350
Total Pipe Cost	15039952	14575072	14909194	14683427	15123264

Table 6.12. Tank Design Values for the Preferred Solutions

Solution ID	Tank Sites	Tank Heights (m)	Tank Diameters (m)	Tank Elevations (m)	Total Tank Volumes (m ³)	Tank Cost
HAN150P1T/150	13	14.19	35.39	47.27	13958	19688609
HAN200P2T/75	25	15	24.34	49.99	13958	3222766
	32	15	24.34	50		
HAN100P3T/23	9	11.13	25.31	34.52	15176	2196241
	12	12.21	21.13	32.55		
	14	9.74	26.31	38.43		
	31	13.76	18.25	49.98		
HAN250P4T/250	32	12.65	19.54	49.84	14007	1960179
	26	14.26	16.87	49.92		
	25	14.8	17.17	49.93		
	25	8.78	15.89	47.34		
HAN100P5T/100	28	8.62	21.27	46.76	15580	2344427
	22	8.68	23.73	49.83		
	31	8.1	24.18	43.71		
	24	8.7	21.7	47.15		

Two graphs are given for 150th solution of trial HAN150P1T. First one is storage tank fullness percent versus time (Figure 6.15). Tank is placed at 12th node (with respect to drawing order) with a label “13” as seen in Figure 6.13 and Table 6.12. It is seen that water level at the end of the day is higher than water level at the beginning of the day, water levels do not stay still more than three hours, tank fill and empty cycles do not occur more than five times and tank is filled and emptied almost fully through the day. When these conditions are met and all constraints are satisfied. With the satisfaction of constraints, good operation of a water tank is performed for the defined network.

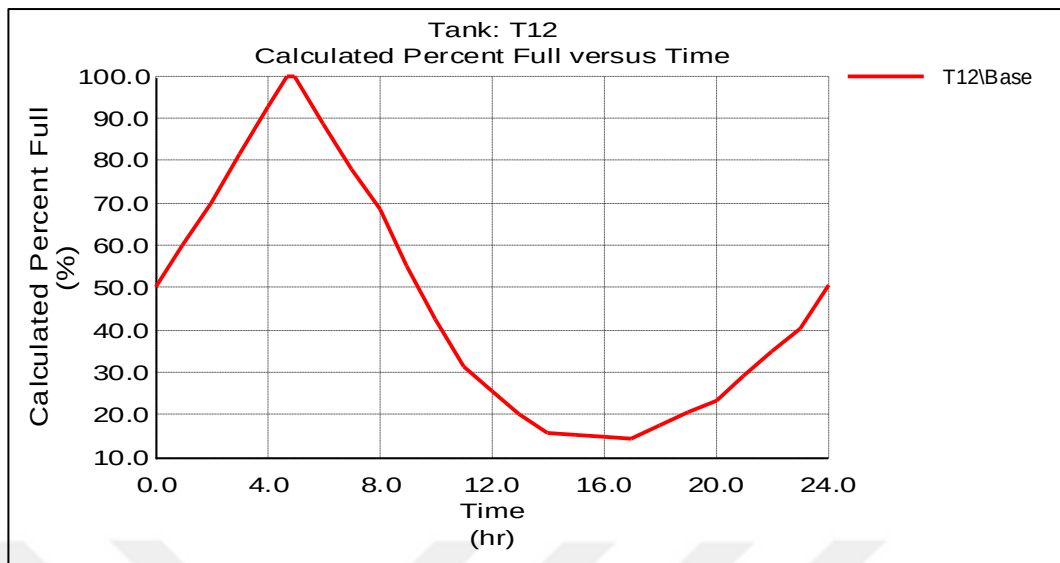


Figure 6.15. Tank Oscillations for Solution

In the second graph, nodal pressures versus time are plotted (Figure 6.16). This graph shows that pressures below 30 meters do not occur in the network through 24 hours meaning that all consumer nodes are supplied with enough pressurized water. This shows that minimum pressure constraint is satisfied.

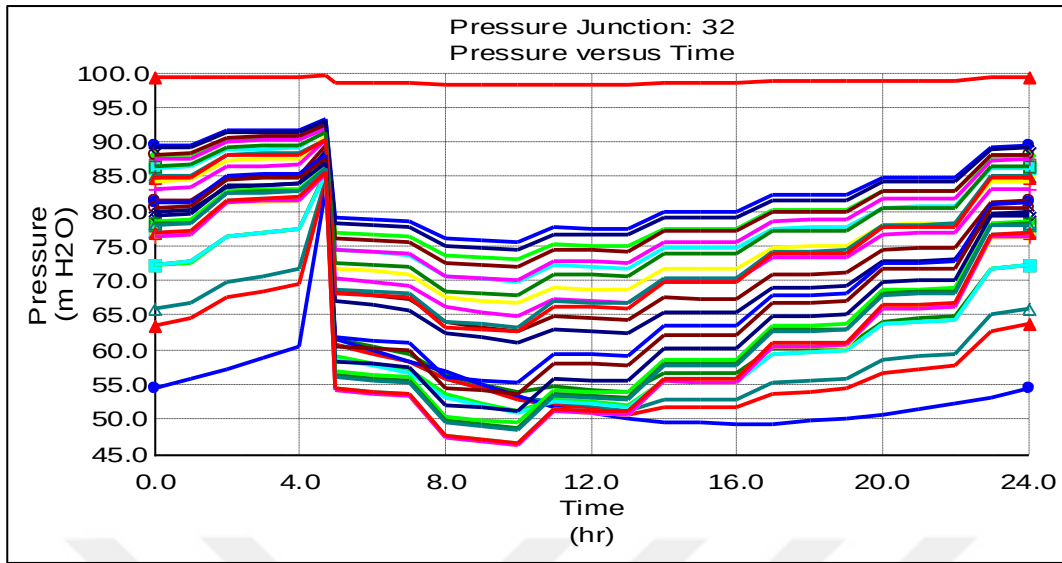


Figure 6.16. Pressure Variations for Solution

Tank oscillation and pressure variation graphs are provided in Appendix B for other solutions coded as HAN200P2T/75, HAN100P3T/23, HAN250P4T/250, and HAN100P5T/100. In these graphs it is again clear that all constraints are satisfied and good operations are performed regardless of tank counts.

A heat map of tank placements is given in Figure 6.17. In this map, a colour ramp from red to green is defined and darker green indicates most preferred tank sites whereas yellow and orange colours indicate a lesser preferred site. Red colour indicates a site that is not preferred. Accordingly, nodes 29, 31 and 13 are the most preferred tank sites, and they are followed by nearby nodes 32, 30, 28 and 26. Nodes like 22, 27 or 14 are less preferred but nodes nearest to source at node 1 are not preferred at all (like nodes 3, 4 or 20). A similar phenomenon to 2-loop case is seen here also. Tanks are placed as far as possible from the source nodes. So, they can provide nearby nodes water easier. Again like 2-loop case, most water demanding nodes (7, 18, 20, 23 and 6) are not widely preferred as a tank site even if they demand approximately 30 % of total consumption.

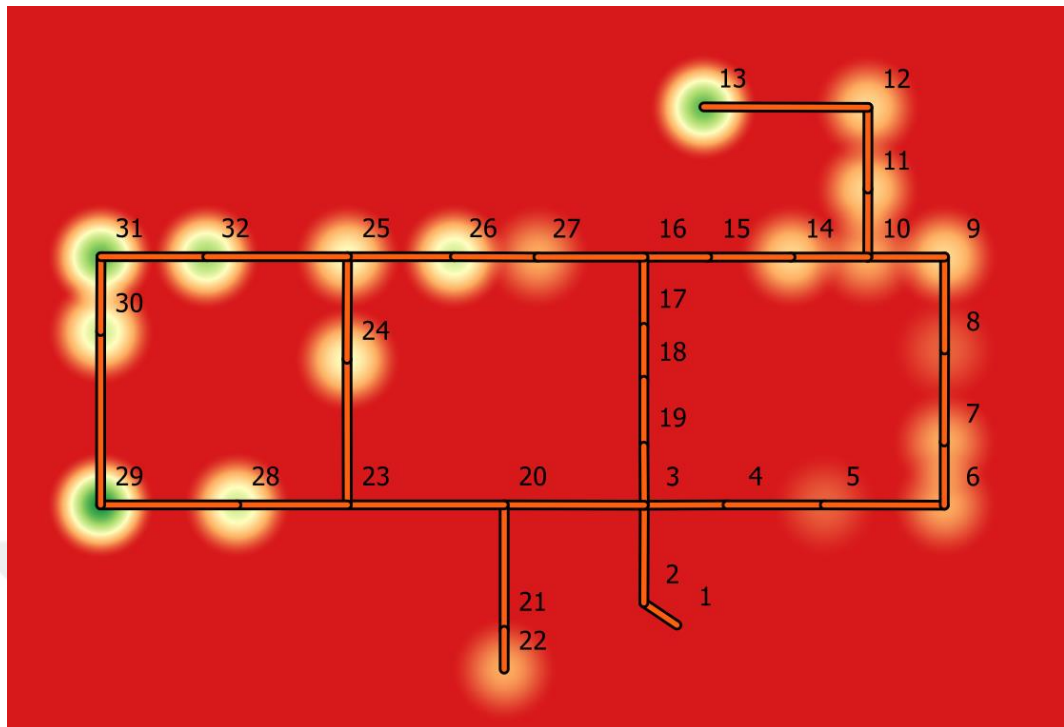


Figure 6.17. Heat Map of Tank Placements for Hanoi Network

In a final graph given in Figure 6.18, behaviour of members in the population through generations are presented. Thick lines represent number of feasible elements in an iteration, while thinner lines represent number of first front elements for that generation. All population sizes are plotted on the same graph.

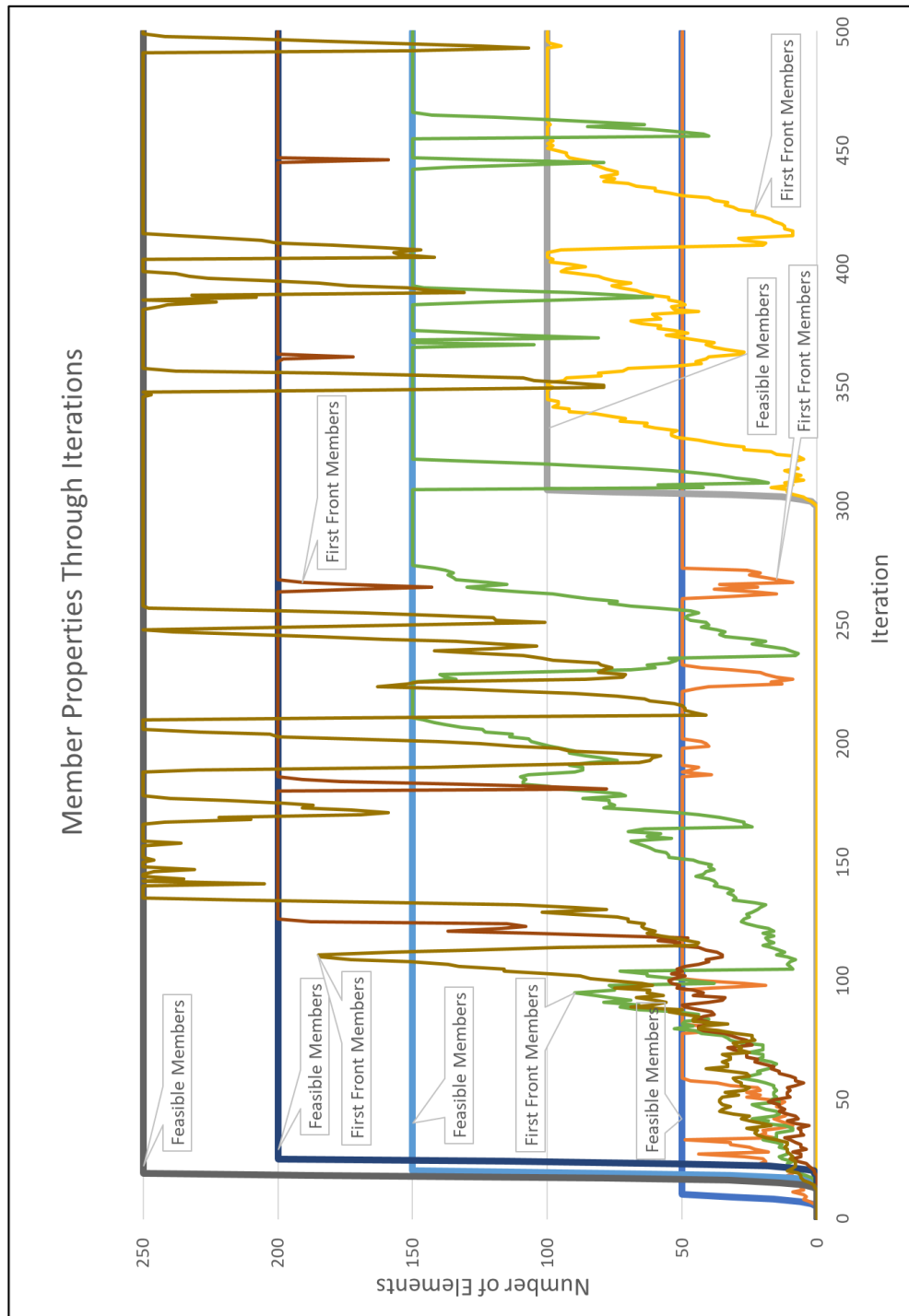


Figure 6.18. Member Properties for HAN1T

6.2.4. Comparison with Published Works

Two comparisons similar to the ones made in Section 6.1.4, are performed and their results are presented in Figure 6.19 and Figure 6.20.

In the first comparison (Figure 6.19), pareto curves obtained by pure pipe diameter optimization (with no tank inclusion) are plotted with best estimate pareto curve (Wang et al. 2015). In addition, two solutions from original Hanoi study are included (Fujiwara and Khang 1990, 1991). finally, several single-optimization results reported in Gençoğlu's study are included as singular points (Gençoğlu 2007).

While producing results for comparison, two trials with different diameter selection lists are used. The lists are compared in Table 6.13 taken from the actual results of the authors (Fujiwara and Khang 1991). First trial is denoted as "OT200P" and second one as "Original Diameter Choices". First list is composed of 14 commercial pipe diameters completed to 16 by using 1 mm increased diameters for 300 and 1000 mm with same costs. On the other hand, second list is composed of uncommercial pipe diameters like 381 mm. By this comparison the best estimate pareto curve is tried to be reached and with second list it is succeeded. From Figure 6.19, two trials can be compared with the best estimate pareto curve and the effect of using different diameters can be seen better. The conclusion is diameter selection lists highly affect the pareto curve, thus the strength of the network. Despite in Wang et.al's study does not give any data about decision variables; it is possible to say that there are definitely non-commercial diameters.

Moreover, it is seen that best estimate pareto curve covers an I_n range of approximately 0.2 and 0.35. The findings presented here cover a narrower but similar range. However, one result of the original Hanoi study and results reported by Gençoğlu are all at a lower cost region having an I_n range between 0.1 and 0.15. As these older studies all focused on minimizing costs, it is normal that they are in the low-cost region. Four of these reported studies (Cunha & Sousa 1999, Savic &

Walters 1997, Gençoğlu 2007 and Fujiwara & Khang 1990) can be added to the best estimate pareto front increasing the I_n coverage of the reported best estimate curve.

Table 6.13. Pipe Diameter Selection Lists for Hanoi Network

OT200P			Original Diameter Choices		
Diameter (in)	Diameter (mm)	Unit Cost (1/\$)	Diameter (in)	Diameter (mm)	Unit Cost (1/\$)
12	300	45.73	12	305	45.73
12	301	45.73	15	381	63.90
14	350	57.62	16	406	70.40
16	400	70.40	17	432	77.10
18	450	84.00	19	483	91.10
20	500	98.39	20	508	98.39
22	550	113.51	22	559	113.51
24	600	129.33	23	584	121.33
26	650	145.83	24	610	129.33
28	700	162.98	25	635	137.50
30	750	180.75	29	737	171.79
32	800	199.12	32	813	199.12
34	850	218.08	35	889	227.77
36	900	237.60	36	914	237.60
40	1000	278.28	38	965	257.67
40	1001	278.28	40	1016	278.28

Second comparison is done to see the constraint effects on Hanoi network. A similar comparison for 2-loop network is performed on Section 6.1.4. Similar findings like envelope of solutions varying between fully constrained and single constrained pareto curves are still observed in this case. The findings are presented in Figure 6.19.

Minimum pressure comparison is again performed between solutions stated by Gençoğlu and two end solutions from “Original Diameter Choices” pareto curve given in Figure 6.19. Two end solutions are referred as HAN0T200P/199 (high cost – high strength) and HAN0T200P/200 (low cost – low strength). Comparison is given in Table 6.14. Similar to 2-loop comparisons, multiobjective optimization provided a variety of results having minimum pressures between 30.05 m and 44.29 m. On the other hand, single objective optimization solutions were able to find low cost – low strength solutions having a minimum pressure around 30 meters.

Table 6.14. Minimum Pressure Comparison

Solution ID	Minimum Pressure (m)
HAN0T200P/199	30.05
HAN0T200P/200	44.29
Fujiwara & Khang (1990)	29.505
Fujiwara & Khang (1990) – 2	35.053
Gençoğlu (2007)	30.01
Abebe&Solomatine (1998)	30.23
Cunha&Sousa (1999)	29.66
Güç (2006)	30.06
Liong&Atiquzzaman (2004)	30.05
Neelakantan&Suribabu (2005)	30.01
Savic&Walters (1997)	29.73

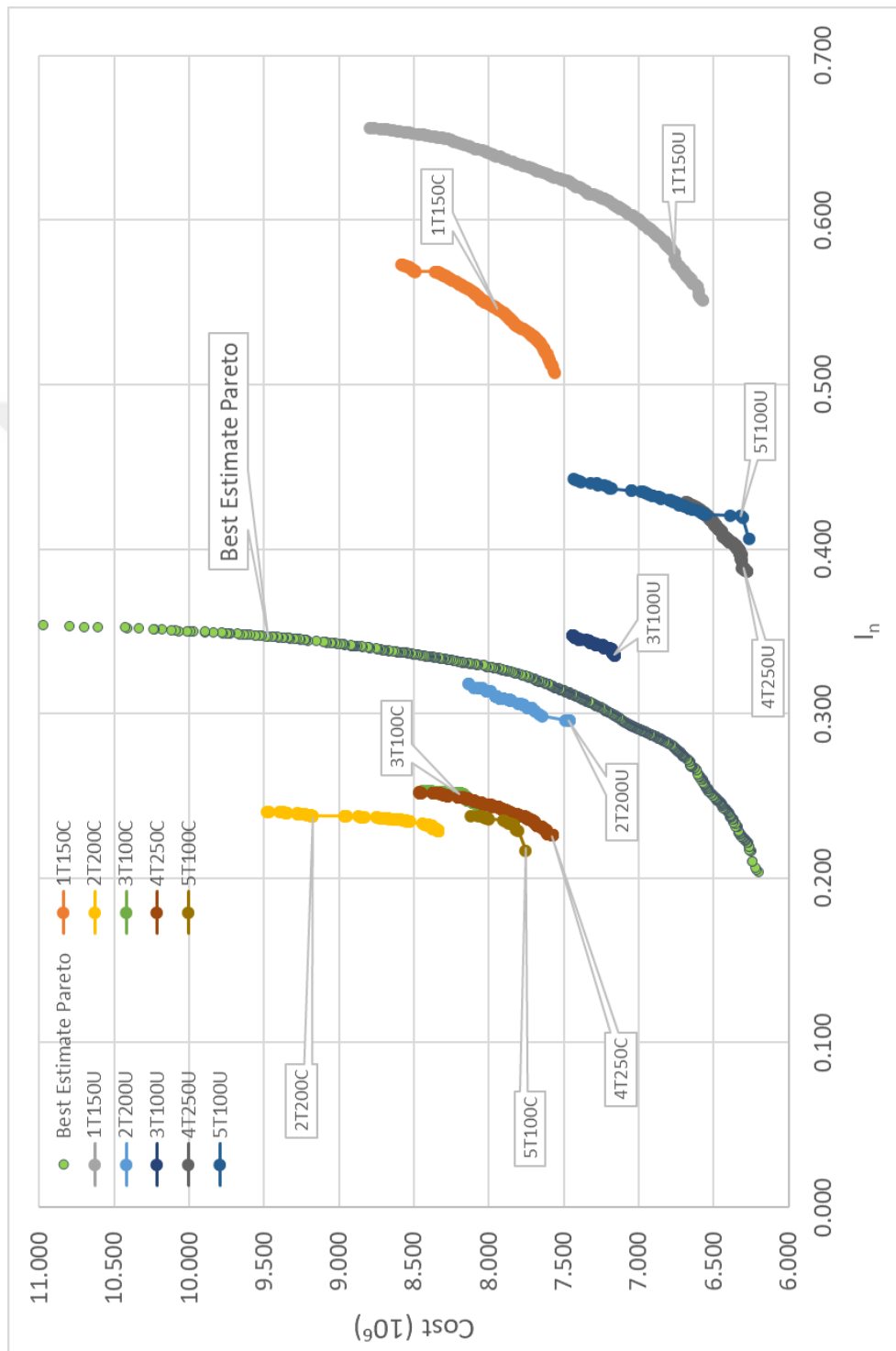


Figure 6.20. Comparison of Constraint Effects to the True Pareto Curve

6.3. Ankara N8.3 Network

6.3.1. Network Data

N8.3 system is a pressure zone located in north-western part of Ankara; with approximate coordinates “40.00° latitude” and “32.80° longitude”. It is a part of İvedik System which is the main water supply system of Ankara city. Water is transmitted from İvedik Treatment Plant to the north parts of the city by pumps through stations P1, P2, P12 and P23 respectively (Figure 6.21). The establishment of İvedik System dates back to 1969 when the first master plan is performed (Schilling and Porter 1990).

N8.3 system lays on a perturbed topography. They are Çiğdemtepe, Şehit Kubilay, Sancaktepe and Yayla from west to east. The hill on Yayla Neighbourhood is separated from other hills with Yozgat Boulevard. Present facilities and neighbourhood borders are shown in Figure 6.22. Actually, N8.3 is a pumped network. It has one pump station (P23) and one storage tank (T53). P23 pumps feed the system on west hills with a Ø500 mm pipe and Yayla hill with a Ø300 mm pipe. Storage tank named T53 is at the west end of the system which is fed by P23 pumps also. The layout of the network and topography of the service area is given in Figure 6.23.

The network is modelled by 712 consumer junctions and 803 pipes. Water source has its base elevation at 1156 meters and the service area is between 1051 and 1142 meters. It is obvious that, in a system with a water source at 1156 meters and with consumers at 1142 meters, it is not possible to maintain operating pressures of 30 meters. The situation of consumers at high elevations near the source is a result of unplanned shanty settlements. As these shanty houses are mostly single-story structures, operating water pressures around 15 meters may be satisfactory. Aforementioned high settlements that are exposed to low pressures are only located near the water source which is a minor part when whole system is concerned. For

the explained reasons, minimum pressure constraint needs to be revised for N8.3 system.

Actual water consumption measurements are performed in 2011 in the network by dividing it into 6 district metered areas (DMA) (Şendil 2013). Measurement dates are as follows:

- 7 September 2011 for West Çiğdemtepe
- 15 September 2011 for East Çiğdemtepe
- 19 September 2011 for North Sancaktepe
- 22 September 2011 for South Sancaktepe and Şehit Kubilay
- 28 September 2011 for Yayla

These six measurements are separately applied to the relevant network parts. Separate and combined demand patterns are given in Figure 6.24. When the consumption patterns are investigated, it is seen that there is only a distinct peak hour in the middle of the day (around 12:00). Moreover, between 03:00 and 05:00 it is seen that lowest consumption occurs with approximately 45% of average consumption. These low consumption values can be considered as water leaks occurring in the area.

With the provided information, the problem is stated as follows: The service area of present N8.3 system is defined as a new settlement area. Development plans defining streets (where new infrastructure will be places) are present. There is a water source with enough capacity at the site where T53 stands. Population and water use estimates are calculated as in Figure 6.24. A design for water distribution network with an elevated storage is needed. The designer seeks multiple solutions that consider costs and strength of the system. Therefore, the methodology developed in this study is used to perform a multiobjective optimization to suit the designer's needs.

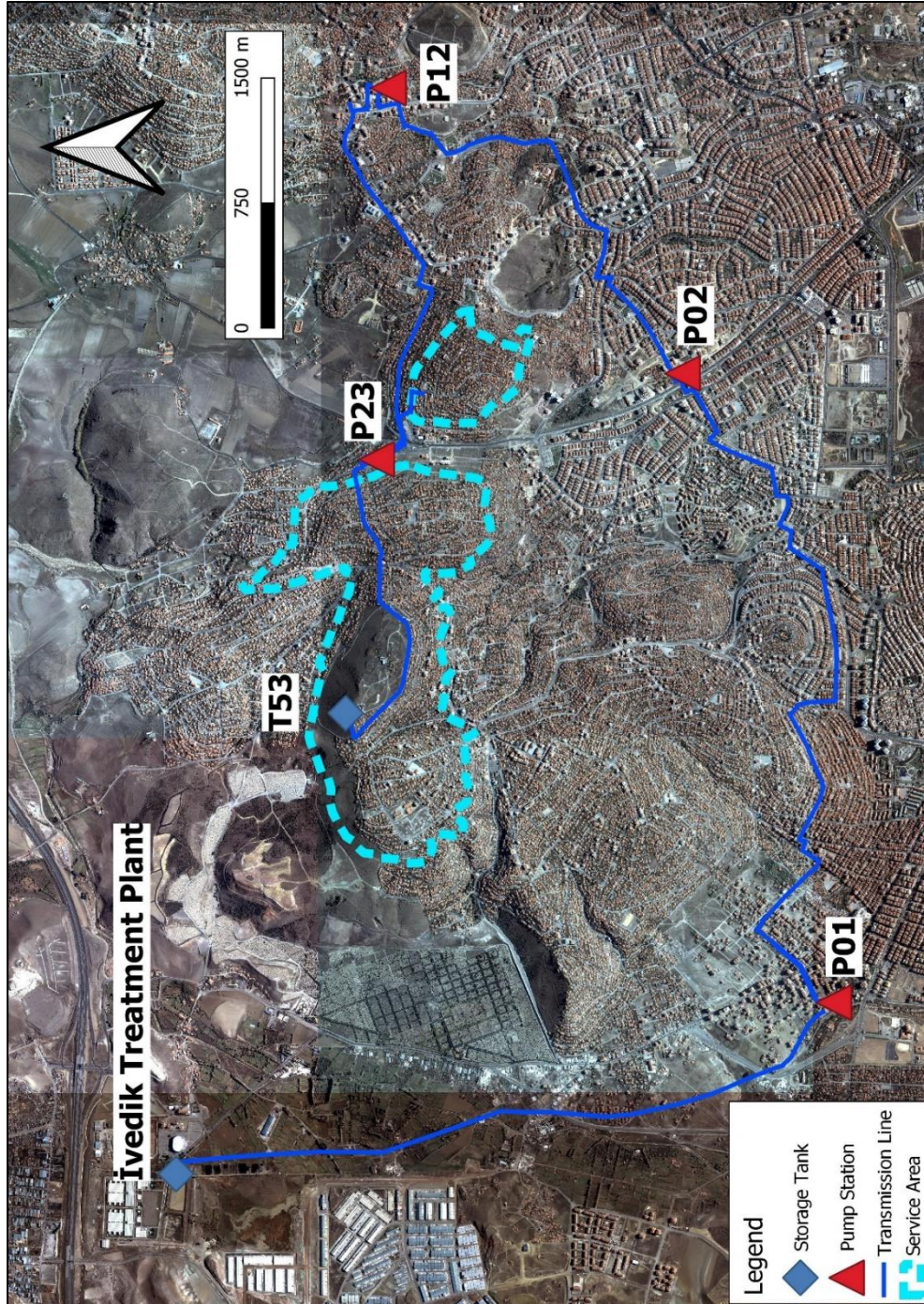


Figure 6.21. N8.3 System as a Part of İvedik System

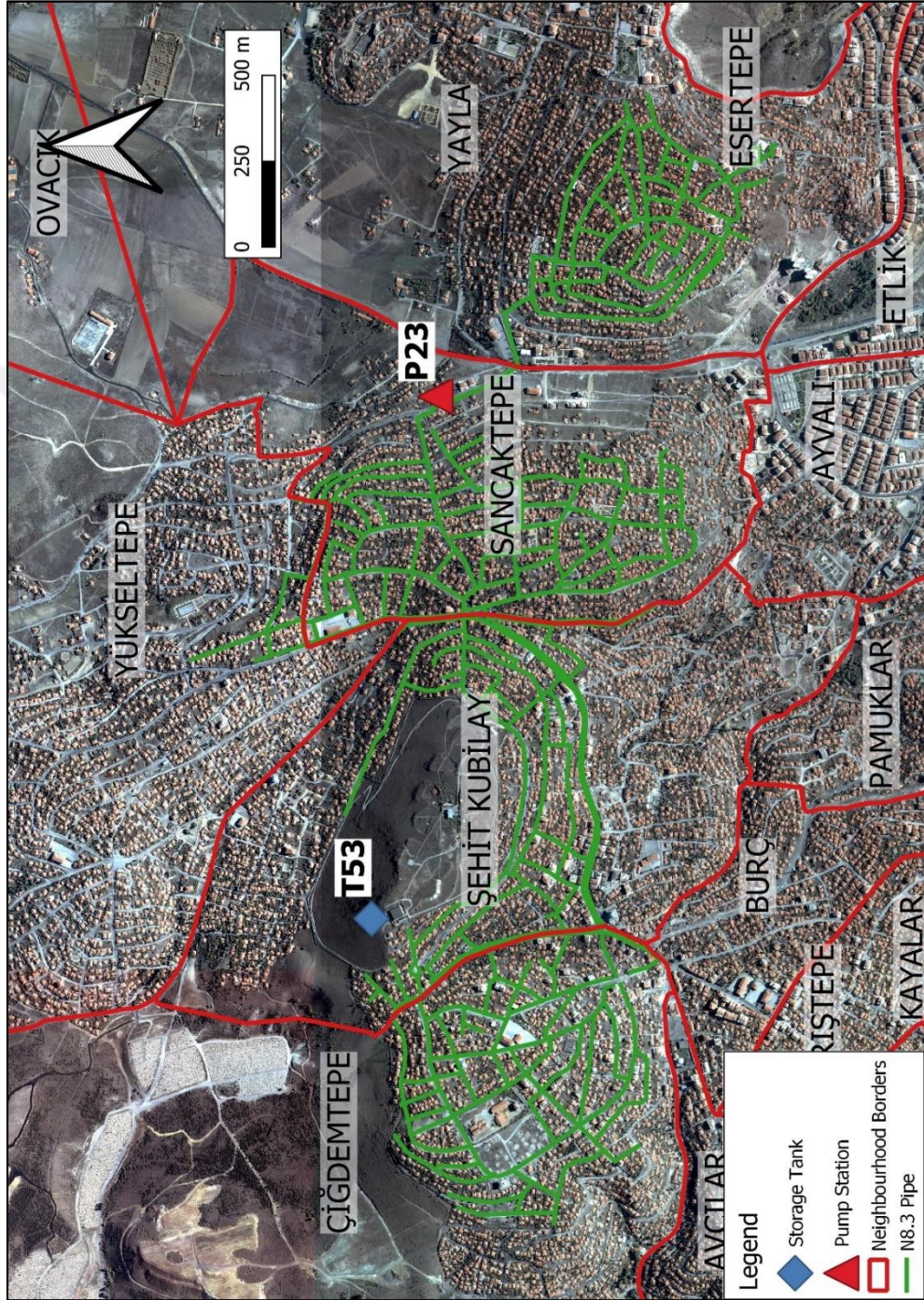


Figure 6.22. Present Facilities of N8.3 System

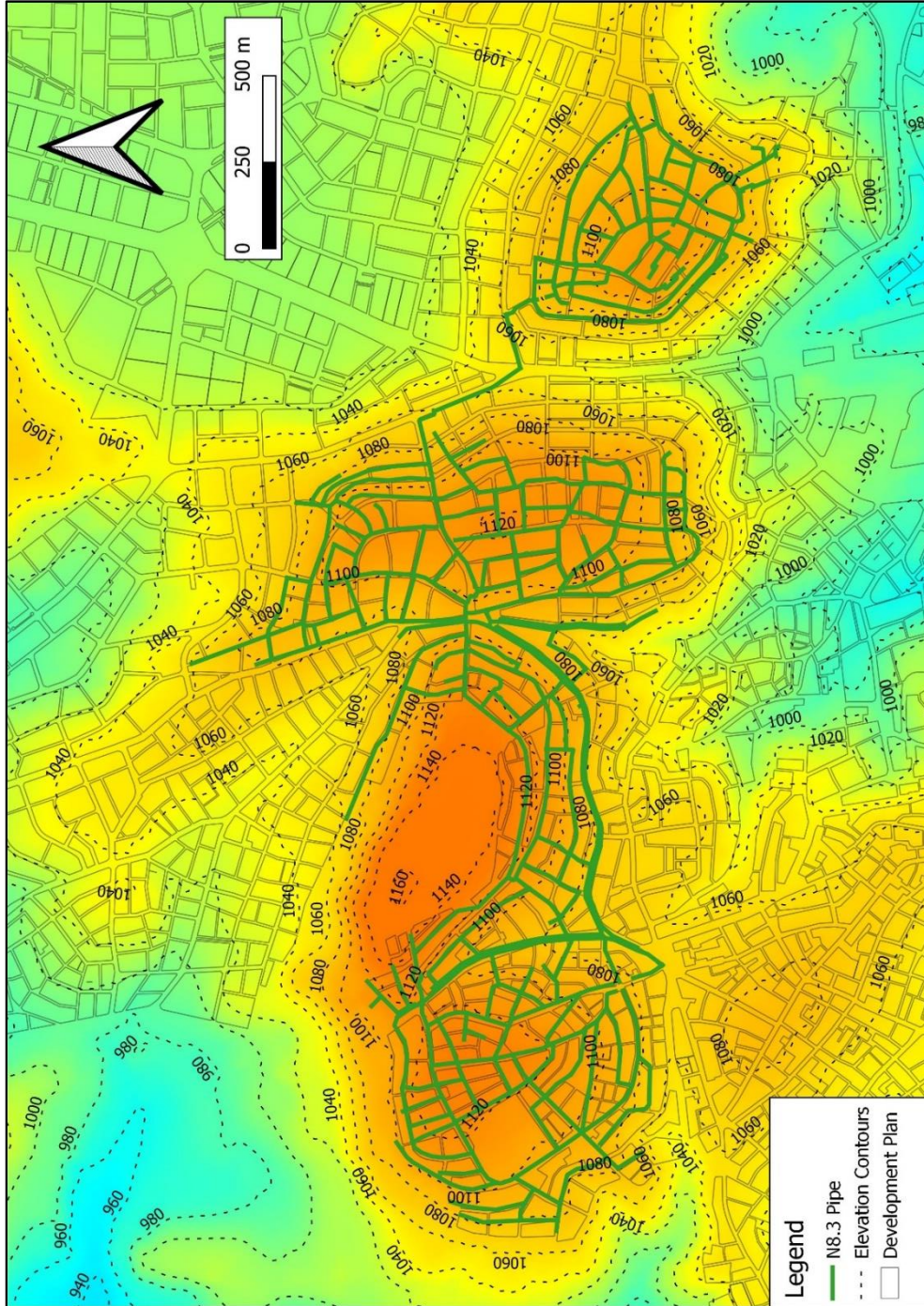


Figure 6.23. Topography of N8.3 Study Area

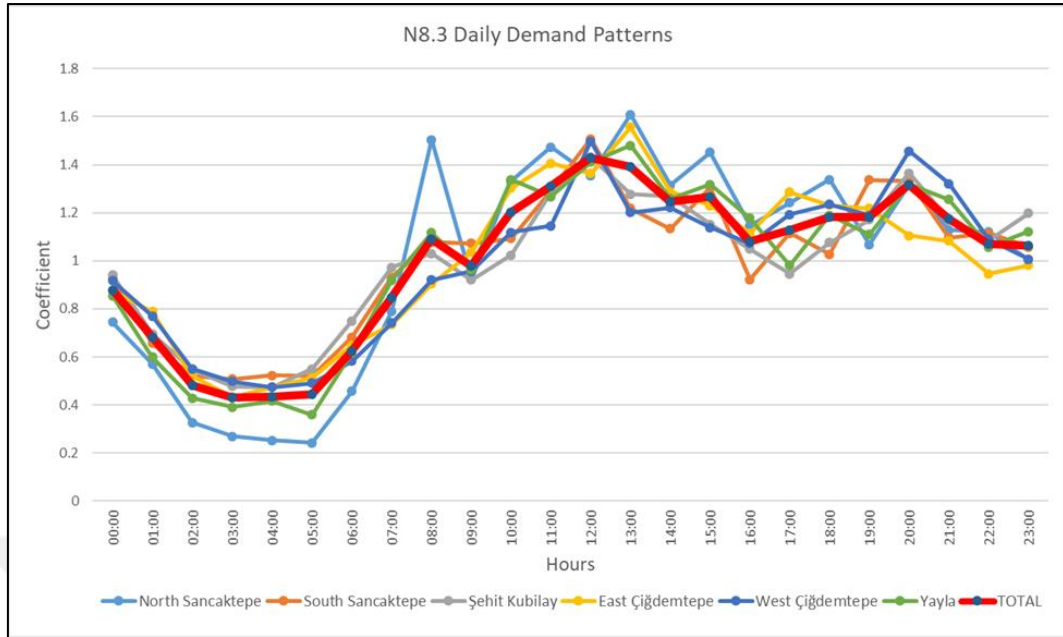


Figure 6.24. Daily Demand Pattern of N8.3 Network

For pipe costs, 2018 ductile iron pipe values from Table 4.8 are used. They are interpolated and extrapolated to fill missing unit costs. For tank costs fitted curve given in Figure 4.13 is used.

With the formula given in Section 4.3, size of search space of N8.3 network for 1 to 2 tanks are given below:

- 1 Tank / ($n_j=712$, $n_p=803$, $n_t=1$, $n_{pd}=16$) – $SS = 3.78 \times 10^{977}$
- 2 Tanks / ($n_j=712$, $n_p=803$, $n_t=2$, $n_{pd}=16$) – $SS = 6.88 \times 10^{982}$

Daily water demand of N8.3 network is 4380 m^3 , and balancing flow amount is 185 m^3 .

6.3.2. Implementation

Similar implementation details given in Section 6.1.2 are applied. The main difference is OCVNEGJ constraint is modified so that only negative pressures are prevented.

N8.3 network is optimized by trying two tank configurations. This makes a total number of solutions of 10. A normalization of 10^7 to cost values are applied.

In this study, pump station and transmission line ($\varnothing 500\text{mm}$) between pump and tank is removed. All present network pipes in the system are connected to each other without a transmission line. Additionally, storage tank is replaced by a reservoir. The problem therefore turned out to a gravity network optimization which tries to find available elevated tank sites inside the network. (Figure 6.22). Nodal elevations are obtained from digital elevation model of the area (Figure 6.23).

Nodal demands are distributed equally within the six relevant DMAs.

6.3.3. Results

The results are summarized in Table 6.15. Also, summary statistics for every run is given in Table 6.16. Following the tables, Figure 6.25 for pareto front and hypervolume is provided for the selected population size of one tank solutions. Rest of the similar figures are given in Appendix C.

The solutions are coded as N83XXXPYT in which XXX denotes the population count and Y indicates tank count. For example, N8350P1T means the trial performed with 50 population members and 1 tank.

In both tank count trials; 250 population solutions generated the largest hypervolume therefore they are defined as optimum population sizes. Also, with two tank counts, solutions did not reach the stopping condition until the maximum iteration count of 2000. Moreover, 250 population 2 tank solutions could not generate 250 distinct first

front elements. However, selected two solutions are the cheapest and strongest (excluding 2L100P2T/19) alternatives within the 10 preferred solutions.

Table 6.15. Summary Results of N8.3 Network

Tank Count	Population Size	HV	Selected Member	I _n	Cost (TL)	Average Pipe Cost (TL)	Average Pipe Diameter (mm)	Average Tank Cost (TL)	Average Tank Volume (m ³)
1	50	0.178	3	0.6412	26817041	26886820	417	573090	719
	100	0.175	90	0.5338	22432230	23550202	384	472494	484
	150	0.309	61	0.6611	24749725	24776298	399	566803	704
	200	0.492	127	0.7214	23525926	23590267	397	471919	481
	250	0.652	250	0.7987	20113405	20312755	349	346074	193
2	50	0.187	50	0.5892	26757991	26079512	412	844900	732
	100	0.324	19	0.6886	23435792	22704810	381	771528	564
	150	0.367	139	0.6037	23372771	24562473	382	898668	856
	200	0.419	19	0.5668	22060694	21229980	361	917318	901
	250	0.938	39	0.6790	20852760	20189164	352	914675	893

Table 6.16. Summary Statistics of N8.3 Network

Tank Count	Population Size	Total Iteration	Total Run Time (s)	Average Iteration Time (s)	Final Percent of F1 Elements
1	50	529	4655	8.80	100.0
	100	689	12837	18.63	100.0
	150	527	15128	28.71	100.0
	200	500	18874	37.75	100.0
	250	1633	86929	53.23	100.0
2	50	1035	9363	9.05	100.0
	100	2000	48382	24.19	25.0
	150	2000	73203	36.60	100.0
	200	2000	94620	47.31	43.5
	250	2000	115891	57.95	19.2

In Figure 6.25, final pareto curve for 250 population 1 tank solutions is given. Remaining 9 pareto curves are given in Appendix C.

The search for optimum solutions can be traced from Figure 6.25. Within many solutions having I_n as low as 0.25 and costs as high as 32 million TL, final pareto optimal front is composed of solutions having I_n around 0.8 and costs around 20.5 million TL.

When the pareto curve in Figure 6.25 is examined, it is seen that the curve is rather narrow in both x and y axes. This is generally observed in other results of this case. The reason for this is an optimum tank is detected by the algorithm, and remaining elements are formed by using that tank properties with varying in pipe diameters only. Only pipe diameter variations with a good tank scheme creates a narrow list of pareto members.

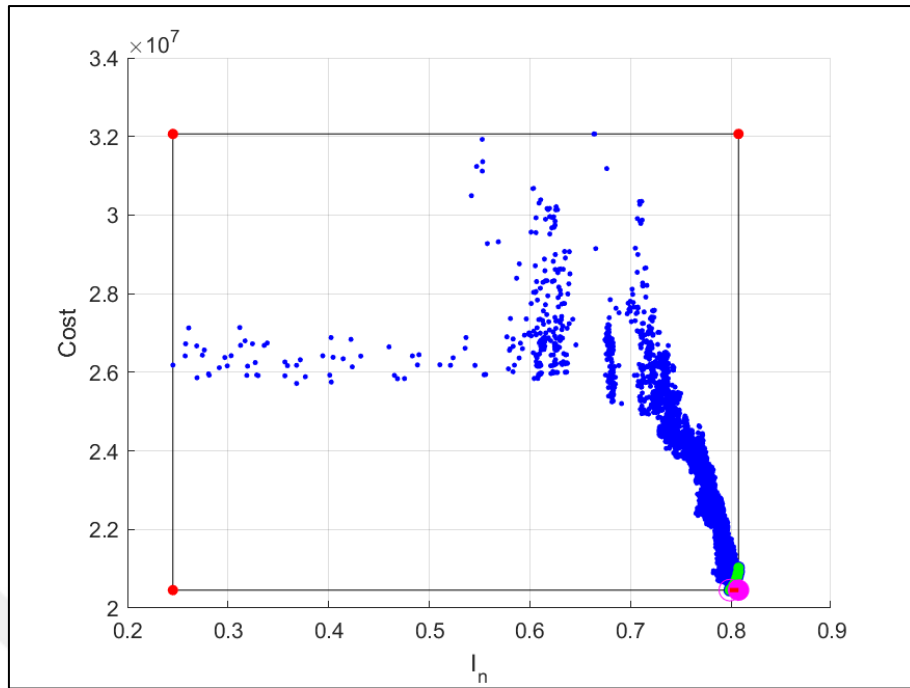


Figure 6.25. Pareto Curve for N83250P1T

Two designs are selected as the preferred solution, namely:

- Solution 250 from 250 population 1 tank configuration
 - $I_n = 0.7987$ Cost= 20457093 TL (Pipe cost = 20113404 TL)
- Solution 39 from 250 population 2 tank configuration
 - $I_n = 0.6790$ Cost= 20852760 TL (Pipe cost = 19938567 TL)

Tank design values (decision variables) for the preferred solutions are given in Table 6.17.

Table 6.17. Tank Design Values for the Preferred Solutions

Solution ID	Tank Sites	Tank Heights (m)	Tank Diameters (m)	Tank Elevations (m)	Total Tank Volumes (m ³)	Tank Cost (TL)
N83250P1T/250	542	3.00	8.92	38.29	187	343688
N83250P2T/39	96	3.00	12.41	48.37	891	914192
	336	4.23	12.61	49.67		

From Figure 6.26, it is seen that tank constraints are satisfied and from Figure 6.27 pressure oscillations for junctions can be observed. Although it is hard to observe junction by junction results within 712 different curves, this plot gives an idea about the minimum and maximum observed operating pressures occurring. As minimum pressure, there is only one junction having approximate 15 meters and one junction between 15 and 20 meters of operating pressure. Also, it is possible to observe the maximum pressures in the network do not exceed 110 meters. However, to observe pressures better, a colour coding is applied on the network depending on the resulting pressures. To observe minimum pressures maximum demand hour is selected as 12:00. Results for this specific hour are colorized with four classes (Figure 6.28). Maximum pressures are observed at 03:00. Colour coding maps and results for two tank designs are given in Appendix C.

Another colour coding map for pipe diameters of N83250P1T/250 is presented in Figure 6.29. Similar map for N83250P2T/39 is given in Appendix C. In these maps pipe diameter choices can be seen visually. It can be concluded that approximately 70 % of the pipes are chosen smaller than 500 mm. This is a natural outcome of cost minimization. However, placing tanks has an effect in choosing smaller diameters. Therefore, tanks can be said to help decrease diameters and maintain the lost strength while decreasing diameters. This can be supported from Table 6.15, where average diameter for 1 tank solutions are 389.2 mm but 377.6 mm for 2 tank solutions. An

overall 11.6 mm decrease in pipes can be achieved by allowing another tank in the network.

In Figure 6.30, 18 different tank sites for 10 solutions obtained from selected results are plotted. Similar to previous sections a trend to place more storage tanks far from the source is observed. However, there is a considerable number of storage tanks near to the source. This is interpreted as the methodology placed storage tanks to the highest available points and achieved to find a feasible spot nearby the source.

In the final figure, Figure 6.31, member properties through generations is given. It is seen that 250-population run has gone under many successive crossover and mutation operations that changed the first front solutions greatly. However, at the end of 1633 iterations, it managed to end with full first front elements resulting a solution set of 250 distinct members. Appendix C contains the similar figure for two tank solutions.

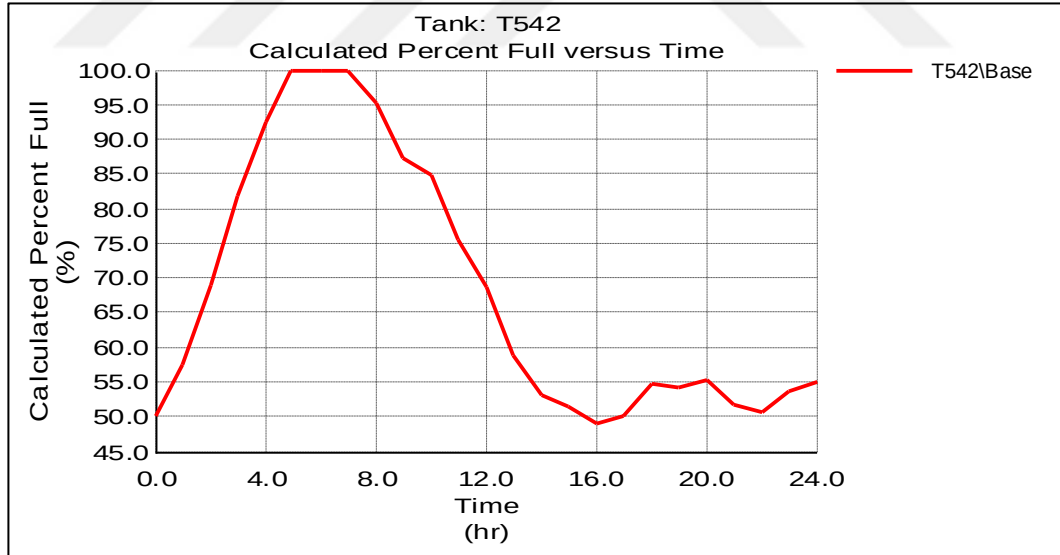


Figure 6.26. Tank Oscillations for Solution N83250P1T/250

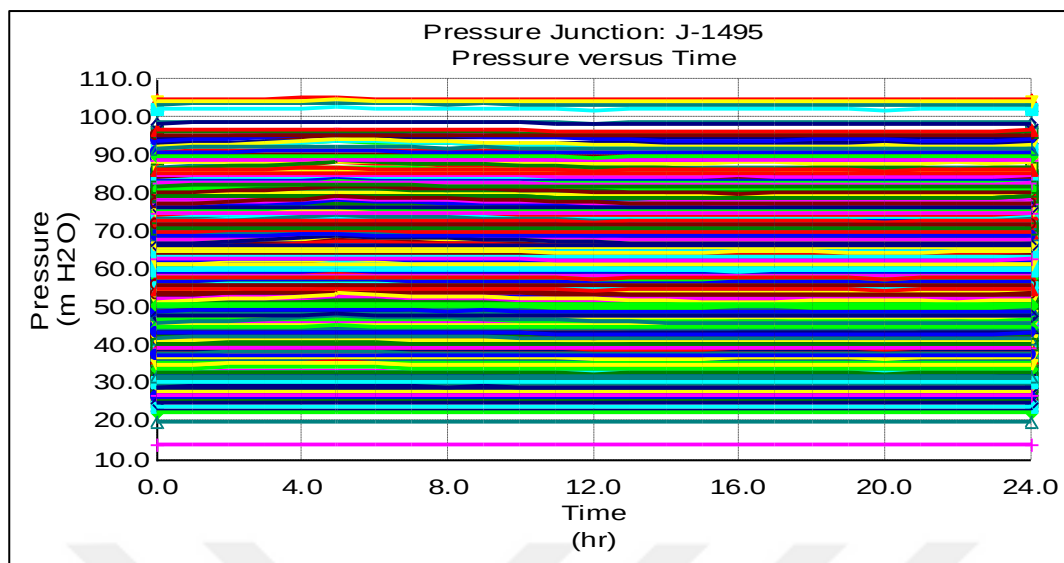


Figure 6.27. Pressure Variations for Solution N83250P1T/250

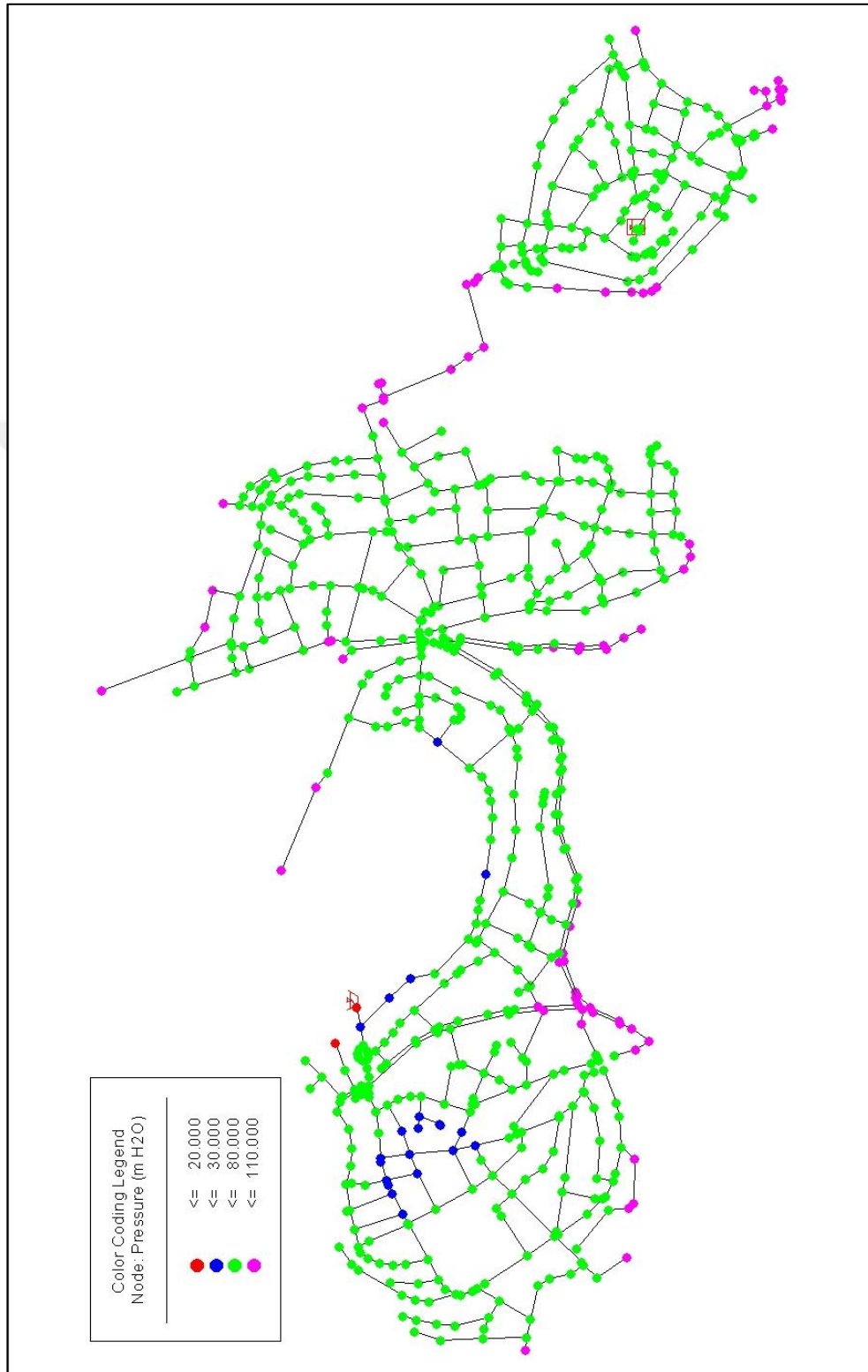


Figure 6.28. Pressure Distribution for N83250P1T/250 at 12:00

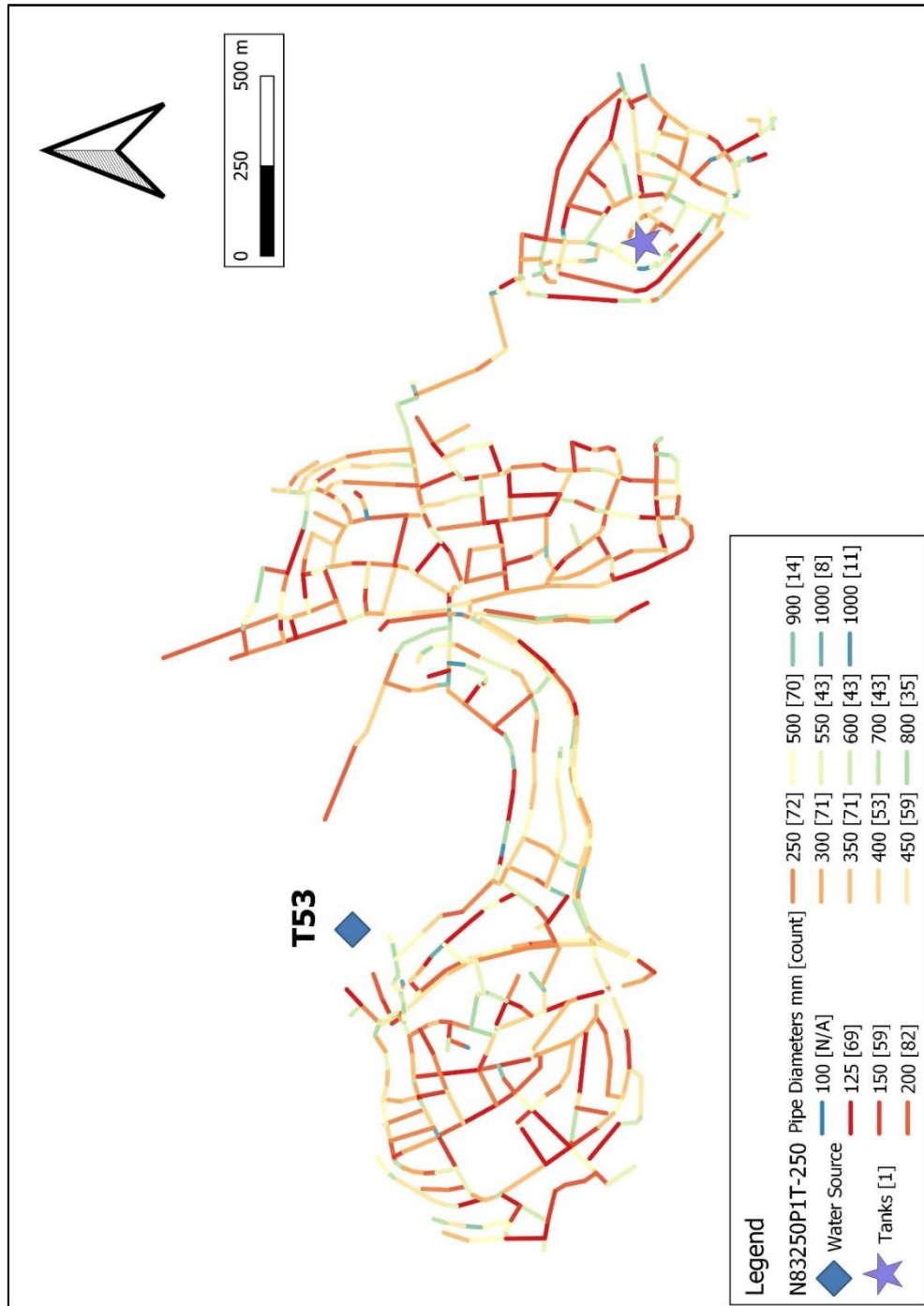


Figure 6.29. Pipe Diameters for N83250P1T/250

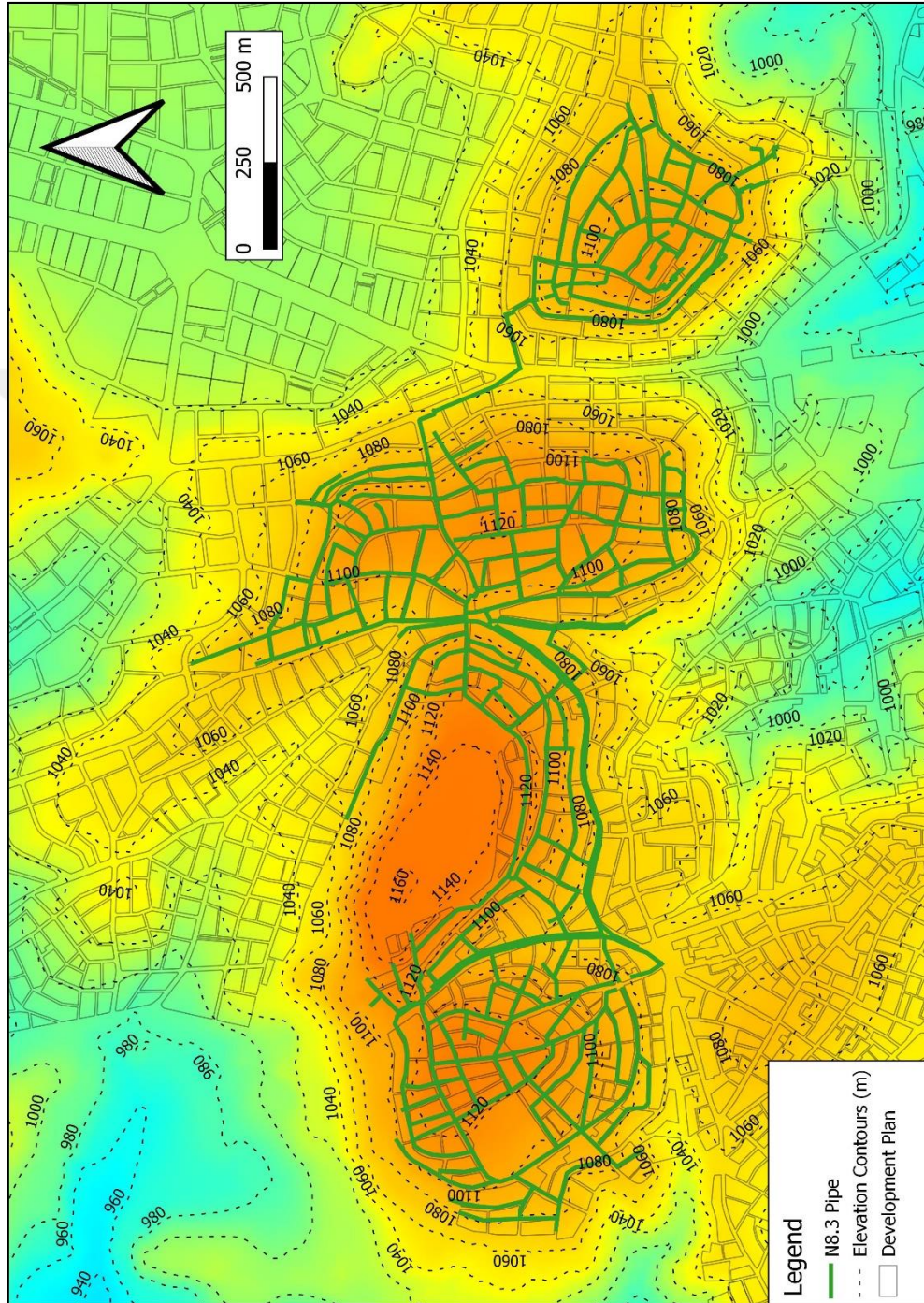


Figure 6.30. Tank Sites for 10 Obtained Solutions

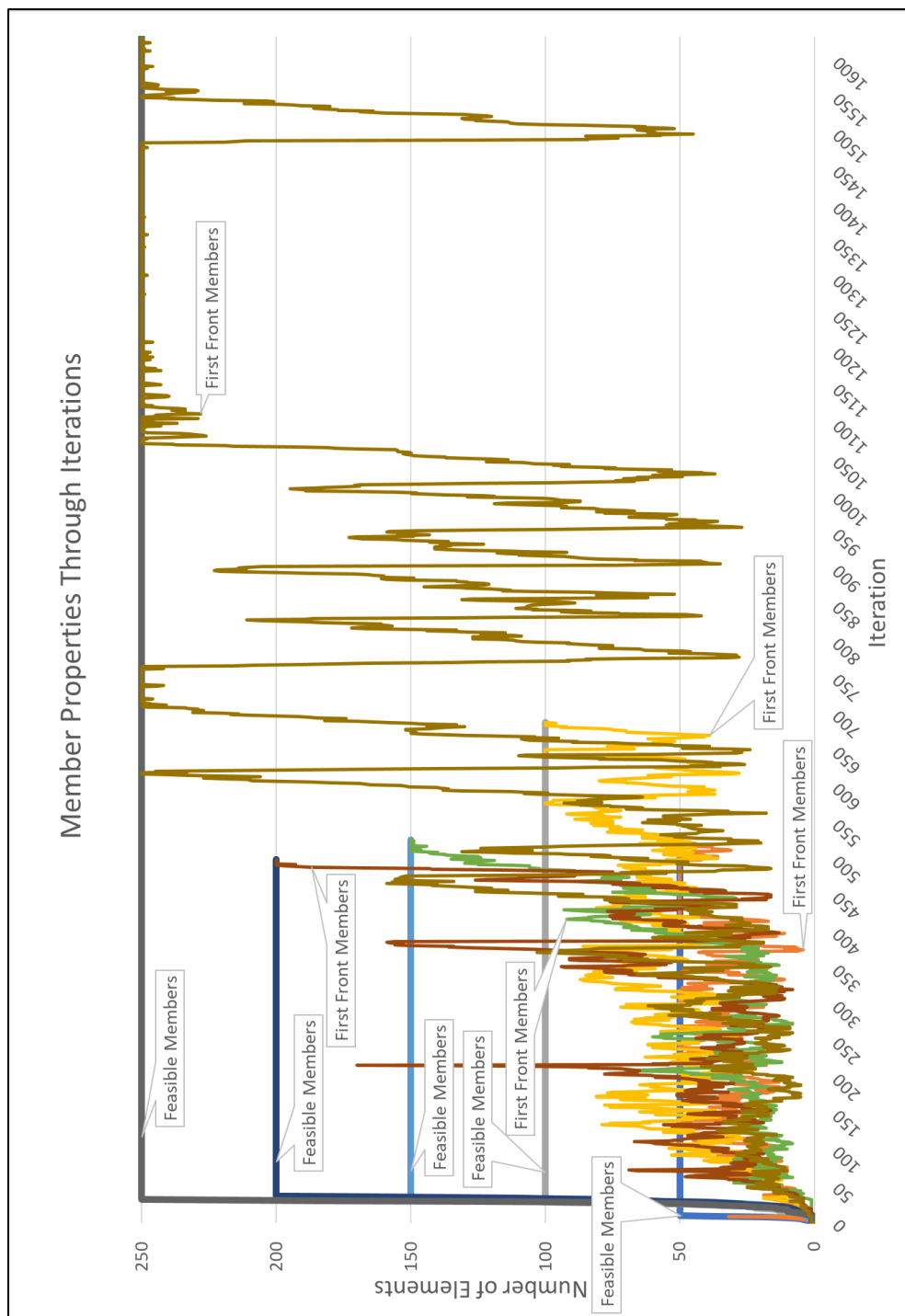


Figure 6.31. Member Properties for N831T

CHAPTER 7

CONCLUSIONS

7.1. Discussions

This study aims to generate cost effective and strong water distribution network designs in a multiobjective optimization manner. The study aims to minimize both total construction costs of the designed system and maximize the strength of the system in terms of network resilience. As water distribution network elements, network pipes and water storage tanks are considered. The methodology is capable of optimizing gravity fed networks but it can be extended to cover pumped networks as well.

During testing the methodology on three different networks, it is clearly seen that constraints take an important role in the process. In Section 6.1.4, it is shown that by adding or cancelling out constraints, final pareto curve shifts significantly. However, adding too many constraints make it harder to find solutions, and cancelling out too many of them result in hydraulically inoperable systems (like storage tanks having no effect).

Definition of the decision variable space (i.e. allowed pipe diameters, allowed storage tank dimensions) has a big effect on the final pareto curves; as it was shown in Hanoi results. What can be recommended at this point is to create a decision space for every variable as realistic as possible to achieve realistic results. Defining a big interval for minimum and maximum values (although the interval contains the realistic values inside) may lead to unrealistic pareto curves.

The effect of storage tanks to the strength of the design is less when compared to pipe diameters. The reason for it is that, many constraints are defined to achieve an operable tank that both stores some water volume and helps peak hour consumptions.

It is sure that by adding storage tanks pipe diameters in the network balancing volume get smaller and it directly effects the cost of the network. This is a good practice when it is considered that around 90 % of construction costs are pipe costs.

Developed methodology has the ability to satisfy the designer's needs. Making careful formulations of any design requirement as constraints should steer the search to find optimal solutions. Unless very strict constraints are not defined, the methodology will be able to find solutions. It is a matter of making several trials to fine tune constraint formulations to satisfy designer's needs and to be able to find feasible solutions.

When the results of the case studies are investigated, it is noted that pipe diameter uniformity coefficient present in the network resilience formulation did not provide precisely uniform diameters in the network. Instead, it provides relatively uniform diameters. This can be overcome by grouping pipes and making pipe change operations based on the groups. Rather than applying such a grouping, overall pipe diameter uniformity for the network can be included as a third objective for maximization. Some similar studies achieved good results with such an approach (Creaco et al. 2016a).

Another important outcome observed from the results is that, total storage tank volumes are much smaller than figures recommended in various technical specifications. For example, for N8.3 2013 network technical specification from Table 2.1 recommends a storage tank volume between 1934 m³ and 2798 m³ however, the methodology generates storages not more than 900 m³. This is because, defined constraints direct solutions to select storages to have volumes approximately equal to daily balancing flows. Storage tank volumes are generated to have just enough volumes to satisfy entering and leaving volume requirements defined as constraints. If storage tank volumes are required to strictly satisfy technical specifications, then constraints defined in the methodology should be redefined to allow some fire and emergency water volume stay still in the tank throughout the

daily simulation. This may lead to water quality problems because of aging water volumes inside tanks.

A separate water quality analysis is not performed in this study, but good water quality is aimed by introducing constraints. With the help of constraints CVEV, CVLV and CVCON, water quality is tried to be preserved by eliminating water aging.

An outcome is the storage tank count limit reached during the trials. This concept can be explained by investigating 2-loop and Hanoi network results. In these results, storage tank counts are increased one by one and pareto optimal results are reached. However, no or very little number of feasible solutions could be observed after some tank counts in these networks. 3 tanks for 2-loop and 5 tank counts for Hanoi network, stand out for case specific tank count limits for these two systems. More tanks cannot be placed in these systems to satisfy defined conditions. Although by increasing tank counts, solutions spaces are also increased accordingly, but it becomes harder to find feasible and pareto optimal solutions in these increased search spaces. This indicates a contraction of the feasible region. Therefore, tank counts are also limited computationally in this perspective.

More issues are observed when more than one tank is optimized. It is seen that when different tank sites are found within first front feasible elements, they begin searching the space from two different tracks resulting a bigger diversity thus a bigger hypervolume. Relevant NSGA-II mechanisms should be studied carefully and precisely to be able to find multiple tank sites in a run. Examples of this phenomena can be seen in results of N83100P1T and N83150P2T in Appendix C.

It is observed that the developed methodology first focuses on finding feasible sites for tanks and tries to preserve these sites as much as possible. After finding feasible sites for tanks, pipe diameter optimization is more focused on. Although necessary tools for tank site change are provided in the methodology, relatively less changes are observed in tank sites throughout the process. This behaviour is more observed in N8.3 trials. With maximum 40-meter elevated tank alternatives, topography

effects are reduced, therefore it became possible to place tanks at lower elevations. These initial found feasible tanks sites are mostly preserved during the optimization process with pipe diameter and tank size optimizations going on. However, if a better tank site is found and members having old tank sites are moved into higher fronts, a big change in pipe diameters occur starting from pipes near the new tank.

Although tank dimensions are explicitly optimized, they may lead to inapplicable storage tank dimension ratios. Preferred designs are investigated if they violate AWWA M42 rules and it is concluded that they do not. However it should be noted that an assessment concerning structural stability is not covered in the scope of this study.

There is a considerable amount of randomness within the nature of evolutionary optimization. Therefore, it may not be possible to achieve good results in each and every run (Mitchell 1999). Results depend on random initial seeds created as initial populations. So, diversity may be lost if initial random seeds fail to find a feasible element within a short period of search. In such a case, the iterations should be stopped and restarted with another random seeds. This is the case applied in this study.

7.2. Suggested Further Studies

There are two types of hydraulic simulation solving techniques. They are, demand driven (node-head analysis) and pressure driven (node-flow analysis) approaches. Up to this point, all hydraulic simulations presented are based on demand driven approach. Also, all design requirements given in technical specifications are based on demand driven approaches. In this approach, it is assumed that all demands are fully satisfied and corresponding hydraulic grades are calculated accordingly. However, in reality that is not the case. There may be times that demands in some nodes cannot be fully satisfied, meaning that less amount of water could be served to those nodes than what is demanded. In such cases pressure driven approaches are

considered. At this point demand satisfaction ratio is introduced computing what percentage of the actual water demanded is met. Demand satisfaction ratio can be included as an additional objective or it can be formulated as a constraint. By doing so, initial population manipulation explained in Section 5.3.2 and constraints like CVPRES can easily be eliminated. Similar studies by Creaco et.al proves these statements (Creaco et al. 2016b).

This study started with the idea of including a transmission line between water source and storage tanks. This transmission line is aimed to be optimized with same rules as other network pipes. As the route of the transmission line is unknown at the initial phase, a shortest path algorithm using present network routes is applied. In addition, possible tank placement sites are extended to territories without a development plan (sites above 1140 m. contours in Figure 6.23). To make this idea possible, a 20m. x 20m. grid is defined in the empty territory, and corner points of this grid is used as a candidate tank site, which also effects the shortest path of transmission line between water source and storage tank. However due to coding difficulties raised from having multiple transmission lines for multiple storage tanks, this idea is left for future enhancements. With these future additions, more realistic designs will be possible which also allowing the designer to choose whether transmission lines should be included in the designs or not.

Pipe diameter uniformity is thought as an additional objective in Section 7.1. With the inclusion of demand satisfaction ratio, NSGA-III or other many-objective optimization methods can be considered. Other possible objectives that can be included may be minimizing water age, minimizing storage and pipe costs separately and maximizing residual chlorine concentration. As stated earlier, many-objective optimization methods are proposed to solve four or more objectives, and these methods may be useful if more objectives are introduced.

Water quality is not studied thoroughly in this study. Only some storage tank limitations are placed to prevent too big tank volume selections which ensures good water quality to some extent. In further studies water aging in the tanks can be

considered as a minimization objective. EPANET solver is capable of making water quality analyses. Additionally, different tank mixing models and tank inlet / outlet connection details can be studied.

Unfortunately, the computation speed of the EPANET EPS solver is not at the desired level. With upcoming versions of EPANET code or by using more powerful computers, better computation performance will be obtained. With increased computation performance it will be possible to run bigger populations for longer generations. This will allow better and more precise estimation of the pareto optimal curves. From another point of view, it will be possible to make more runs instead of a single run for a network configuration. Making numerous computations with different initial population seeds will allow designers statistical measures to be made. By making statistical analysis it may be possible to eliminate any outlier solutions obtained.

With performance improvements defined above, it may also be possible that tank types and tank counts are considered as variables. When tank counts are defined as variables, the application might be as follows: An upper limit for tank counts is set (i.e. half or 1/3 the number of nodes). Random initial populations with random tank counts having a maximum of the defined upper limit. Then the methodology will proceed to promote better elements within the search. Some tank counts in some elements will come to the forefront. This might again indicate some limit for the optimal tank count for specific networks.

Again, considering a high-performance hydraulic solver code, systems with big populations can be run for very long generations to get true pareto curves or best estimates of true pareto curves. By comparing other test runs with these best estimates, it may be possible to apply running performance metrics defined in Figure 2.2. Therefore, different parameter effects (mutation, crossover, population count, number of iterations as stopping criteria, etc.) on performance can be measured in a two-step manner.

REFERENCES

- Alperovits, E., and Shamir, U. (1977). "Design of optimal water distribution systems." *Water Resources Research*, 885–900.
- Ar, K. (2011). "Calibration of water distribution networks." Middle East Technical University.
- Basupi, I., and Kapelan, Z. (2015). "Flexible Water Distribution System Design under Future Demand Uncertainty." *Journal of Water Resources Planning and Management*, 141(4).
- Basupi, I., Kapelan, Z., and Butler, D. (2014). "Reducing life-cycle carbon footprint in the (re)design of water distribution systems using water demand management interventions." *Urban Water Journal*, 11(2), 91–107.
- Batchabani, E., and Fuamba, M. (2014). "Optimal Tank Design in Water Distribution Networks: Review of Literature and Perspectives." *Journal of Water Resources Planning and Management*, 136–145.
- Blasco, X., Herrero, J. M., and Martínez, M. (2008). "A new graphical visualization of n-dimensional Pareto front for decision-making in multiobjective optimization." *Information Sciences*, (172), 3908–3924.
- Çağlar, M. (2015). "Rehabilitation Studies About Storage Tanks: A Case Study." Middle East Technical University.
- Carvalho, A. G., and Araujo, A. E. R. (2009). "Improving NSGA-II with an Adaptive Mutation Operator." *Proceedings of the 11th Annual Conference Companion on Genetic and Evolutionary Computation Conference*, Montreal, 2697–2700.
- Cimorelli, L., Morlando, F., Cozzolino, L., D'Aniello, A., and Pianese, D. (2018). "Comparison Among Resilience and Entropy Index in the Optimal Rehabilitation of Water Distribution Networks Under Limited-Budgets." *Water Resources Management*, 32, 3997–4011.
- Čistý, M., and Bajtek, Z. (2009). "Optimization of the Water Distribution Networks

- with Search Space Reduction.” Slovak University of Technology, Bratislava.
- Coello Coello, C. A., and Montes, E. M. (2002). “Constraint-handling in genetic algorithms through the use of dominance-based tournament selection.” *Advanced Engineering Informatics*, (16), 193–203.
- Craenen, B. G. W., Eiben, A. E., and Marchiori, E. (2001). “How to Handle Constraints with Evolutionary Algorithms.” *The Practical Handbook of Genetic Algorithms, Applications*, L. D. Chambers, ed., Chapman and Hall/CRC, 10.1-10.6.
- Creaco, E., Fortunato, A., Franchini, M., and Mazzola, M. R. (2014). “Comparison between entropy and resilience as indirect measures of reliability in the framework of water distribution network design.” *Procedia Engineering*, (70), 379–388.
- Creaco, E., Franchini, M., and Todini, E. (2016a). “The combined use of resilience and loop diameter uniformity as a good indirect measure of network reliability.” *Urban Water Journal*, 13(2), 167–181.
- Creaco, E., Franchini, M., and Todini, E. (2016b). “Generalized Resilience and Failure Indices for Use with Pressure-Driven Modeling and Leakage.” *Journal of Water Resources Planning and Management*.
- Datta, R., and Deb, K. (2012). *An Adaptive Normalization based Constraint Handling Methodology with Hybrid Bi-Objective and Penalty Function Approach*. Kanpur Genetic Algorithms Laboratory - KanGAL, Kanpur.
- Deb, K. (2000). “An efficient constraint handling method for genetic algorithms.” *Computer Methods in Applied Mechanics and Engineering*, 311–338.
- Deb, K. (2001). *Multi-Objective Optimization using Evolutionary Algorithms*. John Wiley & Sons, West Sussex, England.
- Deb, K., and Agrawal, R. B. (1995). “Simulated Binary Crossover for Continuous Search Space.” *Complex Systems*, 9(2).
- Deb, K., and Agrawal, S. (1999). “A Niched-Penalty Approach for Constraint

- Handling in Genetic Algorithms.” *Proceedings of the International Conference on Artificial Neural Networks and Genetic Algorithms*, Portoroz, 235–243.
- Deb, K., and Jain, H. (2014). “An Evolutionary Many-Objective Optimization Algorithm Using Reference-Point-Based Nondominated Sorting Approach, Part I: Solving Problems With Box Constraints.” *IEEE Transactions on Evolutionary Computation*, 18(4), 577–601.
- Deb, K., and Jain, S. (2002). *Running Performance Metrics for Evolutionary Multi-Objective Optimization*. Kanput Genetic Algorithms Laboratory - KanGAL, Kanpur.
- Deb, K., and Jain, S. (2004). “Evaluating Evolutionary Multi-Objective Optimization Algorithms Using Running Performance Metrics.” *Recent Advances in Simulated Evolution and Learning*, K. C. Tan, M. H. Lim, X. Yao, and L. Wang, eds., World Scientific Publishing, Toh Tuck, 307–326.
- Deb, K., Mohan, M., and Mishra, S. (2003). *A Fast Multi-objective Evolutionary Algorithm for Finding Well-Spread Pareto-Optimal Solutions*. Kanpur Genetic Algorithms Laboratory - KanGAL, Kanpur.
- Deb, K., Pratap, A., Agarwal, S., and Meyarivan, T. (2002). “A Fast and Elitist Multiobjective Genetic Algorithm: NSGA-II.” *IEEE Transactions on Evolutionary Computation*, 187–197.
- Doruk, B. L. (2001). “Identifying efficient pump combinations on water distribution networks management.” Middle East Technical University.
- Dziedzic, R., and Bryan, W. K. (2016). “Performance Index for Water Distribution Networks under Multiple Loading Conditions.” *Journal of Water Resources Planning and Management*, 142(1).
- Eker, İ. (1998). “Analysis of water distribution networks using-informatis systems.” Middle East Technical University.
- Farmani, R., Savic, D. A., and Walters, G. A. (2004). “The Simultaneous Multi-

Objective Optimization of Anytown Pipe Rehabilitation, Tank Sizing, Tank Siting, and Pump Operation Schedules.” *Critical Transitions in Water and Environmental Resources*.

Farmani, R., Walters, G. A., and Savic, D. A. (2005). “Trade-off between Total Cost and Reliability for Anytown Water Distribution Network.” *Journal of Water Resources Planning and Management*, 161–171.

Farmani, R., Walters, G., and Savic, D. (2006). “Evolutionary multi-objective optimization of the design and operation of water distribution network: total cost vs. reliability vs. water quality.” *Journal of Hydroinformatics*, 8(3), 165–179.

Fu, G., Kapelan, Z., and Reed, P. (2012). “Reducing the Complexity of Multiobjective Water Distribution System Optimization through Global Sensitivity Analysis.” *Journal of Water Resources Planning and Management*, 138(3), 196–207.

Fujiwara, O., and Khang, D. B. (1990). “A two-phase decomposition method for optimal design of looped water distribution networks.” *Water Resources Research*, 26(4), 539–549.

Fujiwara, O., and Khang, D. B. (1991). “Correction to ‘A two-phase decomposition method for optimal design of looped water distribution networks’ by Okitsugu Fujiwara and Do Ba Khang.” *Water Resources Research*, John Wiley & Sons, Ltd, 27(5), 985–986.

Gençoğlu, G. (2007). “Developing a Methodology for the Design of Water Distribution Networks Using Genetic Algorithm.” Middle East Technical University, Ankara.

Gençoğlu, G. (2015). “Optimizing pump schedule and valve characteristics of water distribution networks.” Middle East Technical University.

Güç, G. (2006). “Optimization of water distribution networks using genetic algorithm.” Middle East Technical University.

- Halhal, D., Walters, G. A., and Ouazar, D. (1995). "Structured Messy Genetic Algorithm Approach for the Optimal Improvement of Water Distribution Systems." *Genetic Algorithms in Engineering Systems: Innovations and Applications*, 406–411.
- Halhal, D., Walters, G. A., Ouazar, D., and Savic, D. A. (1997). "Water Network Rehabilitation With Structured Messy Genetic Algorithm." *Journal of Water Resources Planning and Management*, 123(3), 137–146.
- Haupt, R. L., and Haupt, S. E. (2004). *Practical Genetic Algorithms*. John Wiley & Sons, Hoboken New Jersey.
- "İçmesuyu Tesisleri Etüt, Fizibilite ve Projelerinin Hazırlanmasına Ait Teknik Şartname." (2013). Ankara.
- Jain, H., and Deb, K. (2014). "An Evolutionary Many-Objective Optimization Algorithm Using Reference-Point-Based Nondominated Sorting Approach, Part II: Handling Constraints and Extending to an Adaptive Approach." *IEEE Transactions on Evolutionary Computation*, 18(4), 602–622.
- Jeong, G., Wicaksono, A., and Kang, D. (2017). "Revisiting the Resilience Index for Water Distribution Networks." *Journal of Water Resources Planning and Management*, 143(8).
- Keleş, G. (2005). "Water Distribution Network Design By Partial Enumeration." Middle East Technical University.
- Kennedy, J., and Eberhart, R. (1995). "Particle Swarm Optimization." *Proceedings of the IEEE international conference on neural networks IV*, Piscataway, 1942–1948.
- Knowles, J., and Corne, D. (2002). "On Metrics for Comparing Non-Dominated Sets." *Proceedings of the 2002 Congress on Evolutionary Computation. CEC '02*, IEEE, 711–716.
- Kukkonen, S., and Deb, K. (2006). *Improved Pruning of Non-Dominated Solutions Based on Crowding Distance for Bi-Objective Optimization Problems*. Kanpur

Genetic Algorithms Laboratory - KanGAL, Kanpur.

- Kukkonen, S., and Deb, K. (2007). *A Fast and Effective Method for Pruning of Non-Dominated Solutions in Many-Objective Problems*. Kanpur Genetic Algorithms Laboratory - KanGAL, Kanpur.
- Liu, H., Savić, D. A., Kapelan, Z., Creaco, E., and Yuan, Y. (2017). “Reliability Surrogate Measures for Water Distribution System Design: Comparative Analysis.” *Journal of Water Resources Planning and Management*, 143(2).
- Maier, H. R., Razavi, S., Kapelan, Z., Matott, L. S., Kasprzyk, J., and Tolson, B. A. (2019). “Introductory overview: Optimization using evolutionary algorithms and other metaheuristics.” *Environmental Modelling & Software*, (114), 195–213.
- Mala-Jetmarova, H., Sultanova, N., and Savic, D. (2017). “Lost in optimisation of water distribution systems? A literature review of system operation.” *Environmental Modelling & Software*, (93), 209–254.
- Mareschal, B., and De Smet, Y. (2009). “Visual PROMETHEE: Developments of the PROMETHEE & GAIA multicriteria decision aid methods.” *2009 IEEE International Conference on Industrial Engineering and Engineering Management*, Hong Kong, 1646–1649.
- Marques, J., Cunha, M., and Savić, D. (2018). “Many-objective optimization model for the flexible design of water distribution networks.” *Journal of Environmental Management*, (226), 308–319.
- Masin, M., and Bukchin, Y. (2008). “Diversity Maximization Approach for Multiobjective Optimization.” *Operations Research*, 56(2), 411–424.
- Mays, L. W., and Tung, Y. K. (2002). *Hydrosystems Engineering and Management*. Water Resources Publications.
- Michalewics, Z., and Schoenauer, M. (1996). “Evolutionary Algorithms for Constrained Parameter Optimization Problems.” *Evolutionary Computation*, 4(1), 1–32.

- Mitchell, M. (1999). *An Introcution To Genetic Algorithms*. MIT Press.
- Monsef, H., Naghashzadegan, M., Farmani, R., and Jamali, A. (2019a).
 “Deficiency of Reliability Indicators in Water Distribution Networks.”
Journal of Water Resources Planning and Management, 145(6).
- Monsef, H., Naghashzadegan, M., Jamali, A., and Farmani, R. (2019b).
 “Comparison of evolutionary multi objective optimization algorithms in
 optimum design of water distribution network.” *Ain Shams Engineering
 Journal*, (10), 103–111.
- Montalvo, I., Izquierdo, J., Pérez, R., and Iglesias, P. L. (2008). “A diversity-
 enriched variant of discrete PSO applied to the design of water distribution
 networks.” *Engineering Optimization*, 40(7), 655–668.
- Nadiroğlu, C. (2014). “Qualitative evaluations about the disinfection capabilities of
 a water distribution network: A model study.” Middle East Technical
 University.
- Nohutçu, M. (2002). “Analysis of water distribution networks with pressure
 dependent demand models.” Middle East Technical University.
- Özand, E. (1939). “Su işleri projesine ait fennî şartname.” *Belediyeler Dergisi*,
 Belediyeler Bankası, Ankara, (41), 17–22.
- Özkan, T. (2001). “Determination of leakages in a water distribution network using
 SCADA data.” Middle East Technical University.
- Özkan, T. (2008). “Leakage control by optimal valve operation.” Middle East
 Technical University.
- Pandey, S., and Mathur, Y. P. (2013). “Optimization of Water Dstribution Network
 Using Particle Swarm Optimization.” *International Journal of Applied
 Engineering Research*, 8(12), 92–101.
- Prasad, T. D., Hong, S.-H., and Park, N. (2003). “Reliability Based Design of
 Water Distribution Networks Using Muti-Objective Genetic Algorithms.”
KSCE Journal of Civil Engineering, 7(3), 351–361.

- Prasad, T. D., and Park, N.-S. (2004). "Multiobjective Genetic Algorithms for Design of Water Distribution Networks." *Journal of Water Resources Planning and Management*, 130(1), 73–82.
- Prasad, T. D., and Tanyimboh, T. T. (2008). "Entropy Based Design of 'Anytown' Water Distribution Network." *Proceedings of the 10th Annual Water Distribution Systems Analysis Conference WDSA2008*, Kruger National Park, 450–461.
- Raad, D. N., Sinske, A. N., and van Vuuren, J. H. (2010). "Comparison of four reliability surrogate measures for water distribution systems design." *Water Resources Research*, 46.
- "Resilience." (n.d.). *Dictionary by Merriam-Webster*, <<https://www.merriam-webster.com/dictionary/resilience>> (Aug. 5, 2019).
- Rossman, L. A. (2000). *EPANET 2 User's Manual*. U.S. Environment Protection Agency, Cincinnati, OH.
- Salehi, S., Jalili Ghazizadeh, M., and Tabesh, M. (2018). "A comprehensive criteria-based multi-attribute decision-making model for rehabilitation of water distribution systems." *Structure and Infrastructure Engineering*, 14(6), 743–765.
- Schilling, K. E., and Porter, E. (Eds.). (1990). *Urban Water Infrastructure*. Springer Netherlands, Dordrecht.
- Seada, H., and Deb, K. (2015). "U-NSGA-III: A Unified Evolutionary Optimization Procedure for Single, Multiple, and Many Objectives: Proof-of-Principle Results." *Evolutionary Multi-Criterion Optimization. EMO 2015. Lecture Notes in Computer Science*, 34–49.
- "Şehir ve Kasaba İçmesuyu Projelerinin Hazırlanmasına Ait Yönetmelik." (1985). İller Bankası Genel Müdürlüğü, Ankara.
- "Şehir ve Kasabalara Getirilecek İçme Suyu Projelerinin Hazırlanmasına Ait Talimatname III." (n.d.). Akın Matbaası, Ankara.

- Şendil, H. (2013). "Operation of Water Distribution Networks." Middle East Technical University, Ankara.
- Shin, S., Lee, S., Judi, D. R., Parvania, M., Goharian, E., McPherson, T., and Burian, S. J. (2018). "A Systematic Review of Quantitative Resilience Measures for Water Infrastructure Systems." *Water*, 10.
- Sivanandam, S. N., and Deepa, S. N. (2008). *Introduction To Genetic Algorithms*. Springer, Berlin.
- Srinivas, N., and Deb, K. (1994). "Multiobjective Optimization Using Nondominated Sorting in Genetic Algorithms." *Journal of Evolutionary Computation*, 221–248.
- Steel Water-Storage Tanks M42 Manual of Water Supply Practices*. (2013). American Water Works Association, USA.
- Tanyimboh, T. T., and Seyoum, A. G. (2016). "Multiobjective evolutionary optimization of water distribution systems: Exploring diversity with infeasible solutions." *Journal of Environmental Management*, (183), 133–141.
- Tanyimboh, T. T., and Templeman, A. B. (2000). "A Quantified Assessment of the Relationship Between the Reliability and Entropy of Water Distribution Systems." *Engineering Optimization*, 33(2), 179–199.
- "The Water Distribution System Hydraulic and Water Quality Analysis Toolkit." (2019). <<https://github.com/OpenWaterAnalytics/EPANET>> (Sep. 16, 2019).
- Todini, E. (2000). "Looped water distribution networks design using a resilience index based heuristic approach." *Urban Water* 2, 115–122.
- Vachhani, V. L., Dabhi, V. K., and Prajapati, H. B. (2016). "Improving NSGA-II For Solving Multi Objective Function Optimization Problems." *International Conference on Computer Communication and Informatics*, Coimbatore.
- Walski, T. M., Brill, E. D., Gessler, J., Goulter, I., Jeppson, R. M., Lansey, K., Lee, H., Liebman, J. C., Mays, L., Morgan, D. R., and Ormsbee, L. (1987). "Battle of the Network Models: Epilogue." *Journal of Water Resources Planning and*

- Management*, 113(2), 191–203.
- Walters, G. A., Halhal, D., Savic, D., and Ouazar, D. (1999). “Improved Design of Anytown Distribution Network Usng.” *Urban Water*, 1(1), 23–38.
- Wang, Q., Guidolin, M., Savic, D., and Kapelan, Z. (2015). “Two-Objective Design of Benchmark Problems of a Water Distribution System via MOEAs: Towards the Best-Known Approximation of the True Pareto Front.” *Journal of Water Resources Planning and Management*, 141(3), 4014060.
- Wang, Q., Wang, L., Huang, W., Wang, Z., Liu, S., and Savić, D. A. (2019). “Parameterization of NSGA-II for the Optimal Design of Water Distribution Systems.” *Water*, 11(5).
- While, L., Hingston, P., Barone, L., and Huband, S. (2006). “A Faster Algorithm for Calculating Hypervolume.” *IEEE Transactions on Evolutionary Computation*, 10(10), 29–38.
- Yağız, E. (1999). “Hydraulic modelling of water distribution networks.” Hydraulic modelling of water distribution networks.
- Yeniay, Ö. (2005). “Penalty Function Methods for Constrained Optimization with Genetic Algorithms.” *Mathematical and Computational Applications*, 10(1), 45–56.
- Yıldız, E. (2011). “Evaluation of performance and optimum valve settings for pressure management using forecasted daily demand curves by artificial neural networks.” Middle East Technical University.
- Yu, P. L. (1973). “A Class of Solutions for Group Decision Problems.” *Management Science*, 19(8), 936–946.
- Zadorojniy, A., Masin, M., Greenberg, L., Shir, O. M., and Zeidner, L. (2012). “Algorithms for Finding Maximum Diversity of Design Variables in Multi-Objective Optimization.” *Procedia Computer Science*, (8), 171–176.
- Zeleny, M. (1973). “Compromise Programming.” *Multiple Criteria Decision Making*, J. L. Cochrane and M. Zeleny, eds., University of South crolina,

Columbia, 262–301.

Zeleny, M. (2011). “Multiple Criteria Decision Making (MCDM): From Paradigm Lost to Paradigm Regained?” *Journal of Multi-Criteria Decision Analysis*, (18), 77–89.

Zheng, J., Shen, R., and Zou, J. (2012). “Enhancing Diversity for NSGA-II in Evolutionary Multi-Objective Optimization.” *8th International Conference on Natural Computation*, Chongqing, 654–657.

Zitzler, E., and Thiele, L. (1998). “Multiobjective Optimization Using Evolutionary Algorithms - A Comparative Case Study.” *Fifth International Conference on Parallel Problem Solving from Nature (PPSN-V)*, A. E. Eiben, T. Bäck, M. Schoenauer, and H.-P. Schwefel, eds., Springer, Berlin, 292–301.

Zitzler, E., and Thiele, L. (1999). “Multiobjective Evolutionary Algorithms: A Comparative Case Study and the Strength Pareto Approach.” *IEEE Transactions on Evolutionary Computation*, 3(4), 257–271.

APPENDICES

A. APPENDIX FOR 2L SOLUTIONS

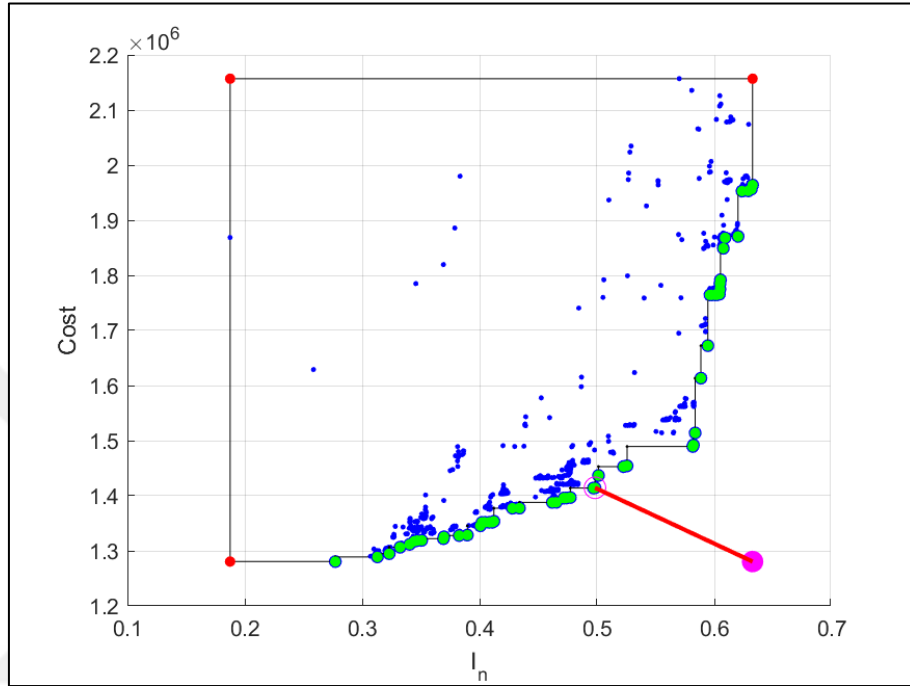


Figure A1. Pareto Curve for 2L50P1T

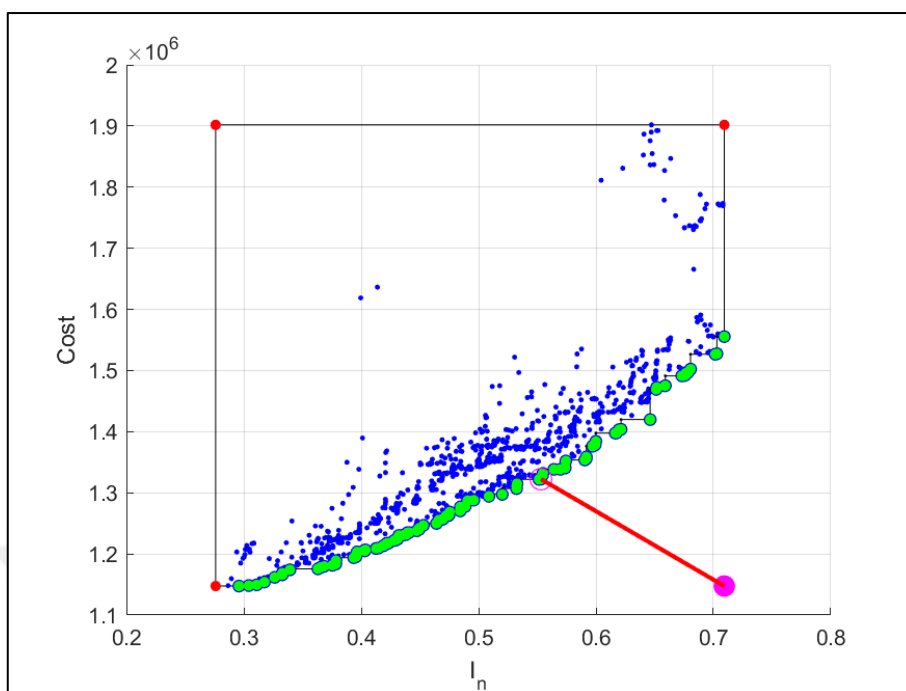


Figure A2. Pareto Curve for 2L100P1T

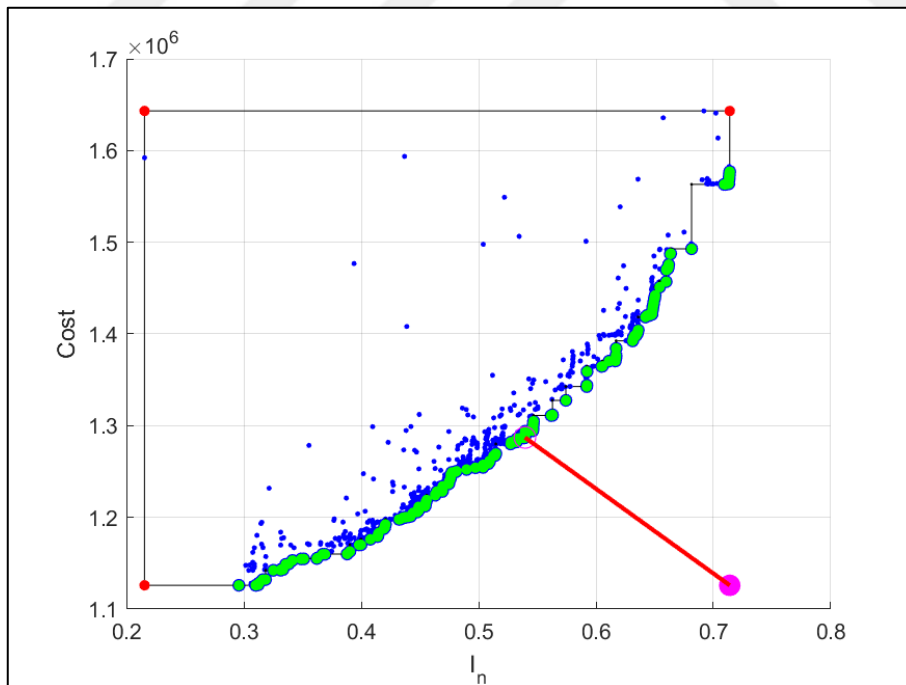


Figure A3. Pareto Curve for 2L200P1T

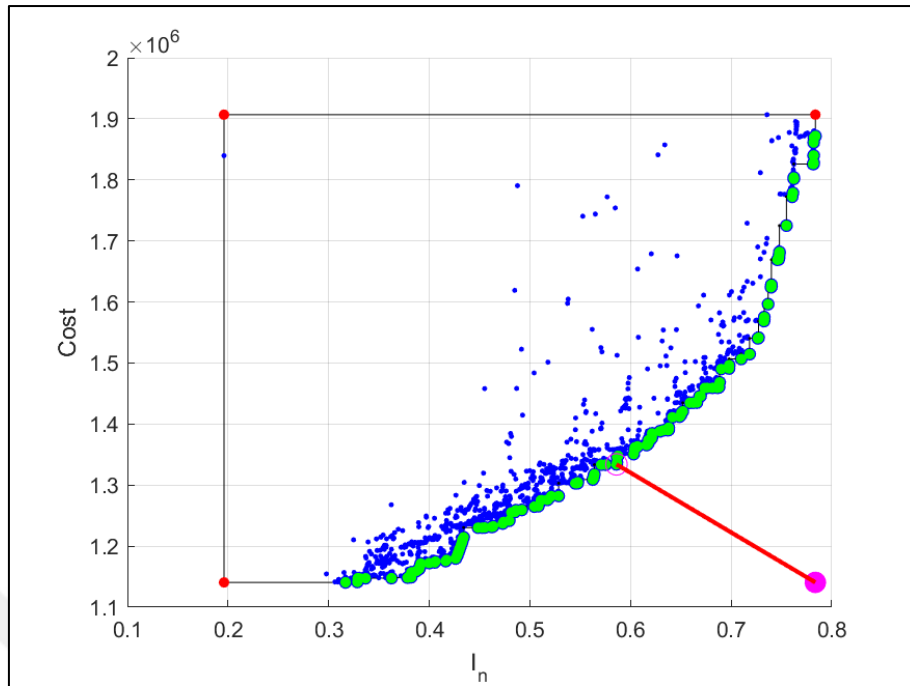


Figure A4. Pareto Curve for 2L250P1T

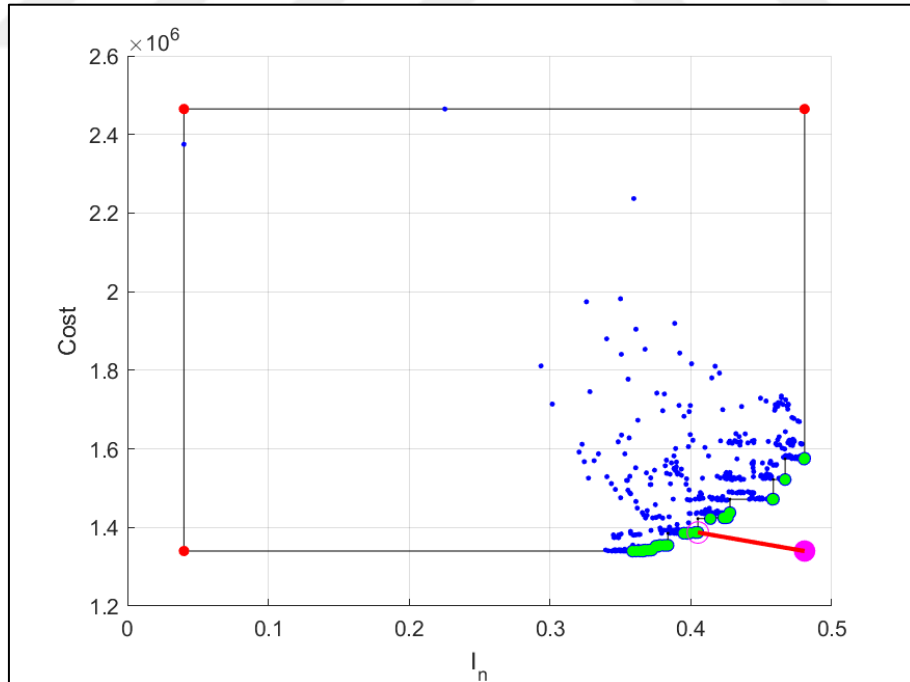


Figure A5. Pareto Curve for 2L50P2T

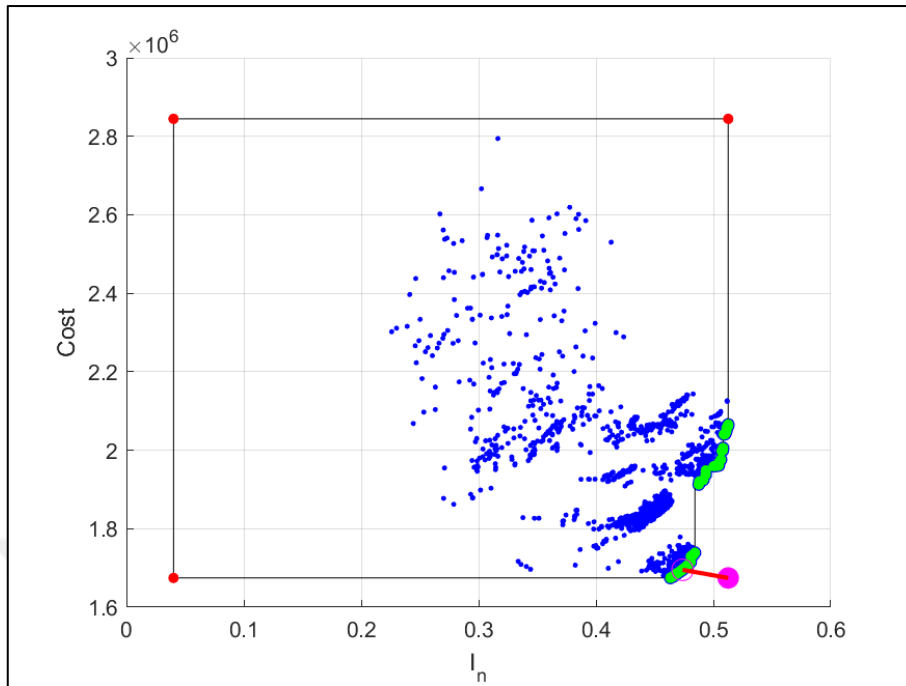


Figure A6. Pareto Curve for 2L100P2T

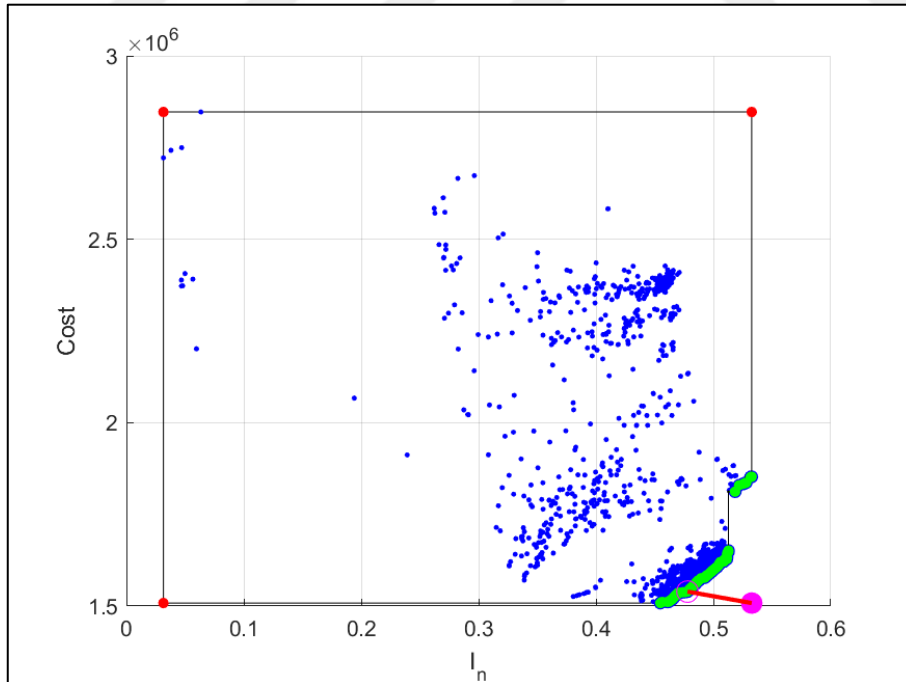


Figure A7. Pareto Curve for 2L150P2T

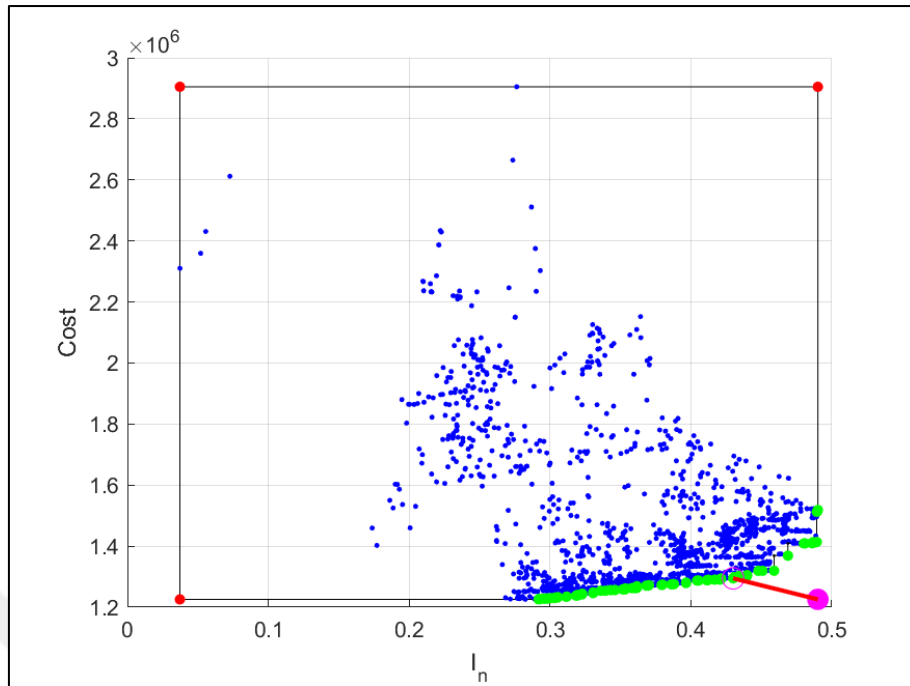


Figure A8. Pareto Curve for 2L200P2T

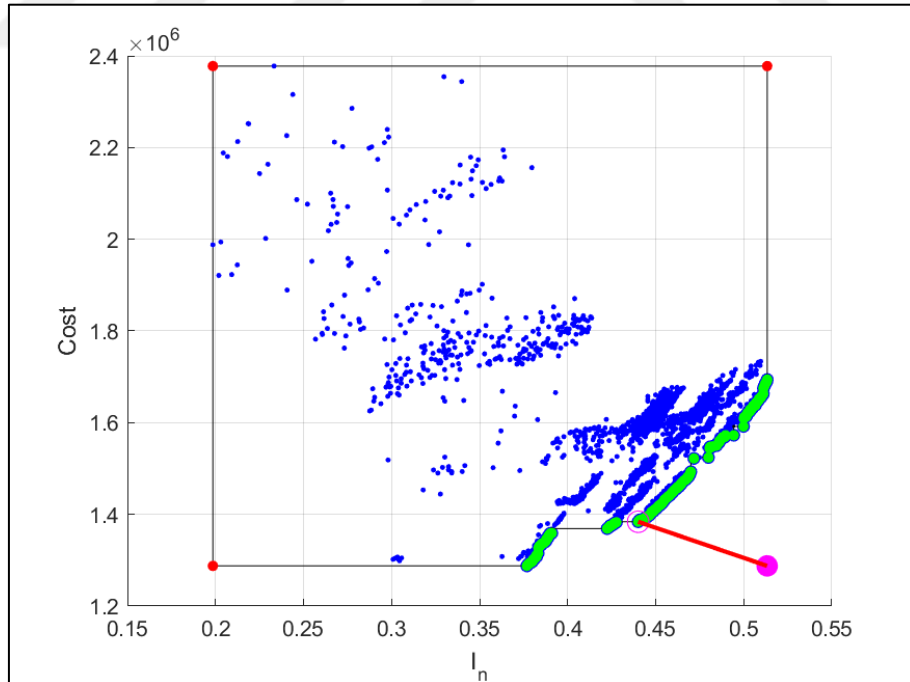


Figure A9. Pareto Curve for 2L250P2T

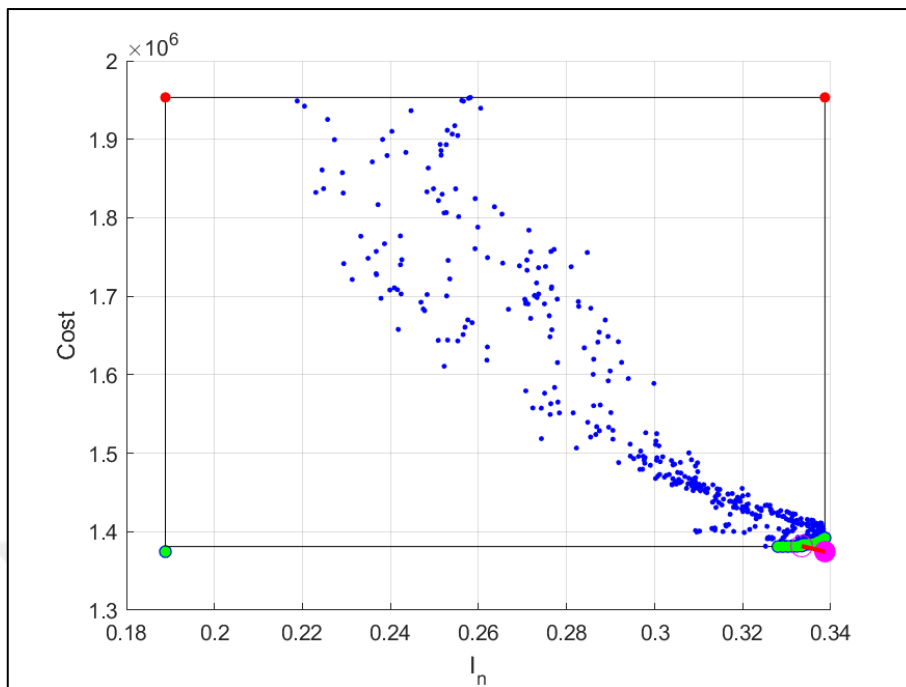


Figure A10. Pareto Curve for 2L50P3T

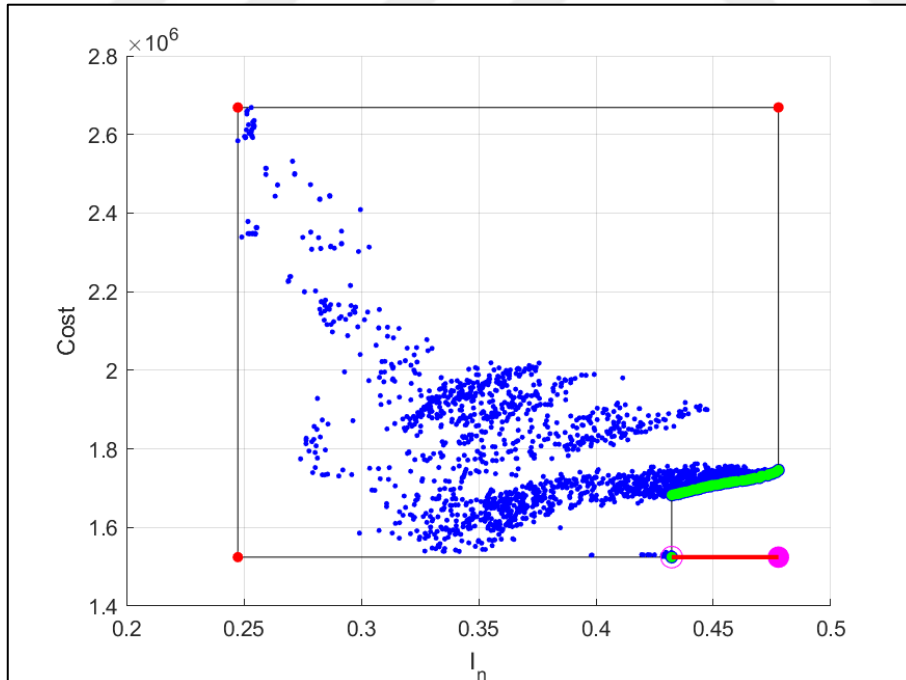


Figure A11. Pareto Curve for 2L100P3T

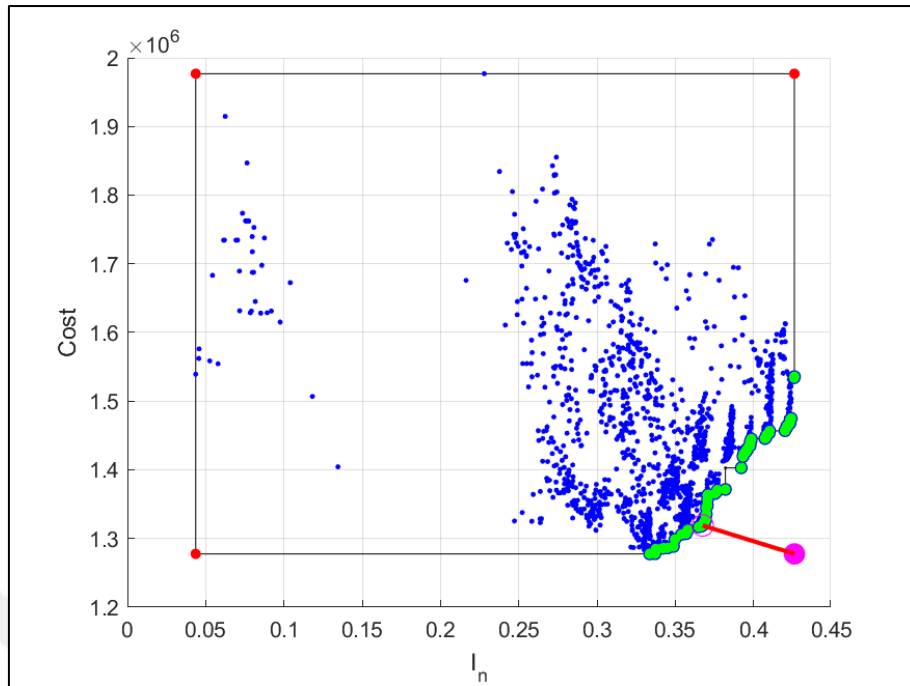


Figure A12. Pareto Curve for 2L150P3T

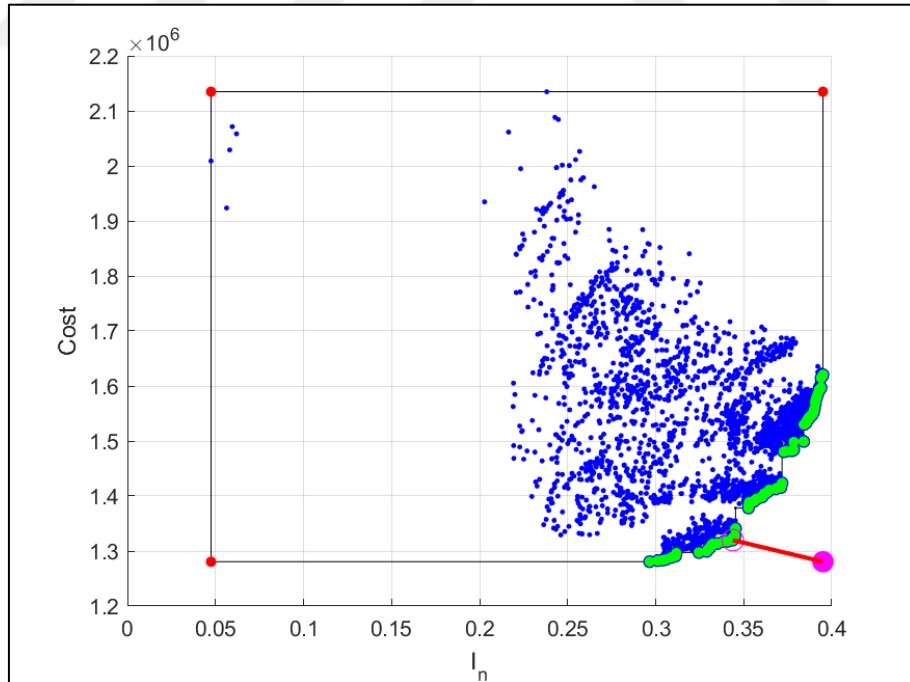


Figure A13. Pareto Curve for 2L200P3T

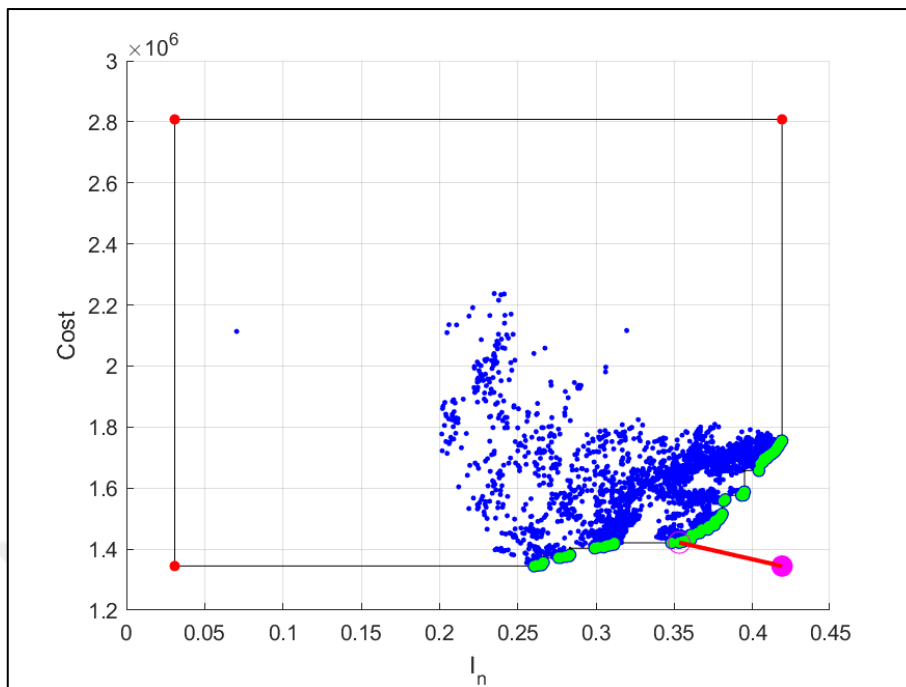


Figure A14. Pareto Curve for 2L250P3T

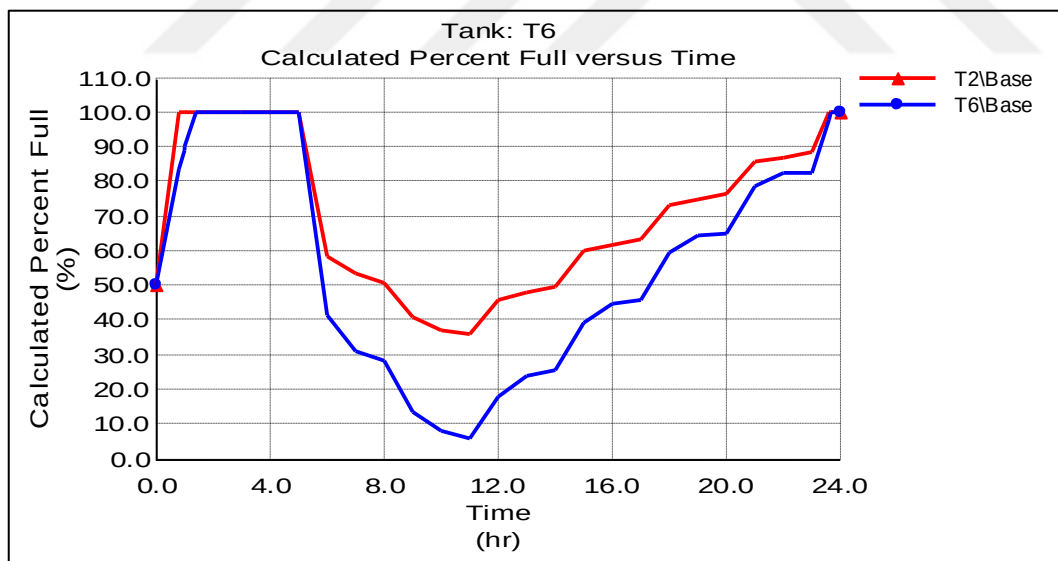


Figure A15. Tank Oscillations for Solution 2L200P2T/175

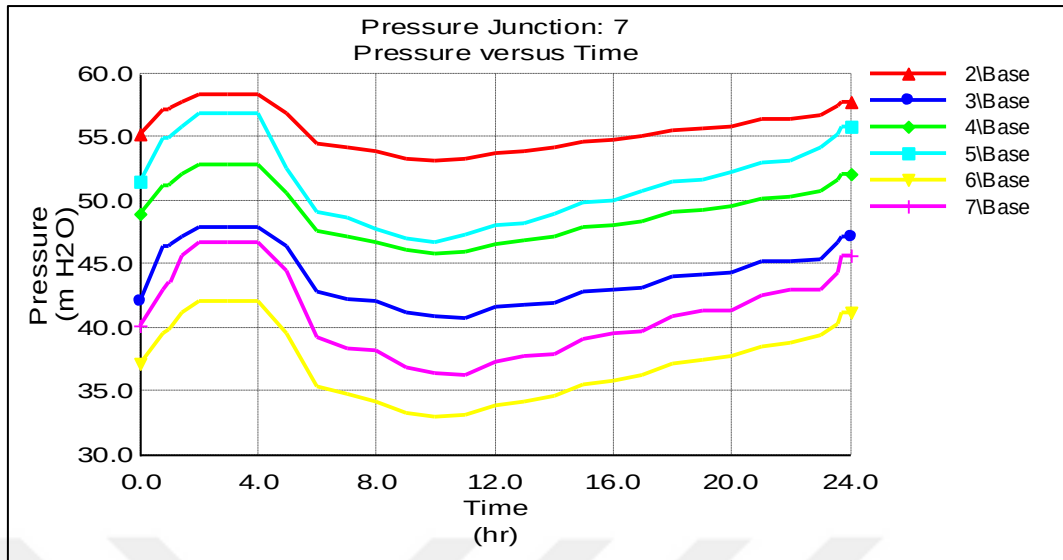


Figure A16. Pressure Variations for Solution 2L200P2T/175

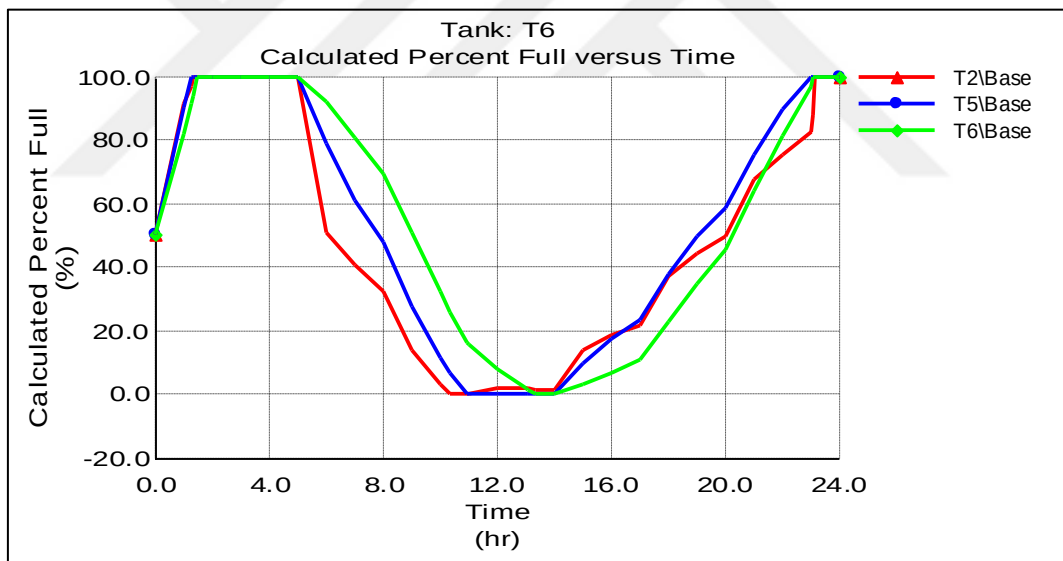


Figure A17. Tank Oscillations for Solution 2L250P3T/235

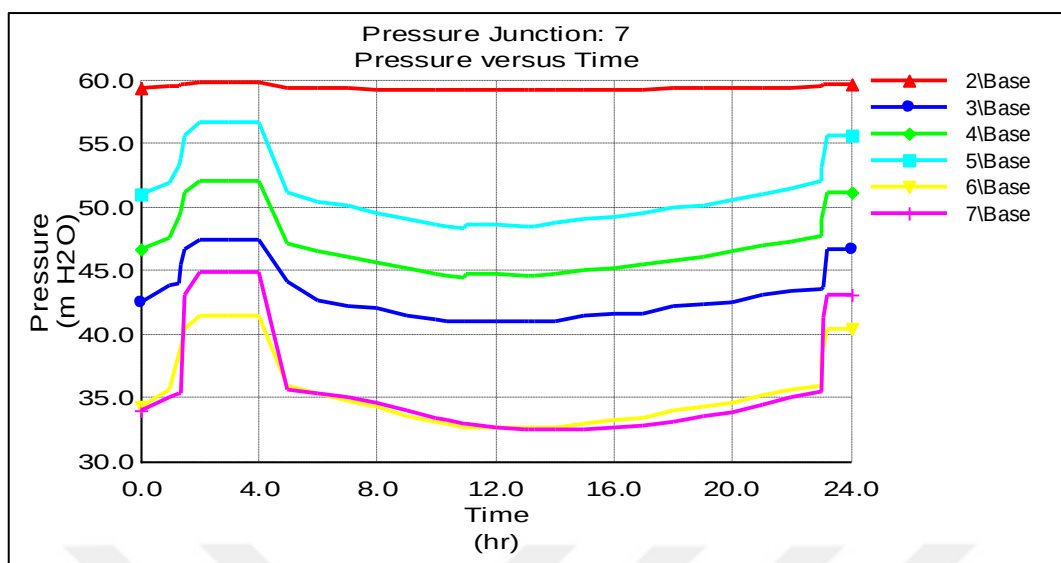


Figure A18. Pressure Variations for Solution 2L250P3T/235

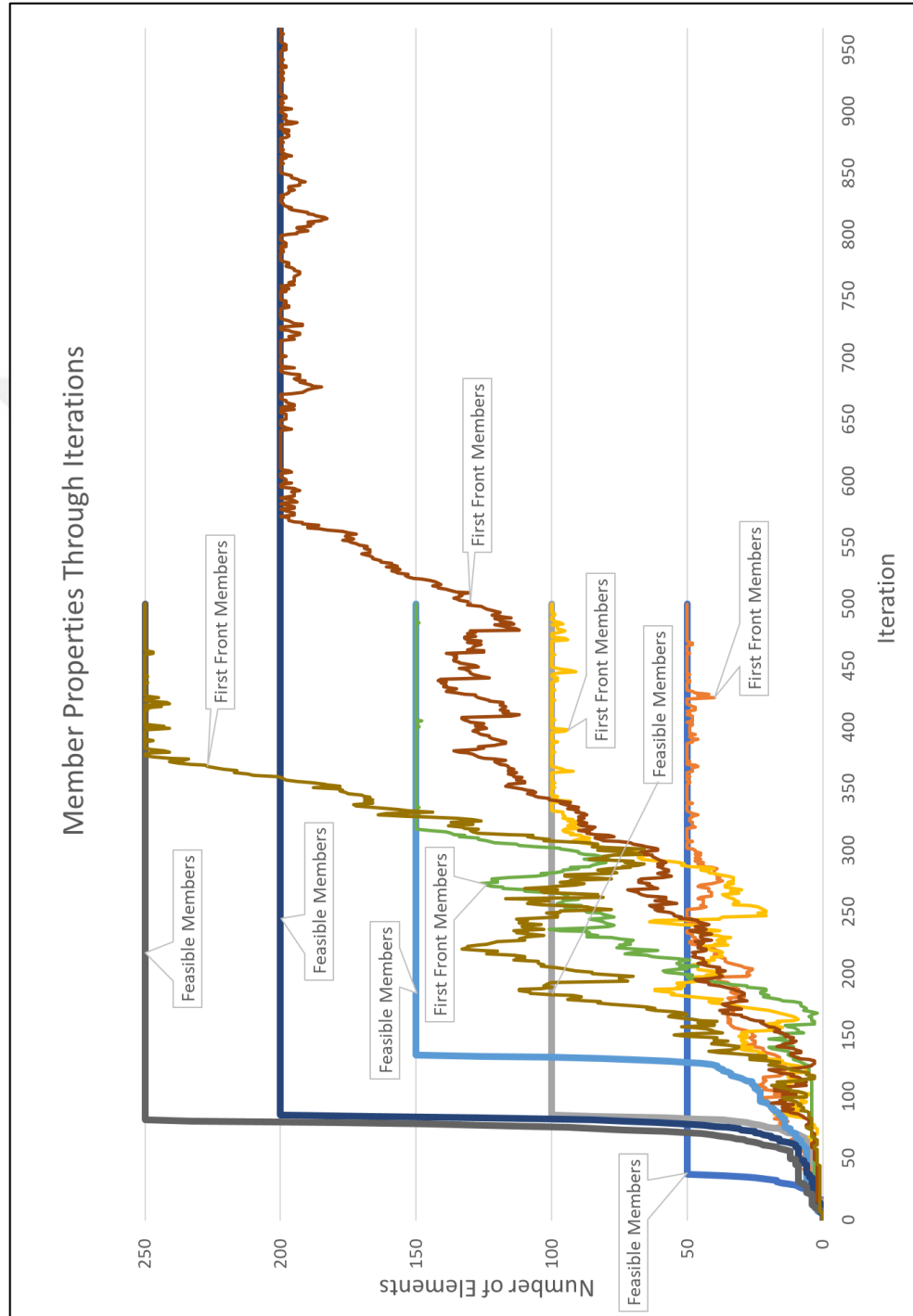


Figure A19. Member Properties for 2L2T

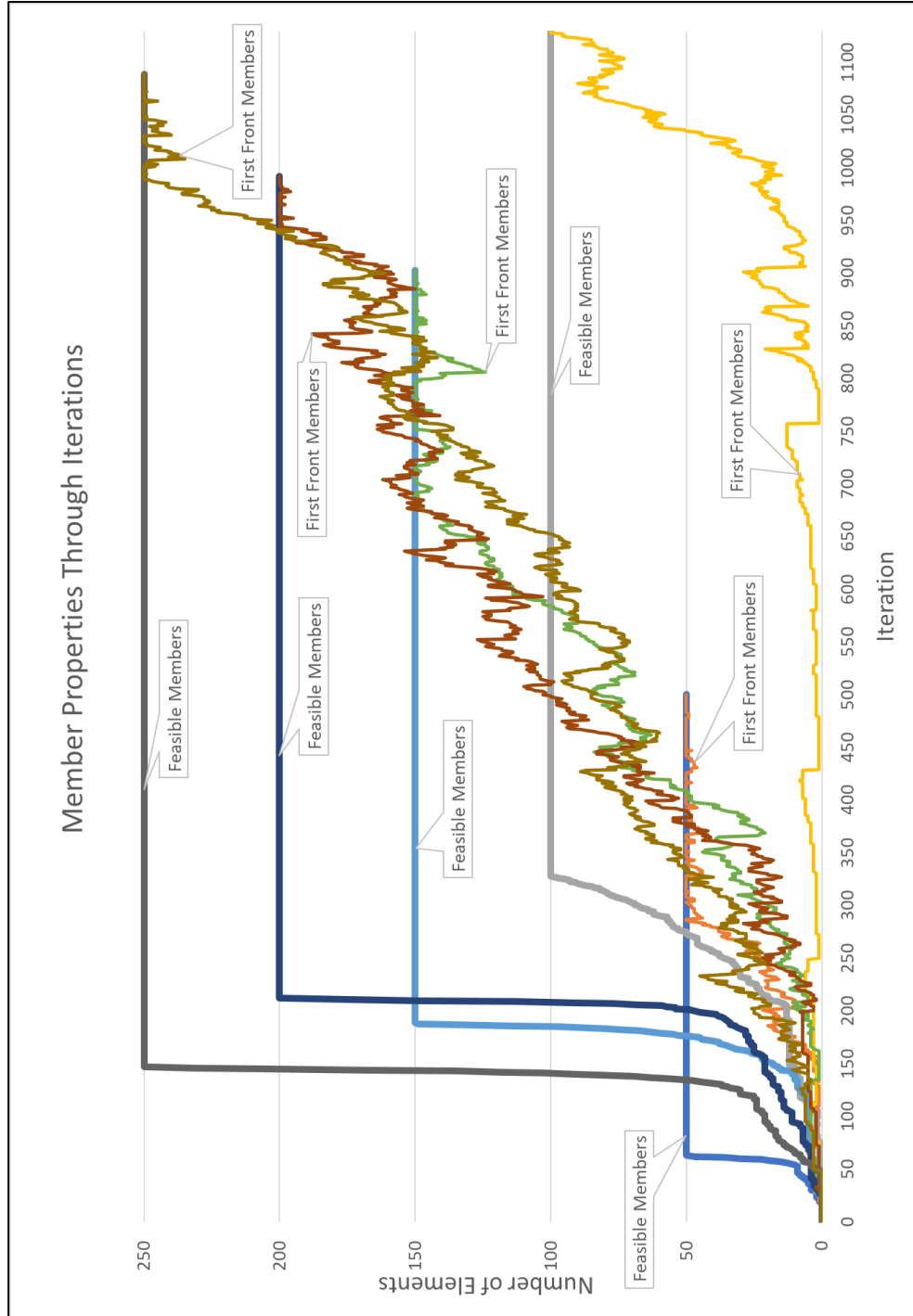


Figure A20. Member Properties for 2L3T

B. APPENDIX FOR HAN SOLUTIONS

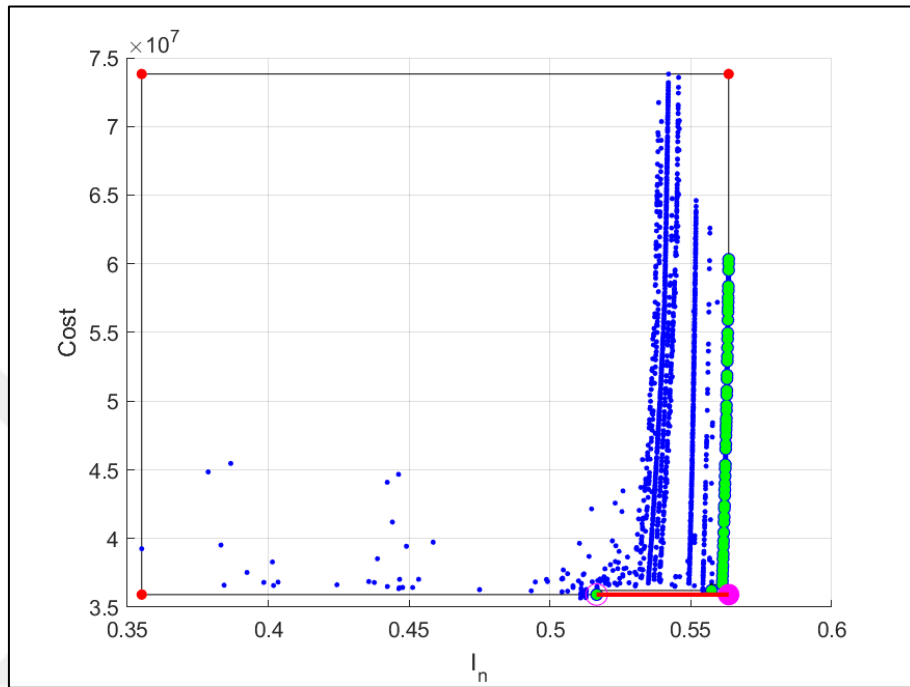


Figure B1. Pareto Curve for HAN50P1T

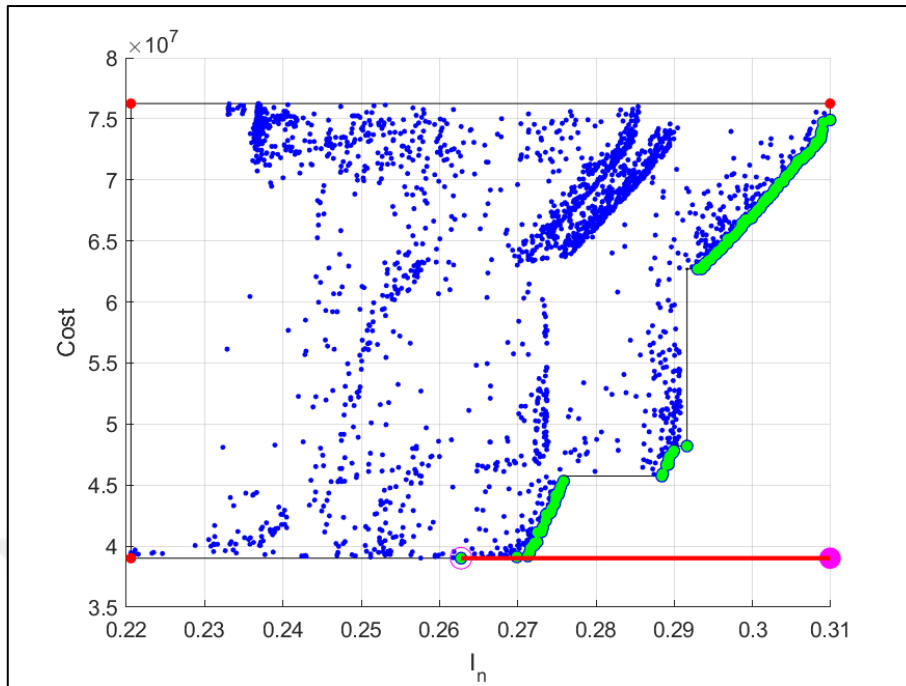


Figure B2. Pareto Curve for HAN100P1T

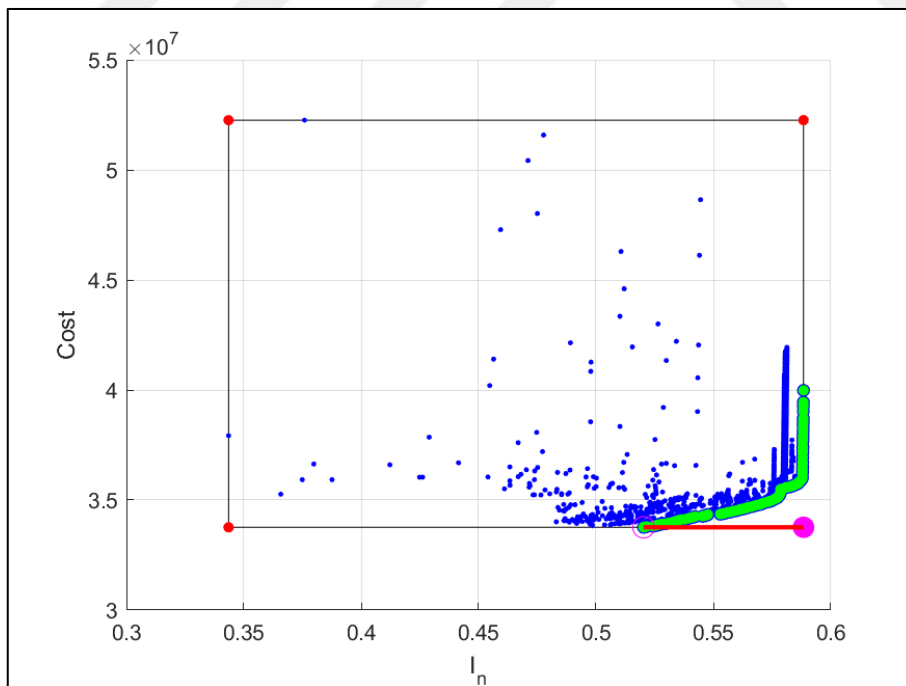


Figure B3. Pareto Curve for HAN200P1T

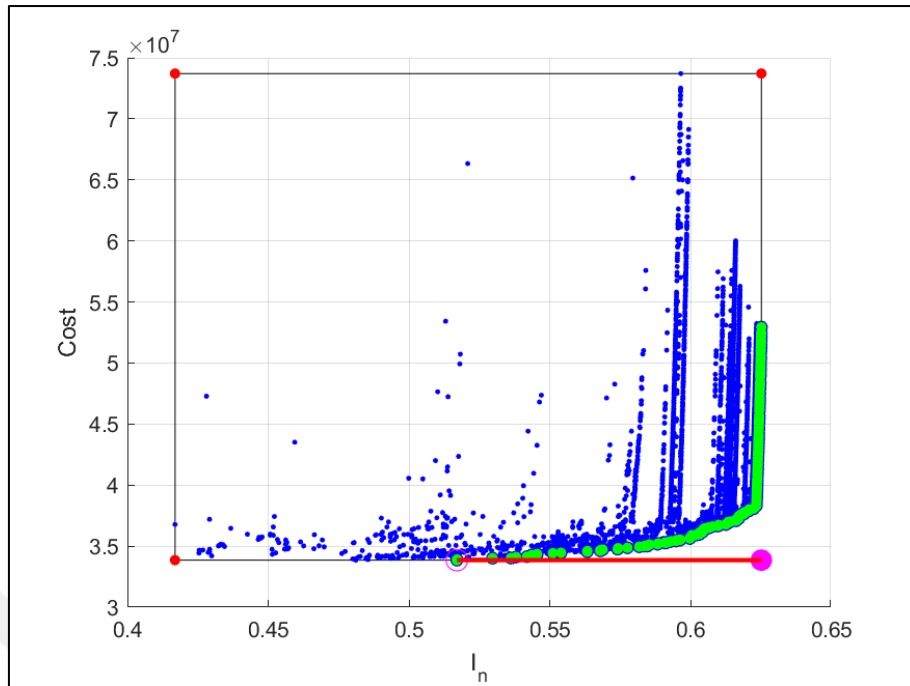


Figure B4. Pareto Curve for HAN250P1T

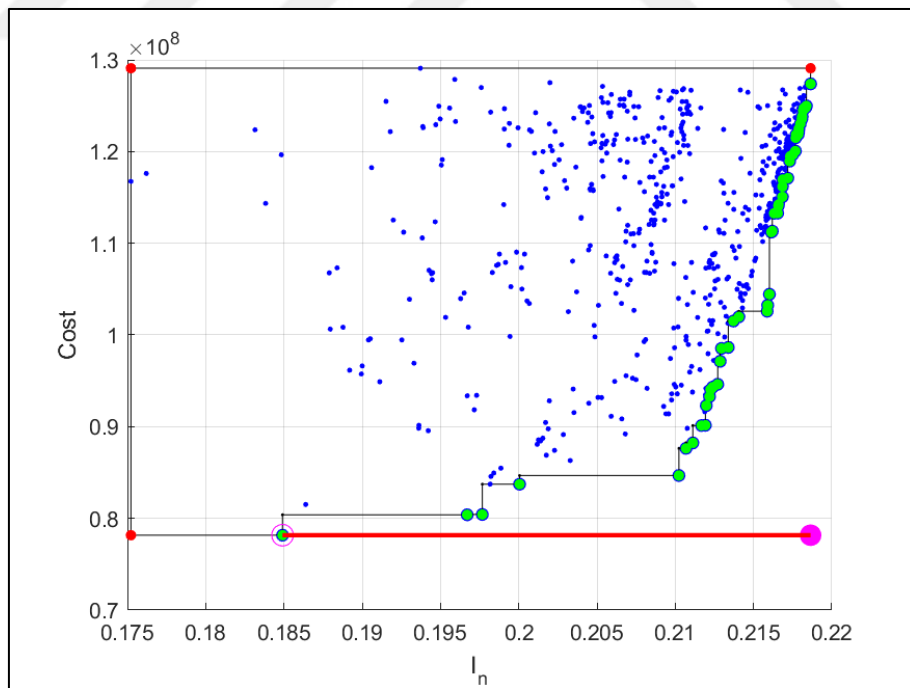


Figure B5. Pareto Curve for HAN50P2T

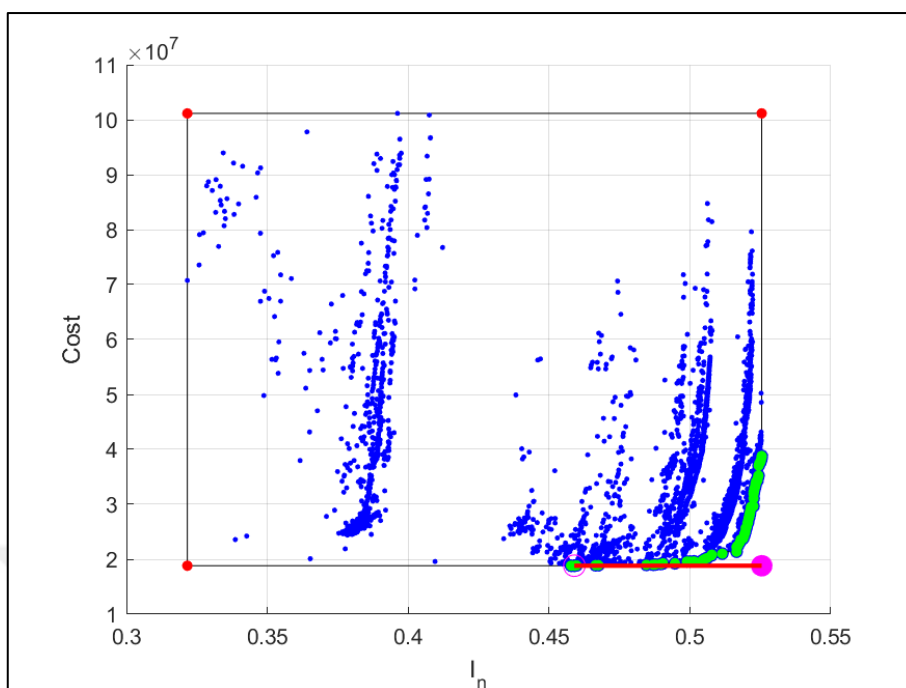


Figure B6. Pareto Curve for HAN100P2T

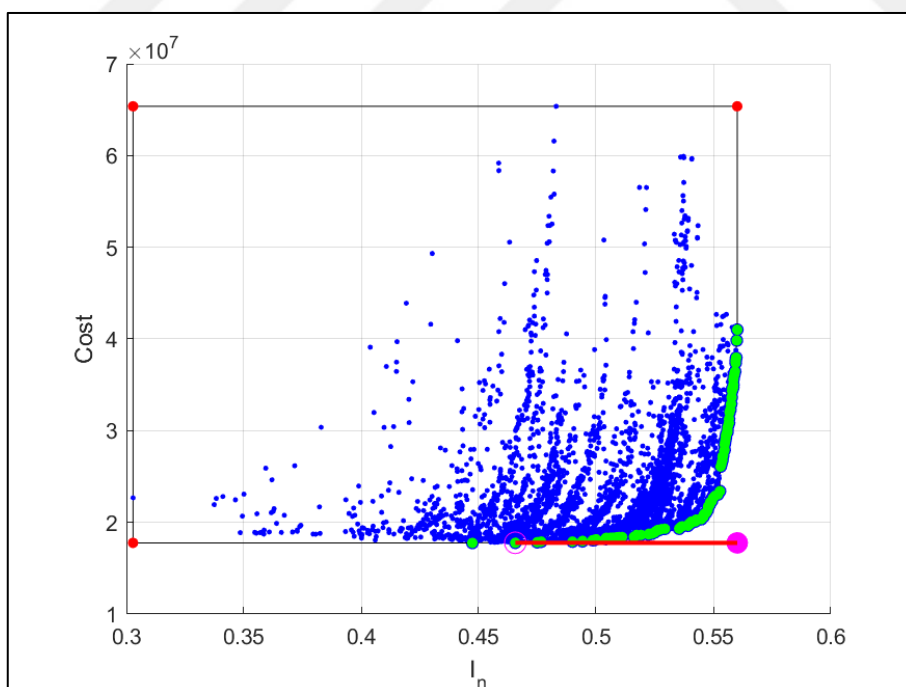


Figure B7. Pareto Curve for HAN150P2T

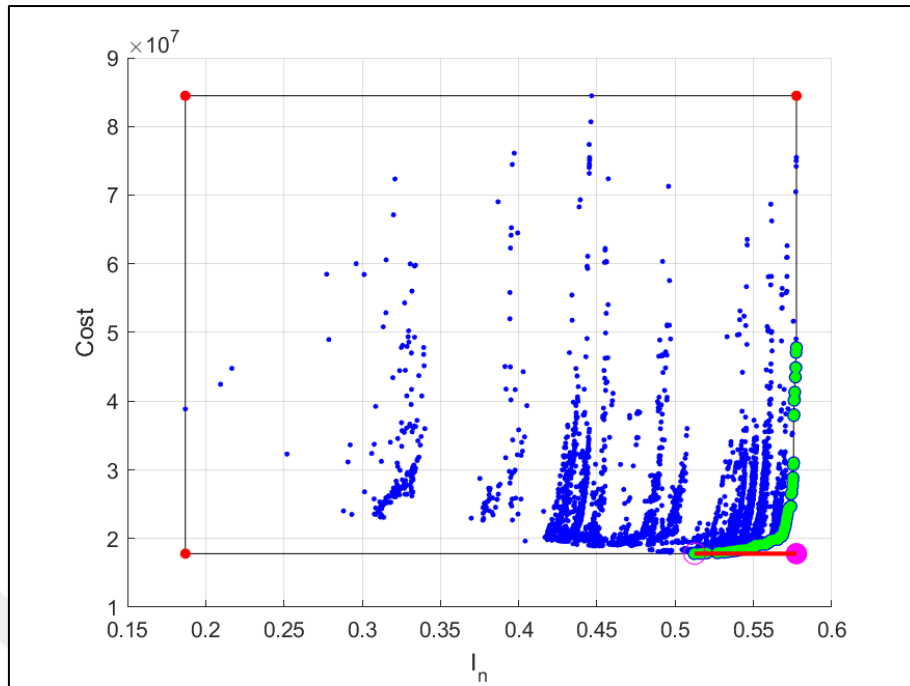


Figure B8. Pareto Curve for HAN200P2T

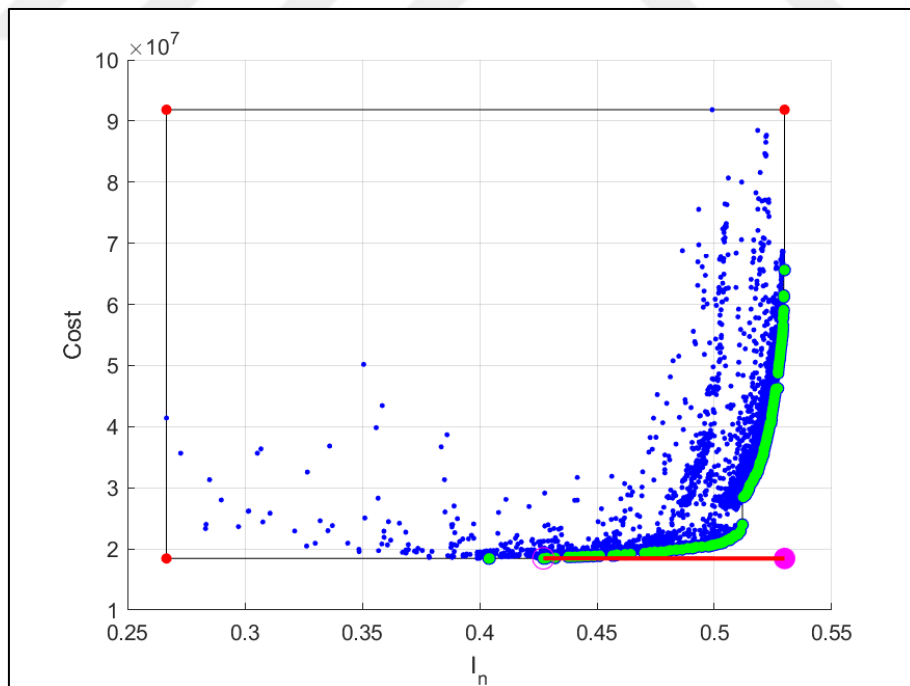


Figure B9. Pareto Curve for HAN250P2T

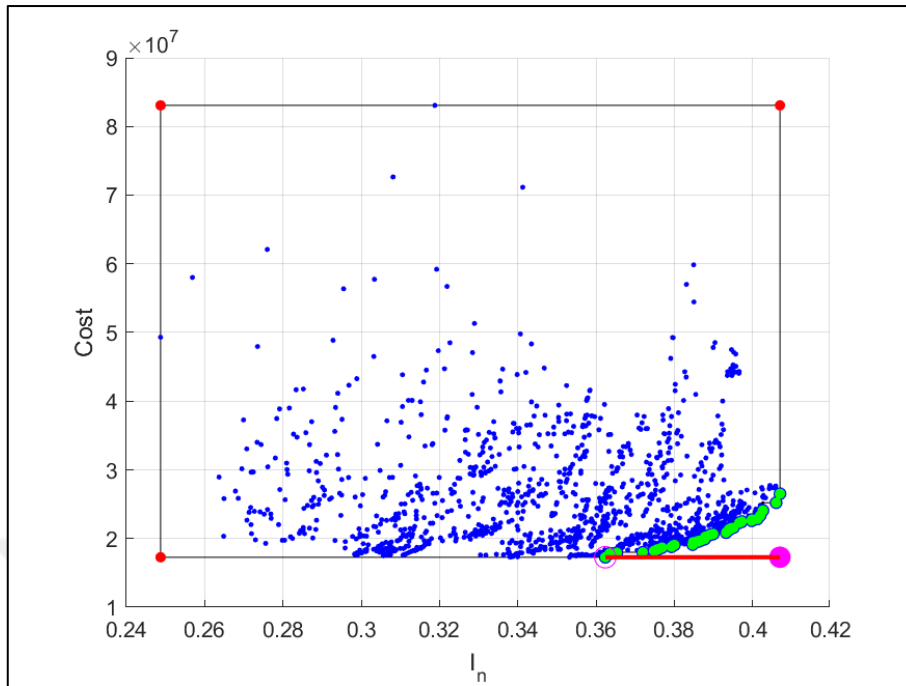


Figure B10. Pareto Curve for HAN50P3T

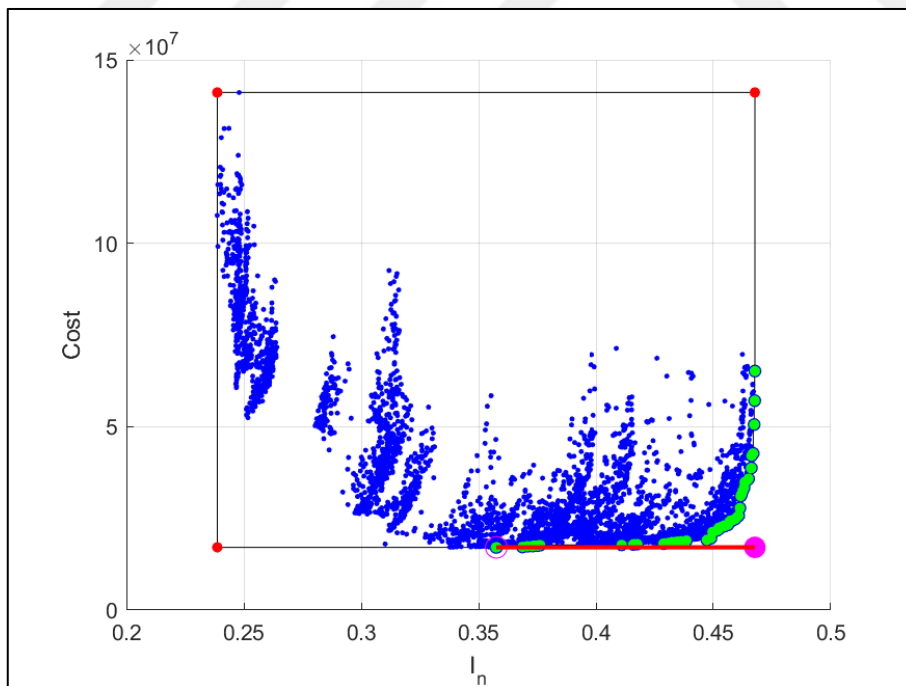


Figure B11. Pareto Curve for HAN100P3T

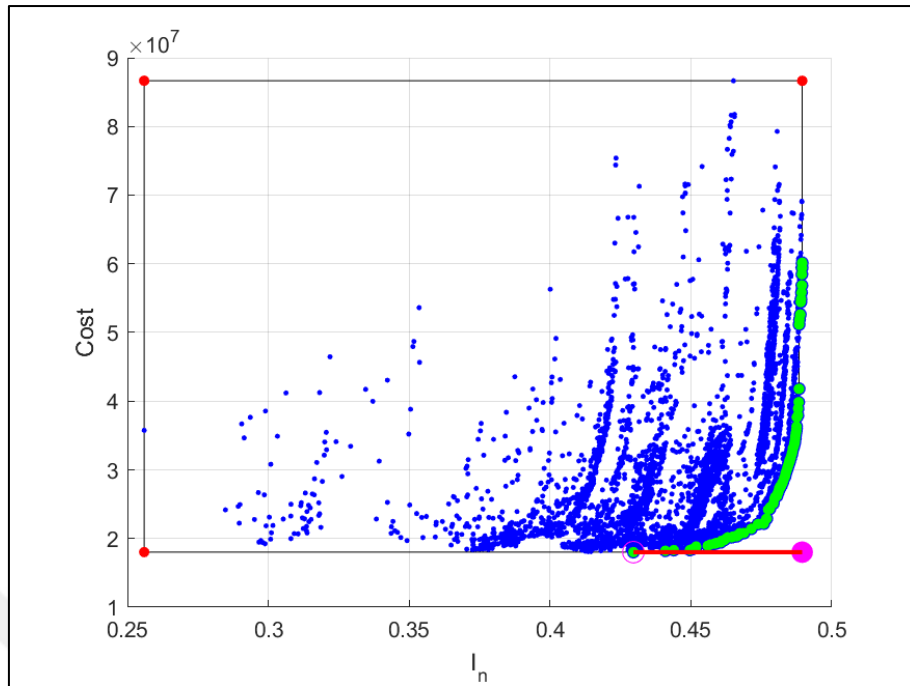


Figure B12. Pareto Curve for HAN150P3T

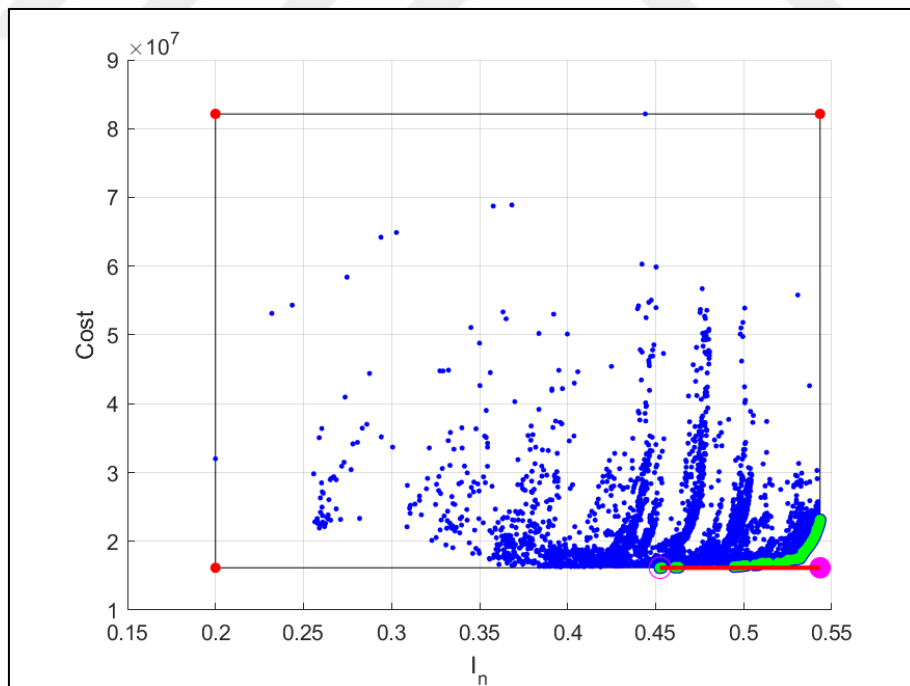


Figure B13. Pareto Curve for HAN200P3T

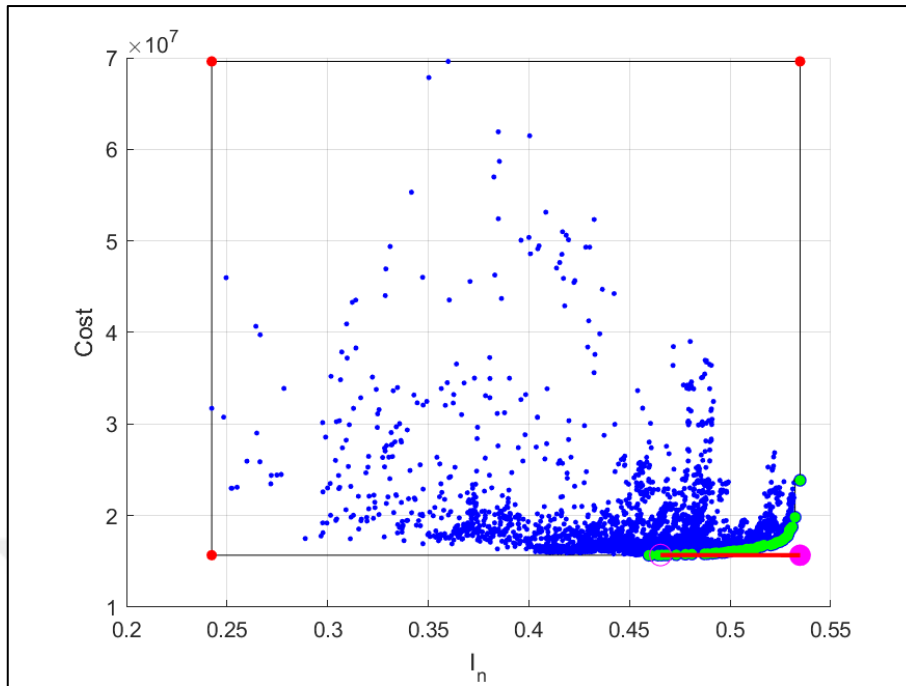


Figure B14. Pareto Curve for HAN250P3T

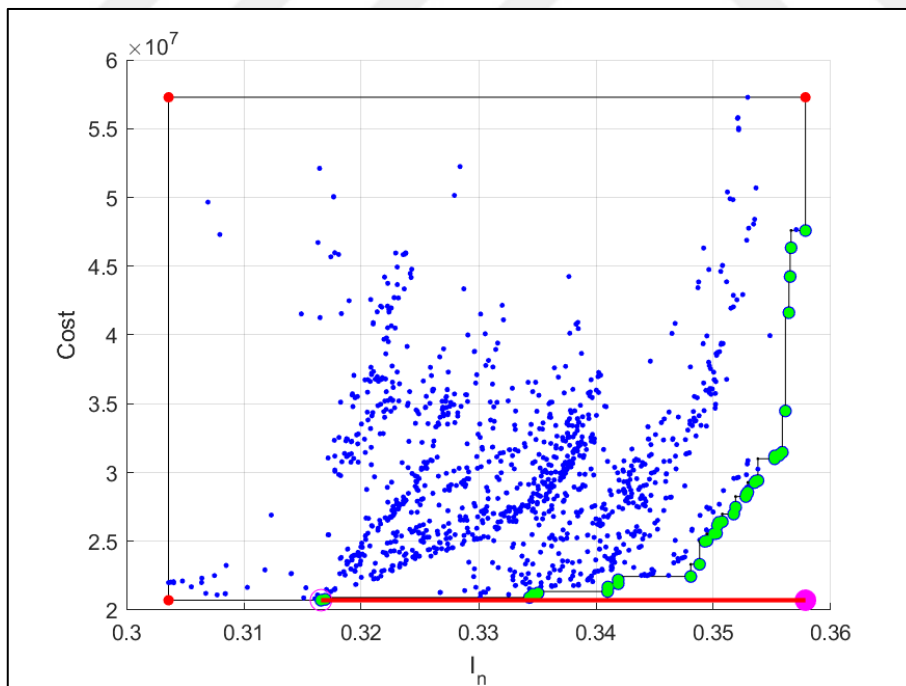


Figure B15. Pareto Curve for HAN50P4T

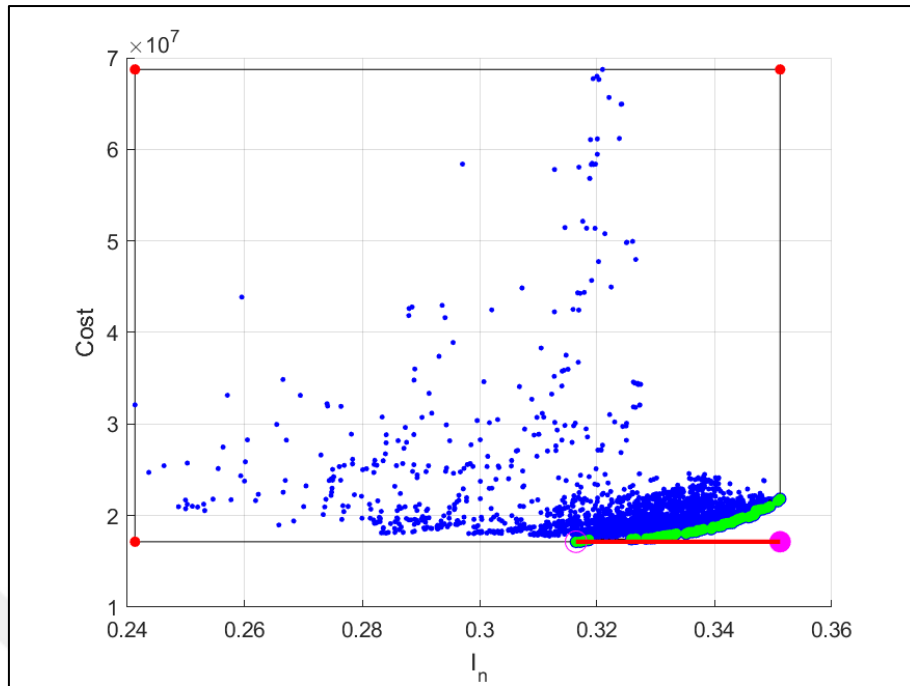


Figure B16. Pareto Curve for HAN100P4T

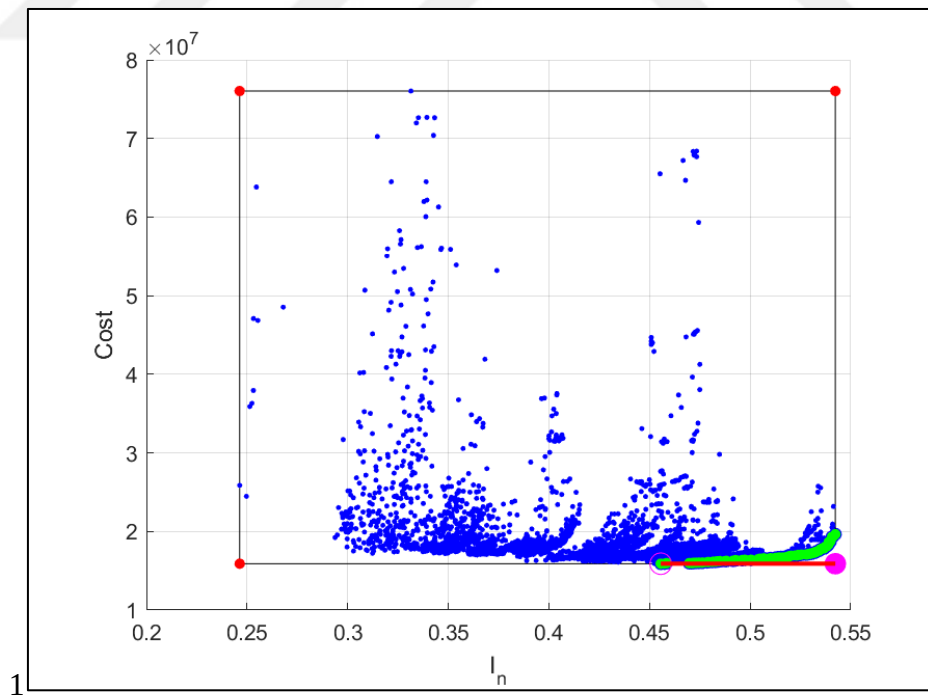


Figure B17. Pareto Curve for HAN150P4T

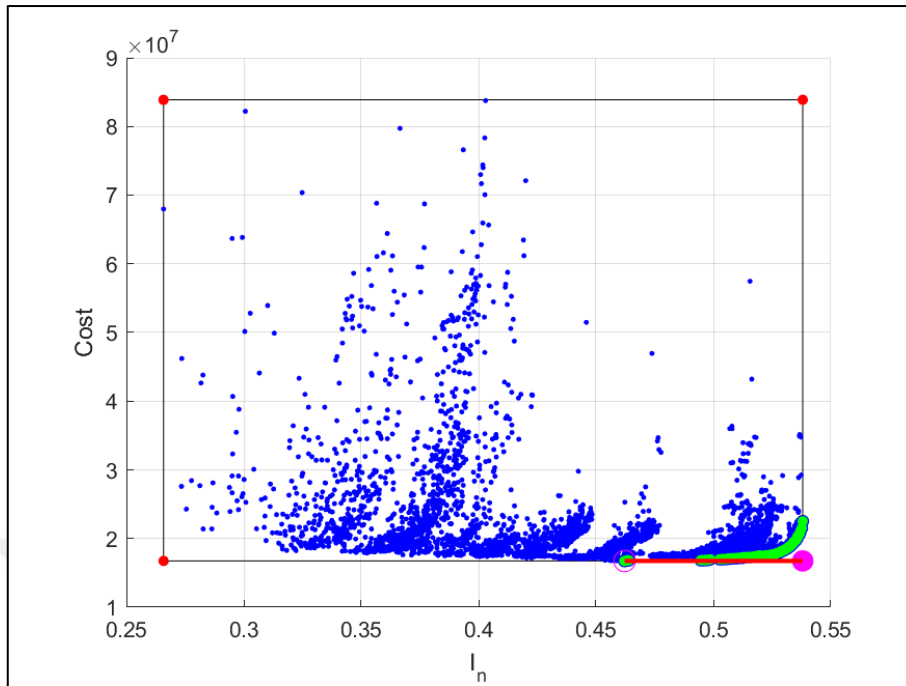


Figure B18. Pareto Curve for HAN200P4T

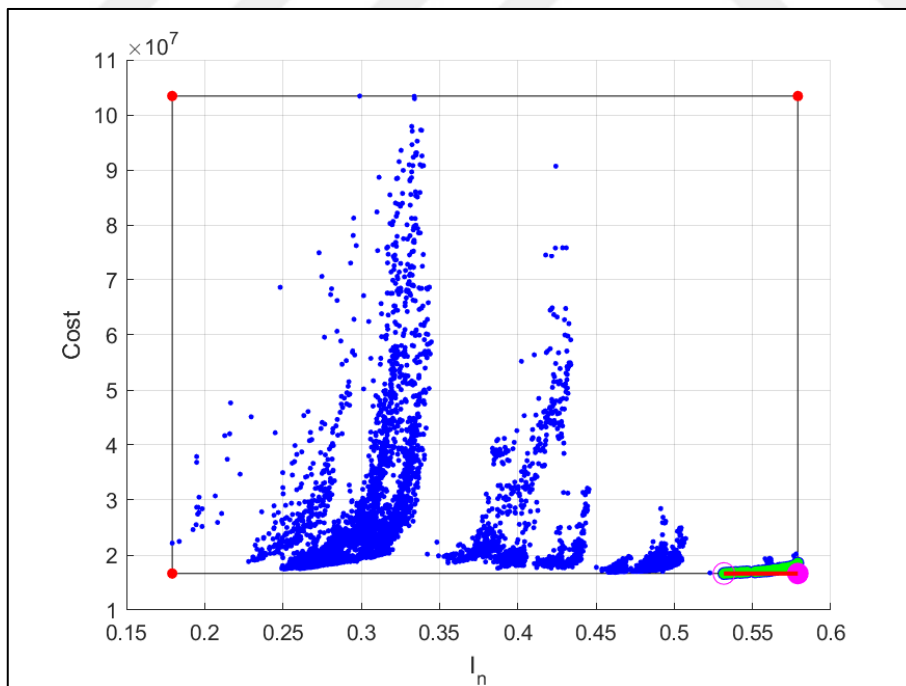


Figure B19. Pareto Curve for HAN250P4T

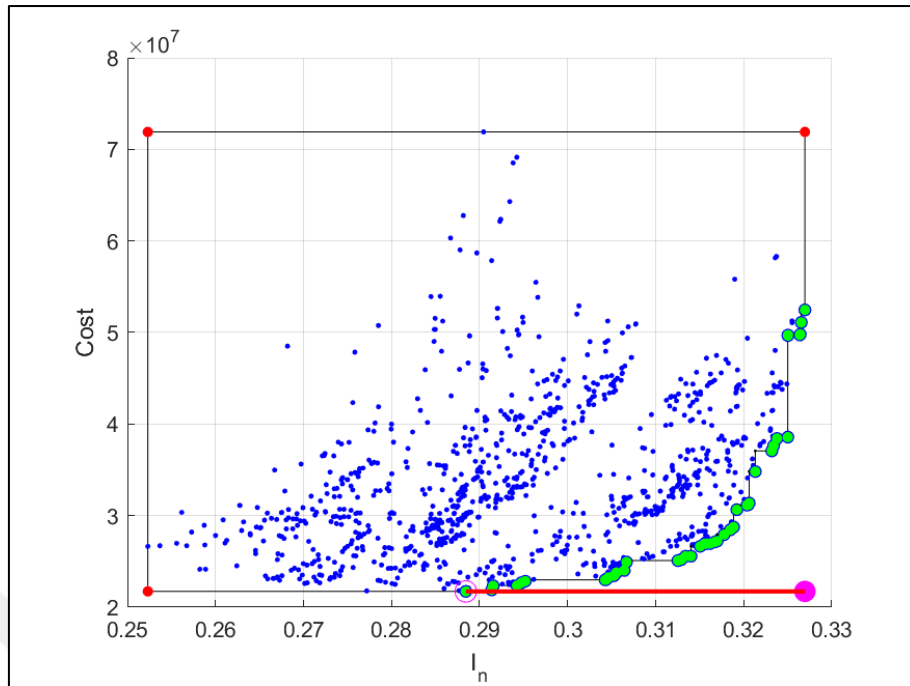


Figure B20. Pareto Curve for HAN50P5T

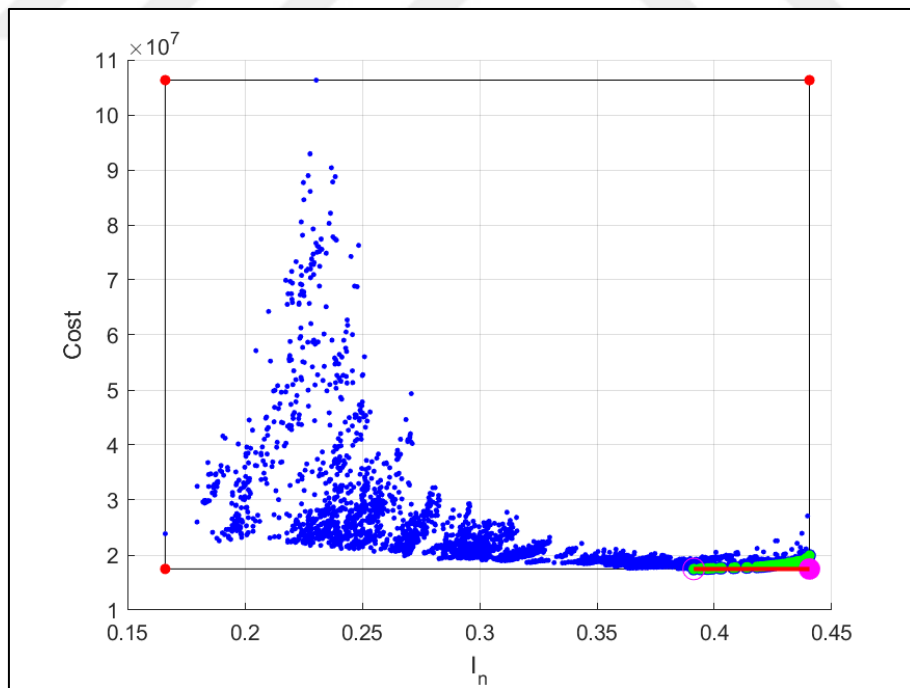


Figure B21. Pareto Curve for HAN100P5T

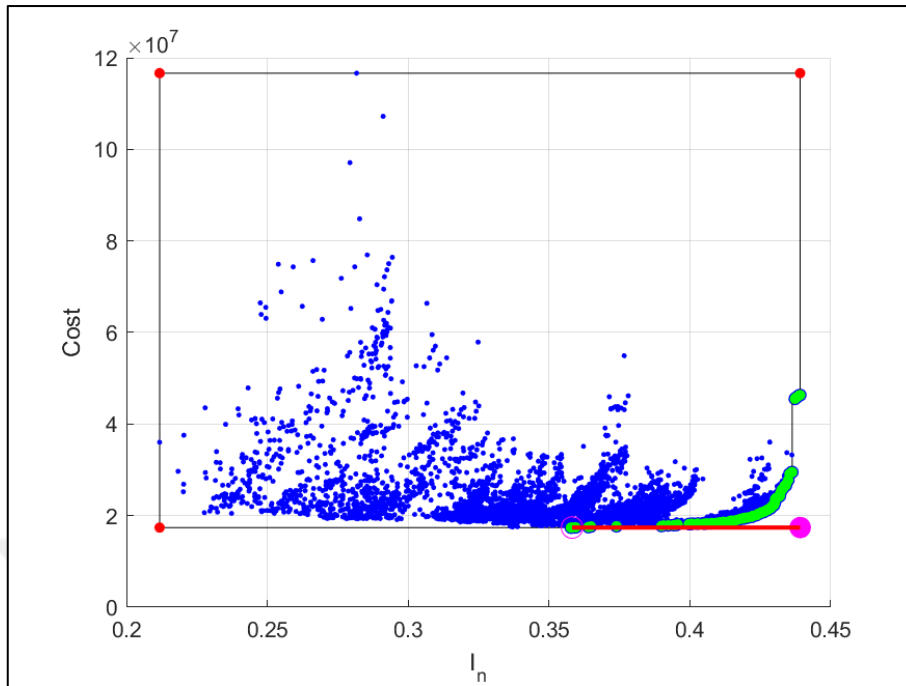


Figure B22. Pareto Curve for HAN150P5T

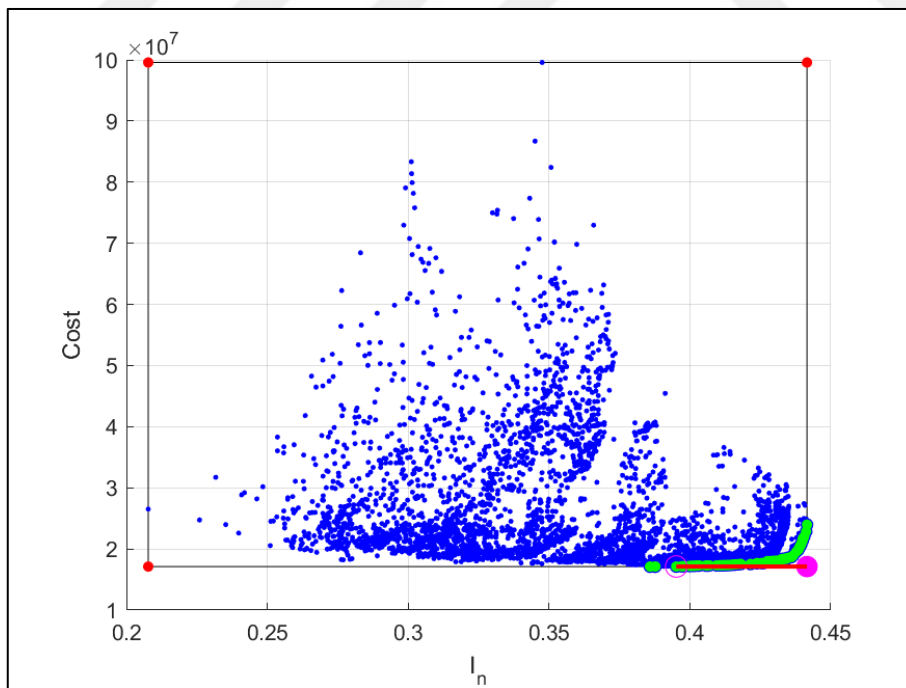


Figure B23. Pareto Curve for HAN200P5T

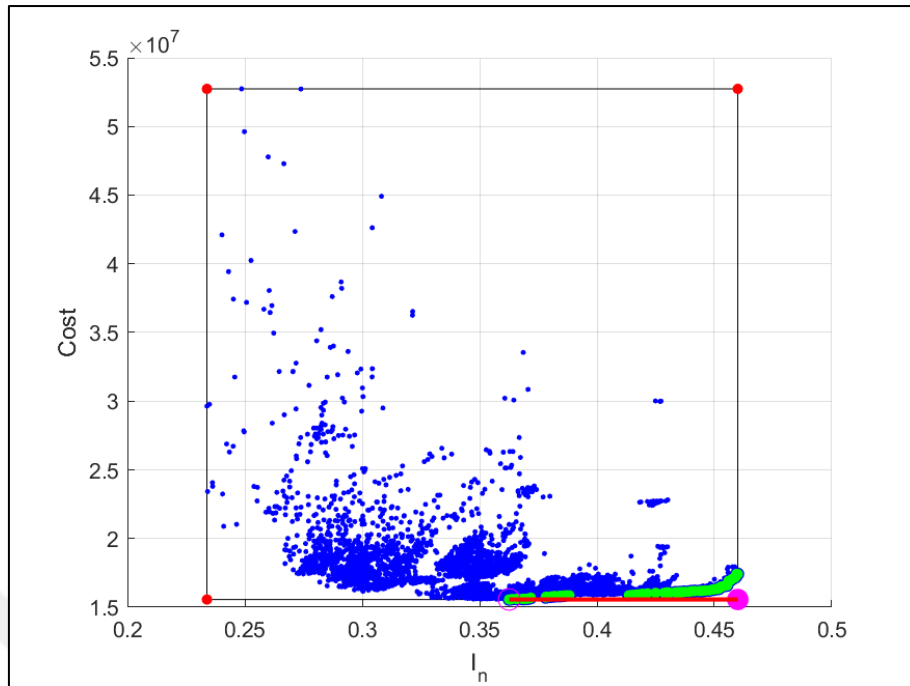


Figure B24. Pareto Curve for HAN250P5T

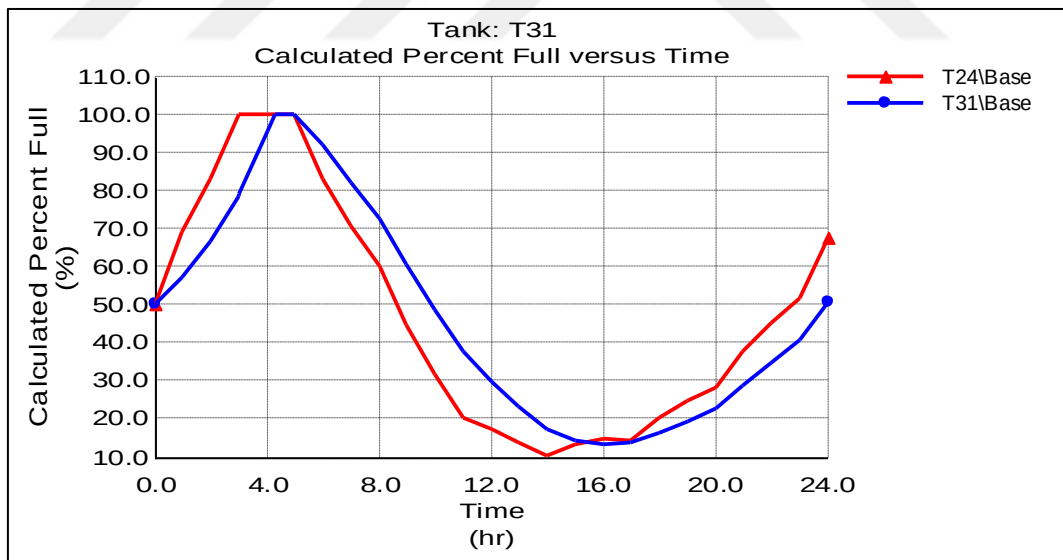


Figure B25. Tank Oscillations for Solution HAN200P2T/75

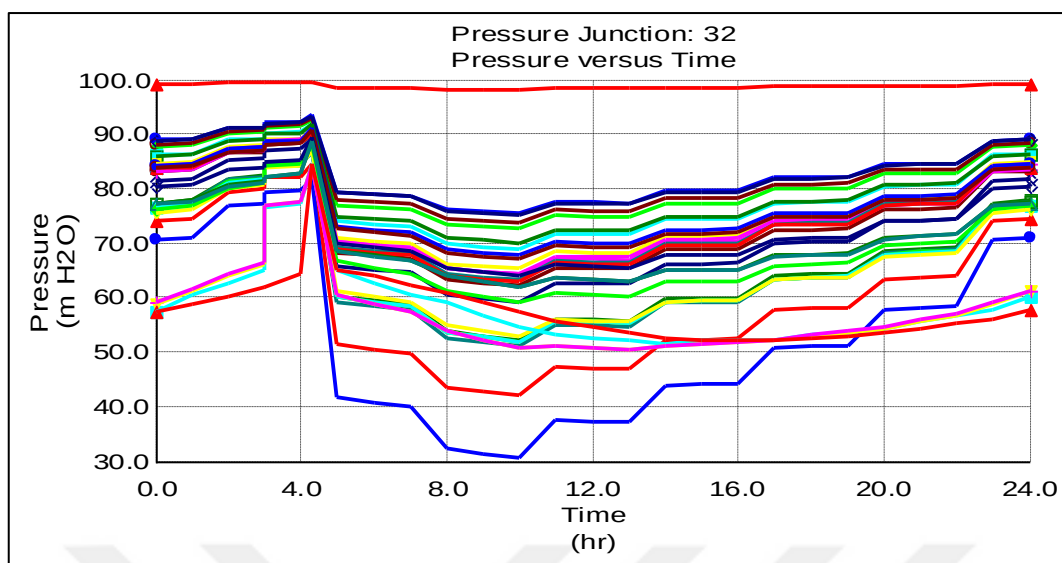


Figure B26. Pressure Variations for Solution HAN200P2T/75

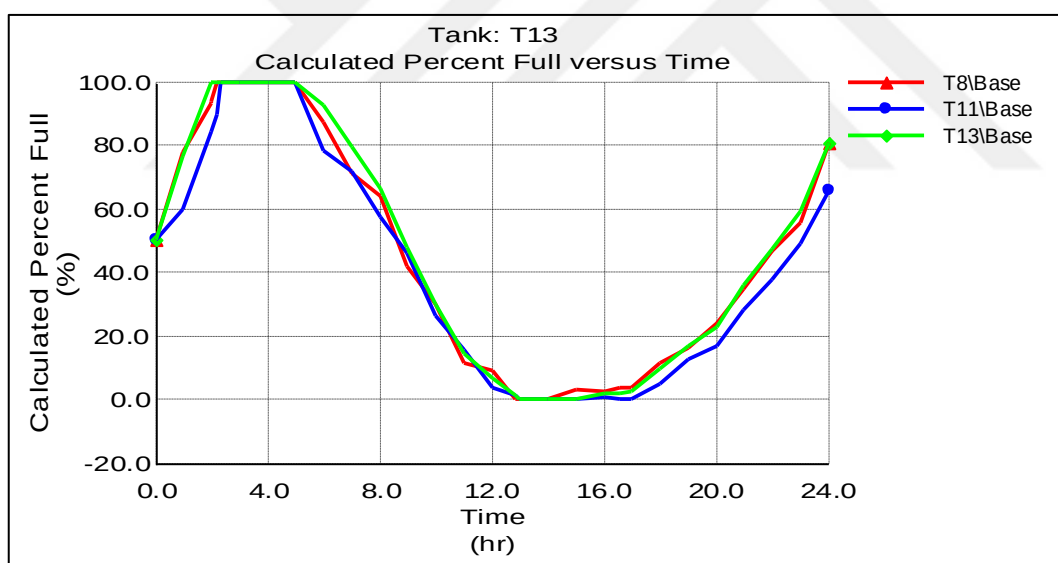


Figure B27. Tank Oscillations for Solution HAN100P3T/123

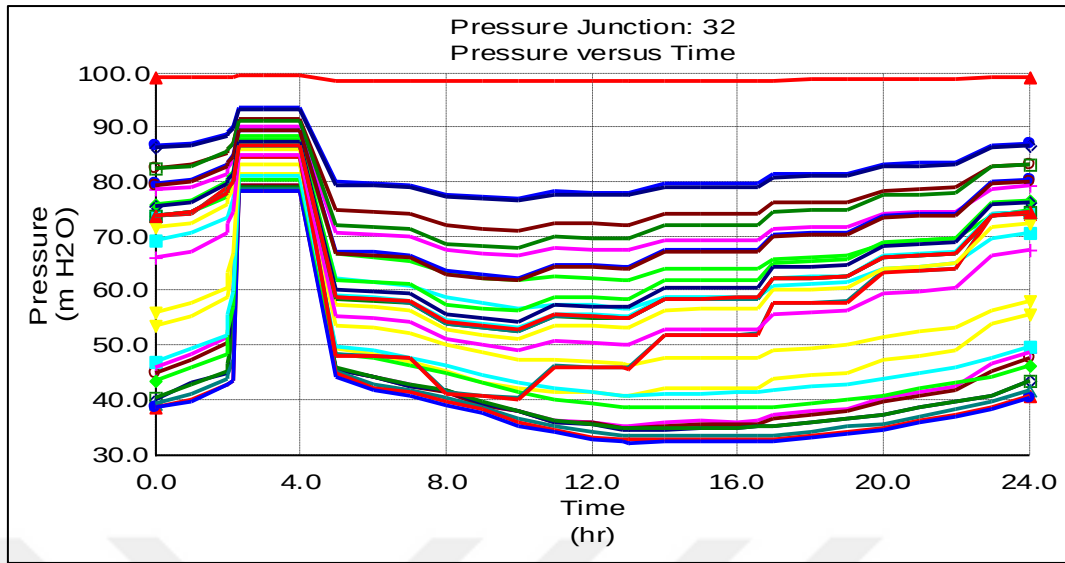


Figure B28. Pressure Variations for Solution HAN100P3T/123

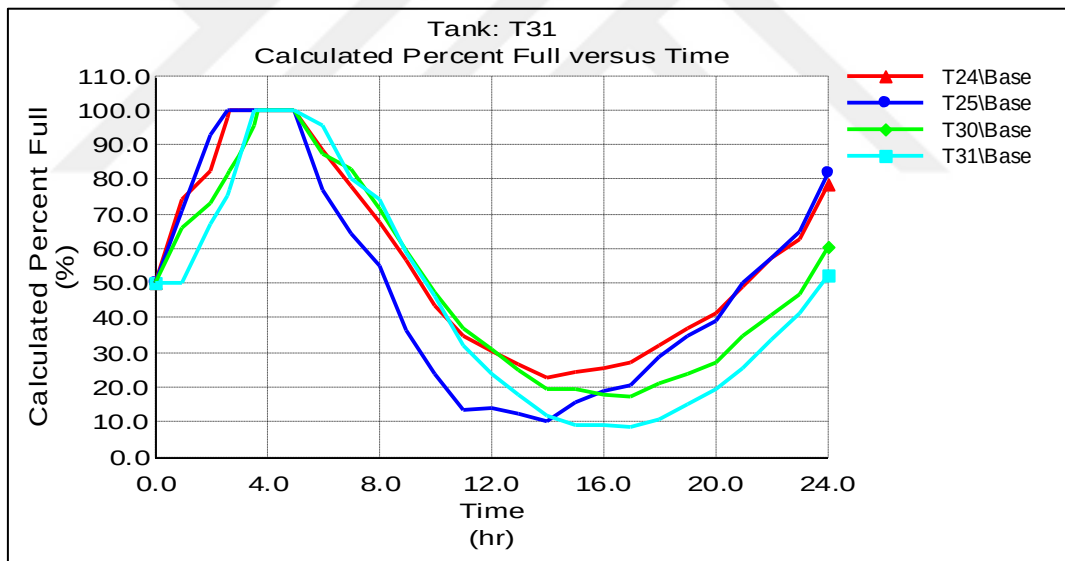


Figure B29. Tank Oscillations for Solution HAN250P4T/250

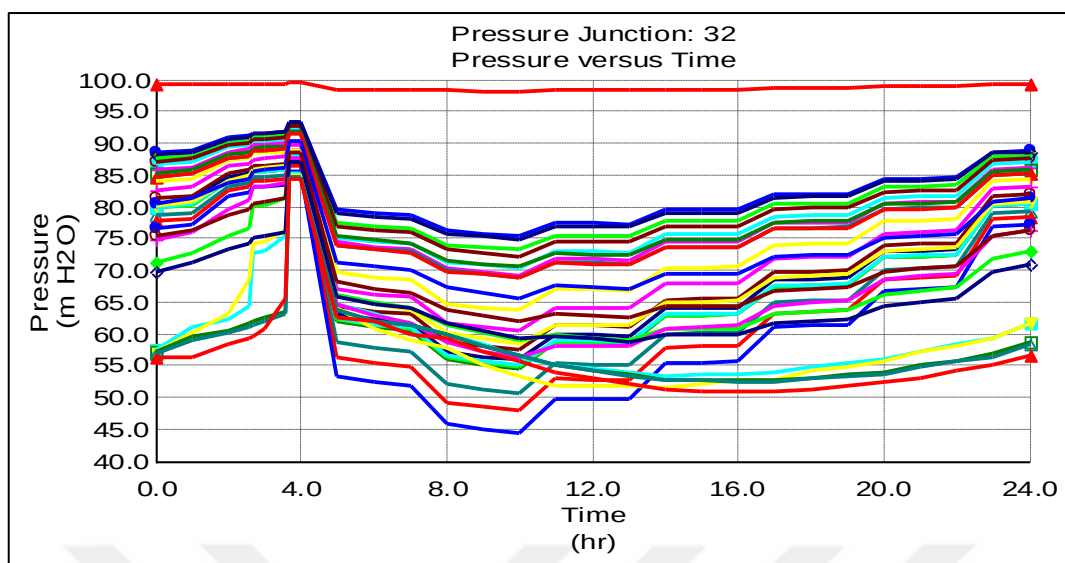


Figure B30. Pressure Variations for Solution HAN250P4T/250

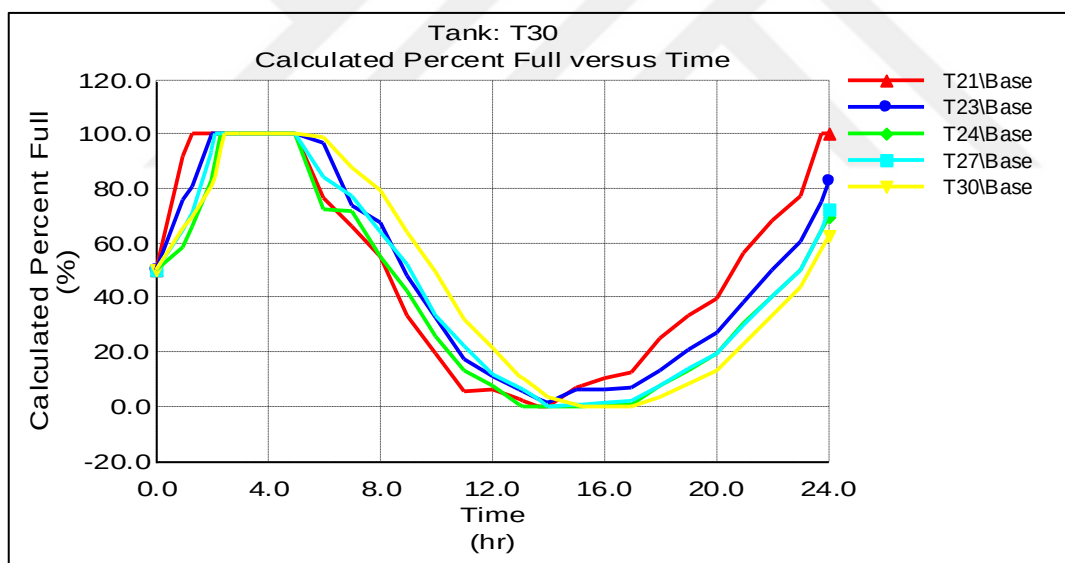


Figure B31. Tank Oscillations for Solution HAN100P5T/100

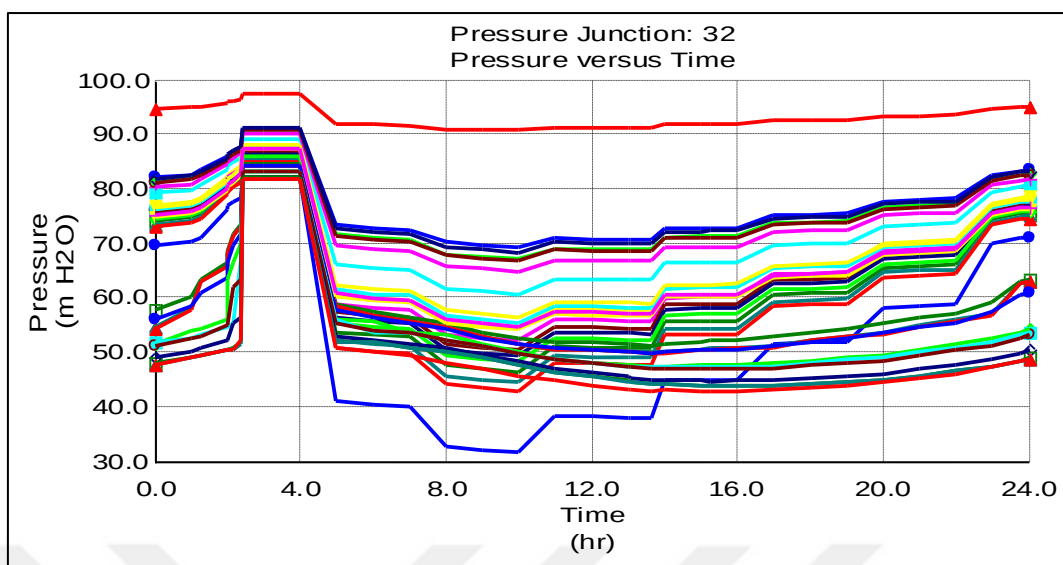


Figure B32. Pressure Variations for Solution HAN100P5T/100

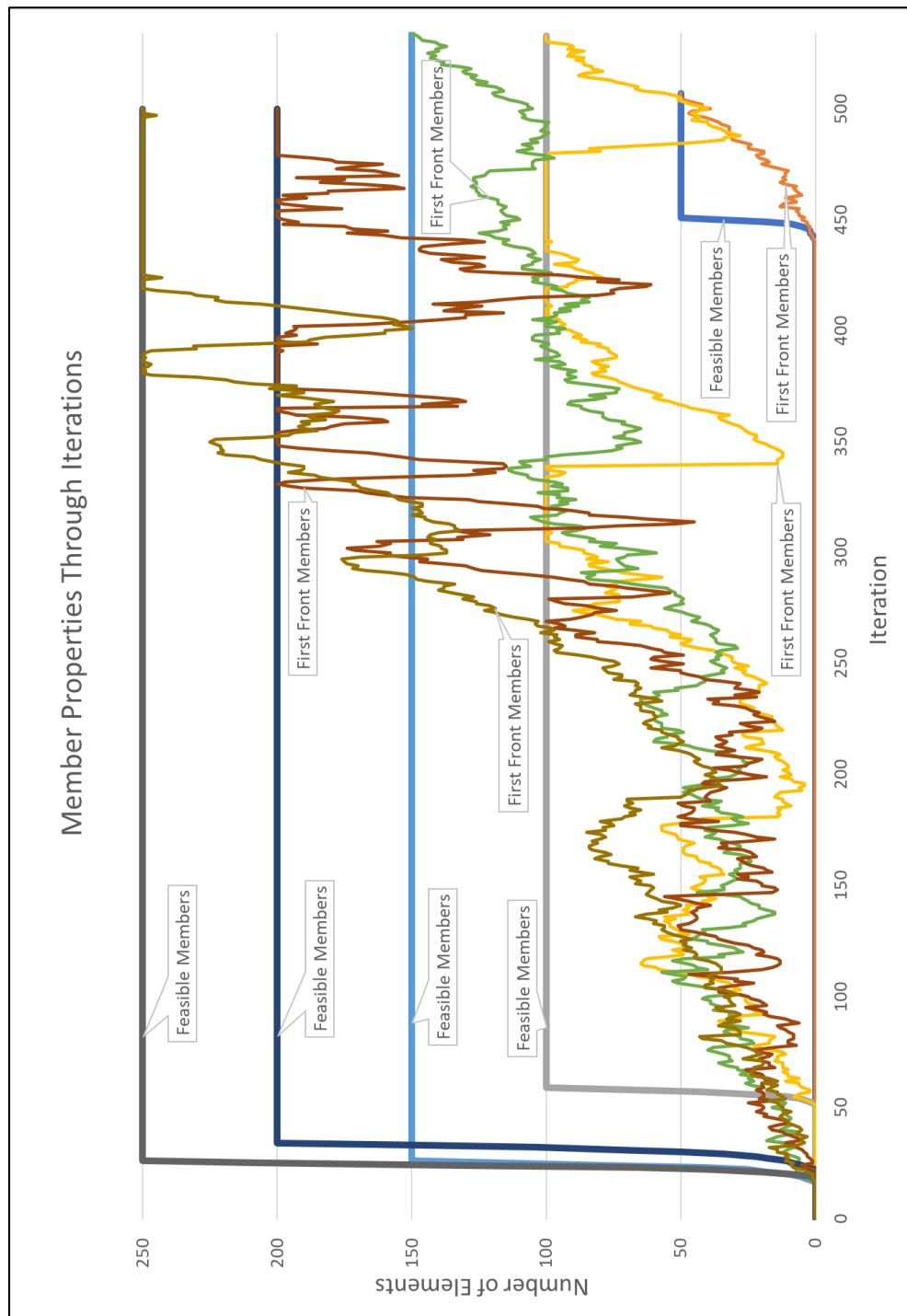


Figure B33. Member Properties for HAN2T

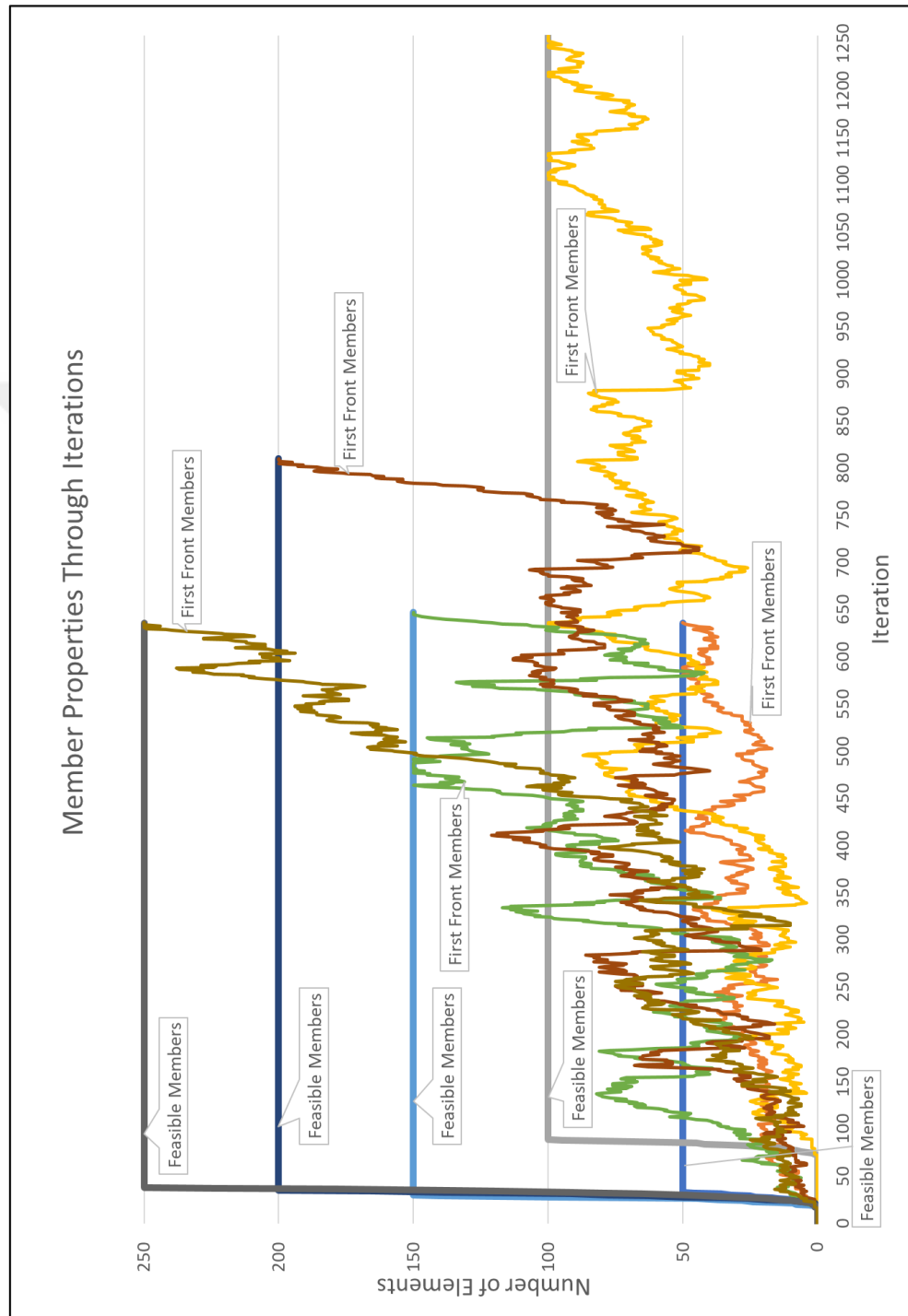


Figure B34. Member Properties for HAN3T

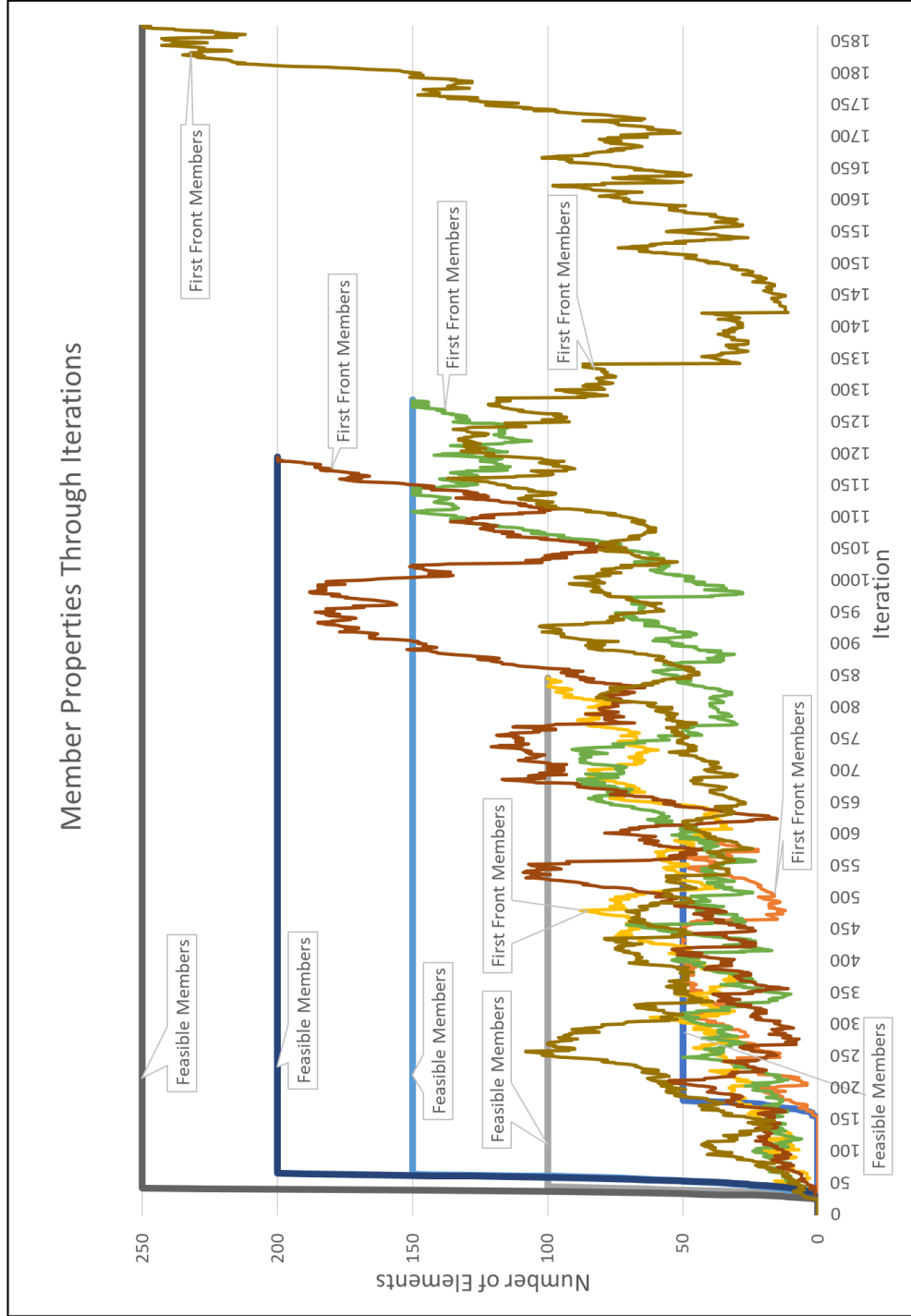


Figure B35. Member Properties for HAN4T

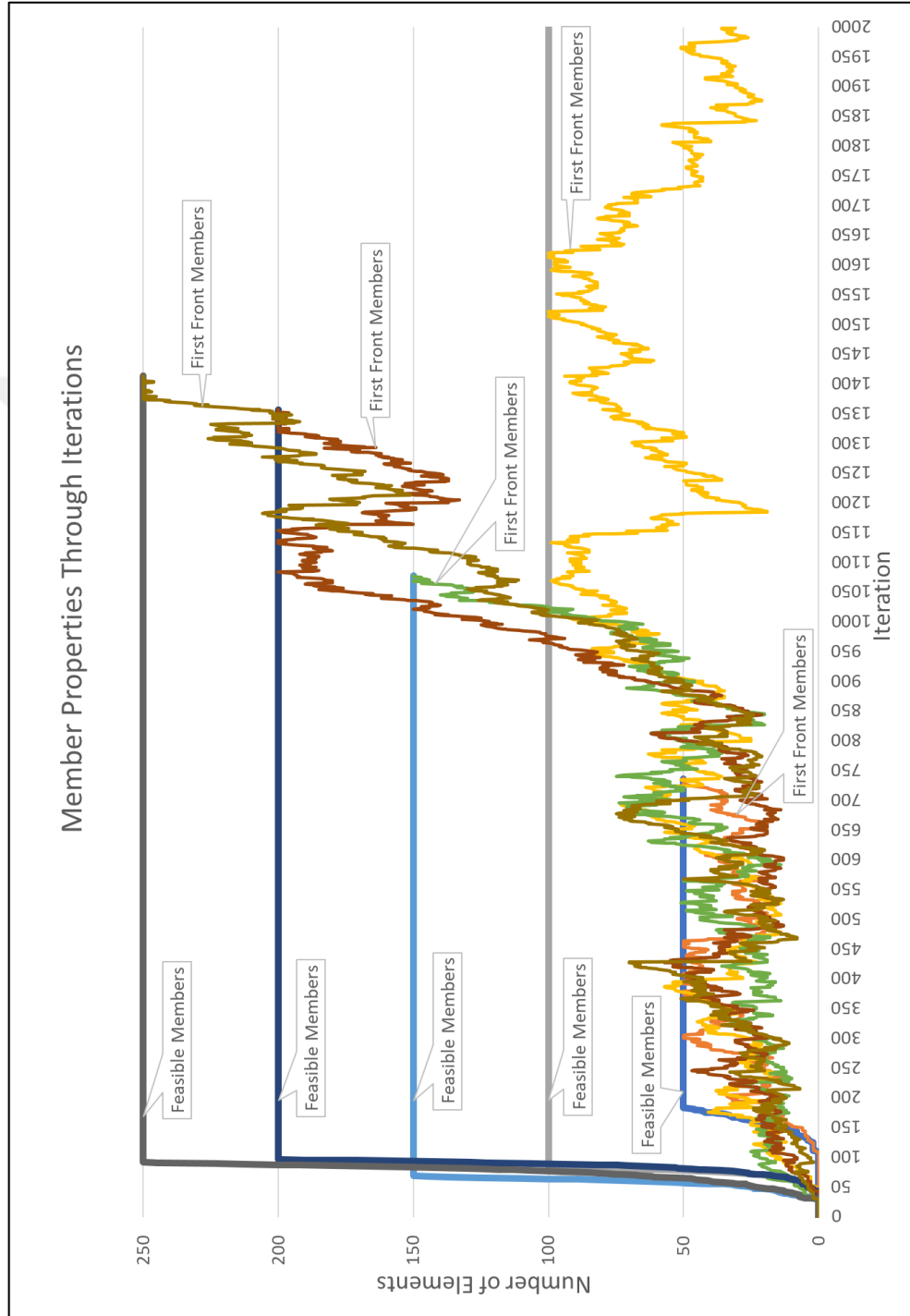


Figure B36. Member Properties for HAN5T

C. APPENDIX FOR N83 SOLUTIONS

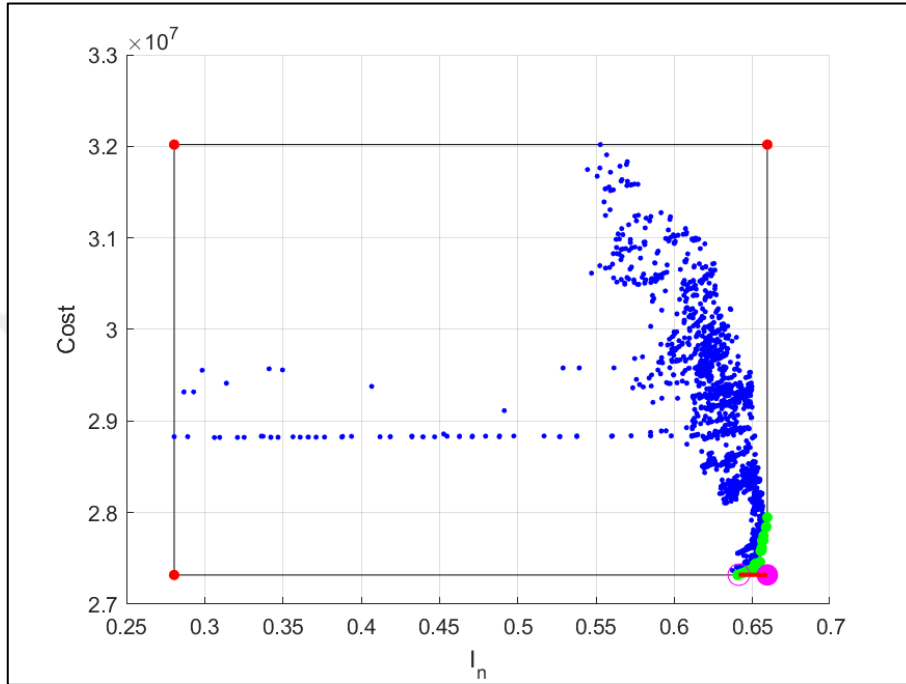


Figure C1. Pareto Curve for N8350P1T

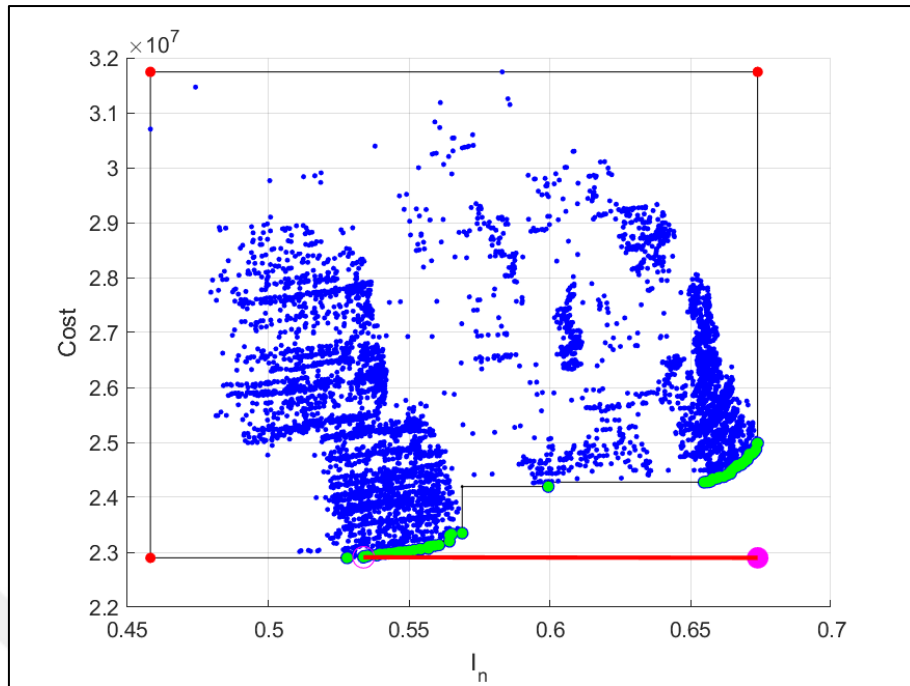


Figure C2. Pareto Curve for N83100P1T

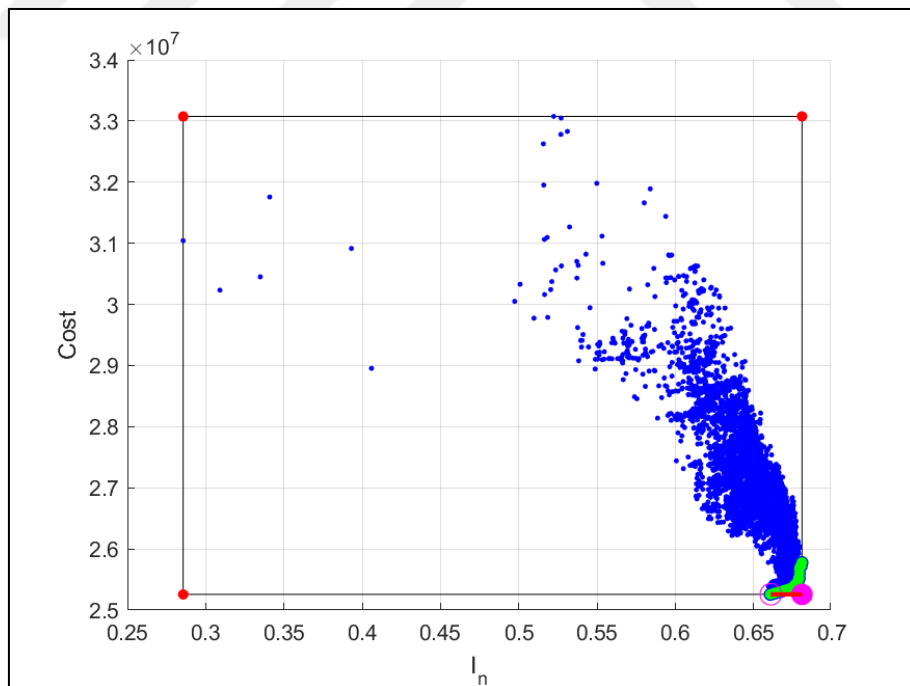


Figure C3. Pareto Curve for N83150P1T

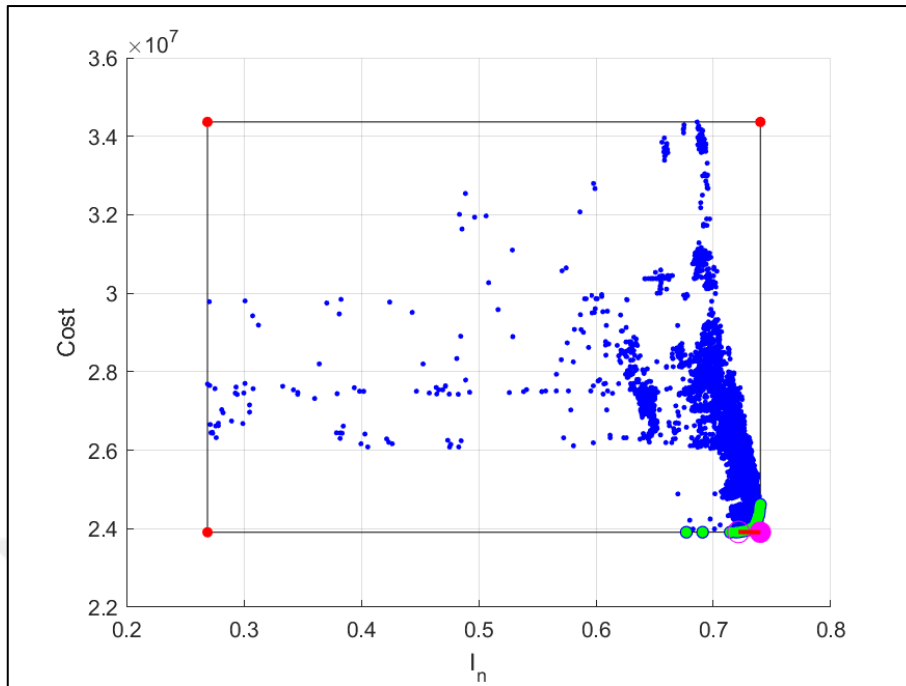


Figure C4. Pareto Curve for N83200P1T

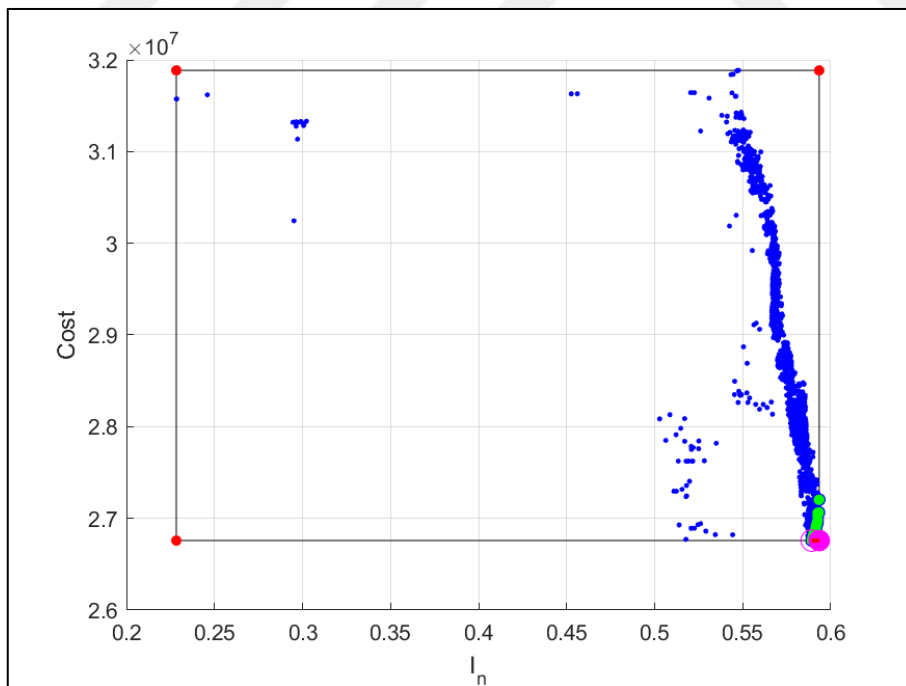


Figure C5. Pareto Curve for N8350P2T

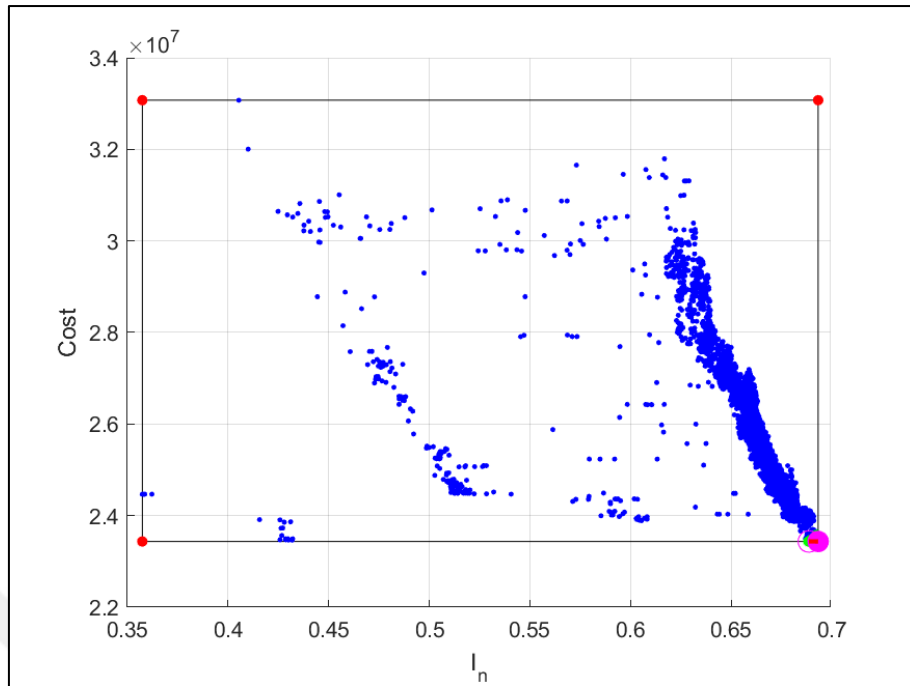


Figure C6. Pareto Curve for N83100P2T

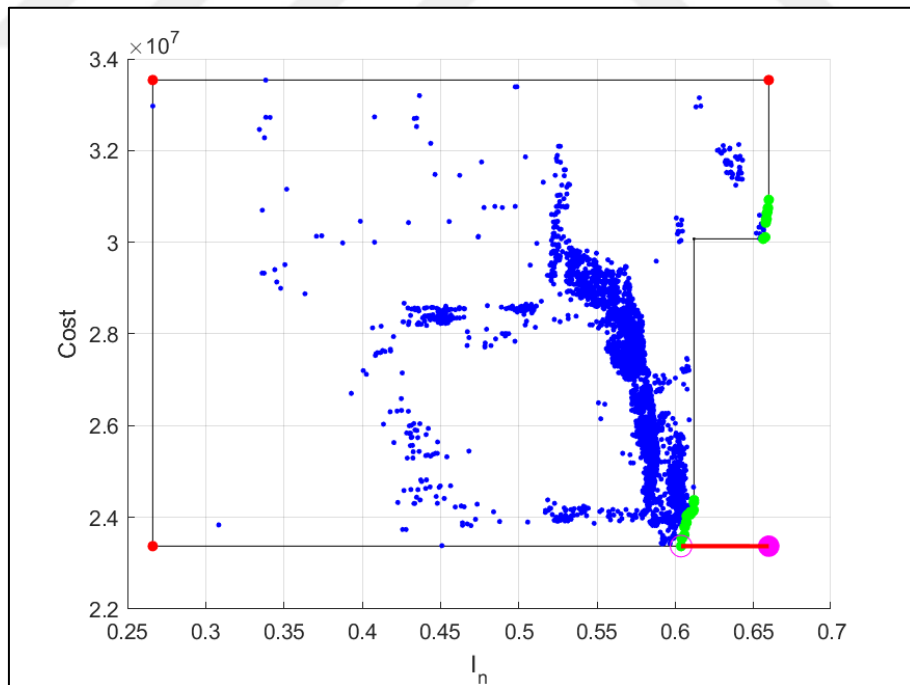


Figure C7. Pareto Curve for N83150P2T

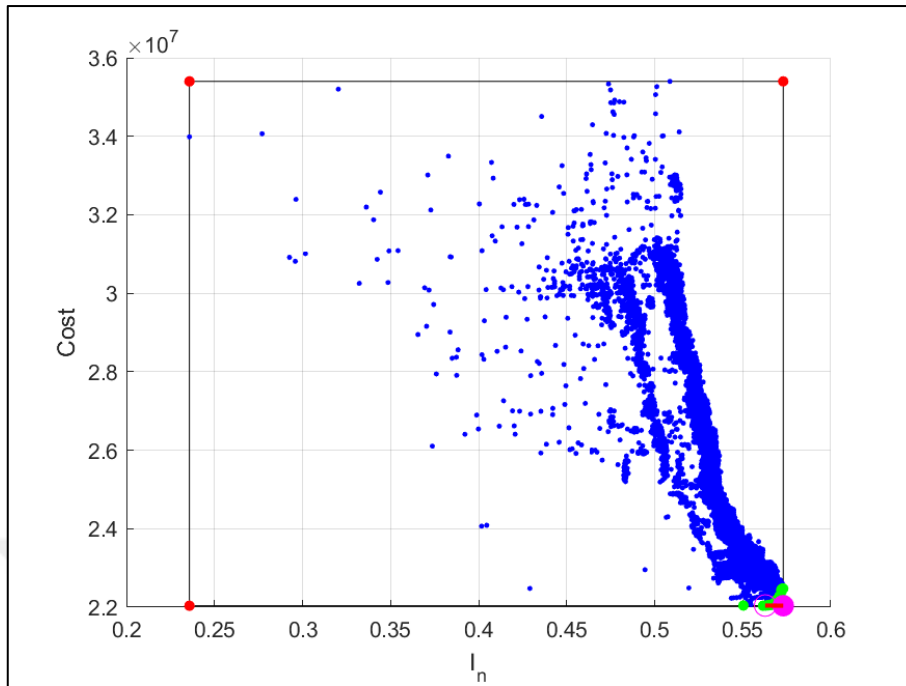


Figure C8. Pareto Curve for N83200P2T

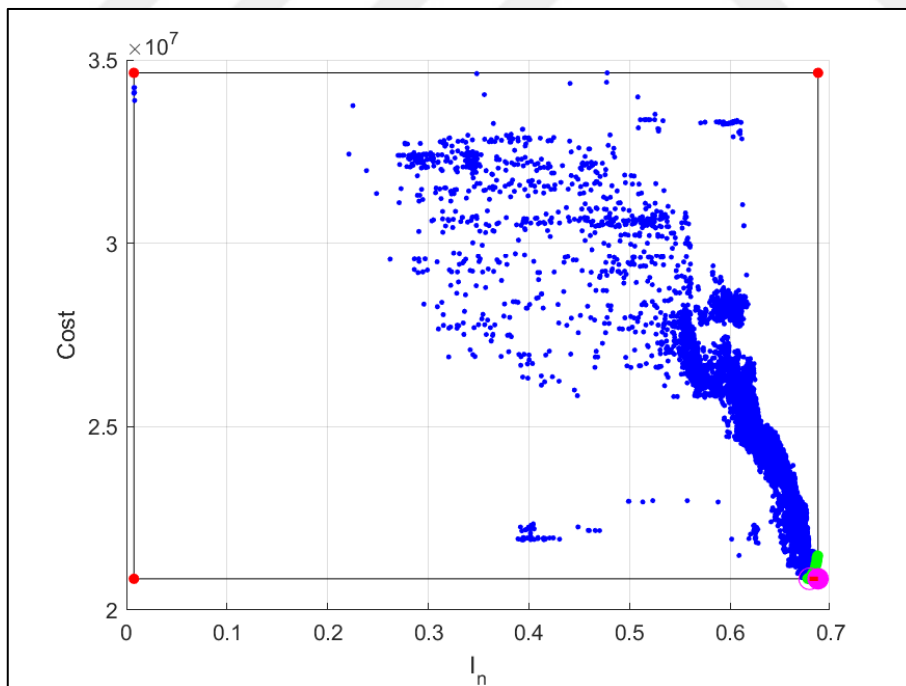


Figure C9. Pareto Curve for N83250P2T

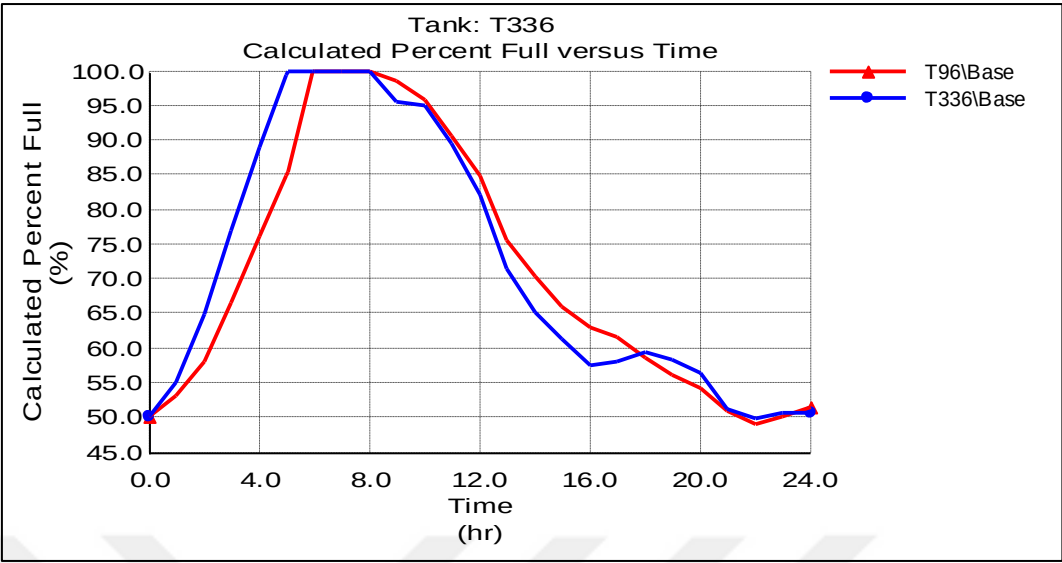


Figure C10. Tank Oscillations for Solution N83250P2T/39

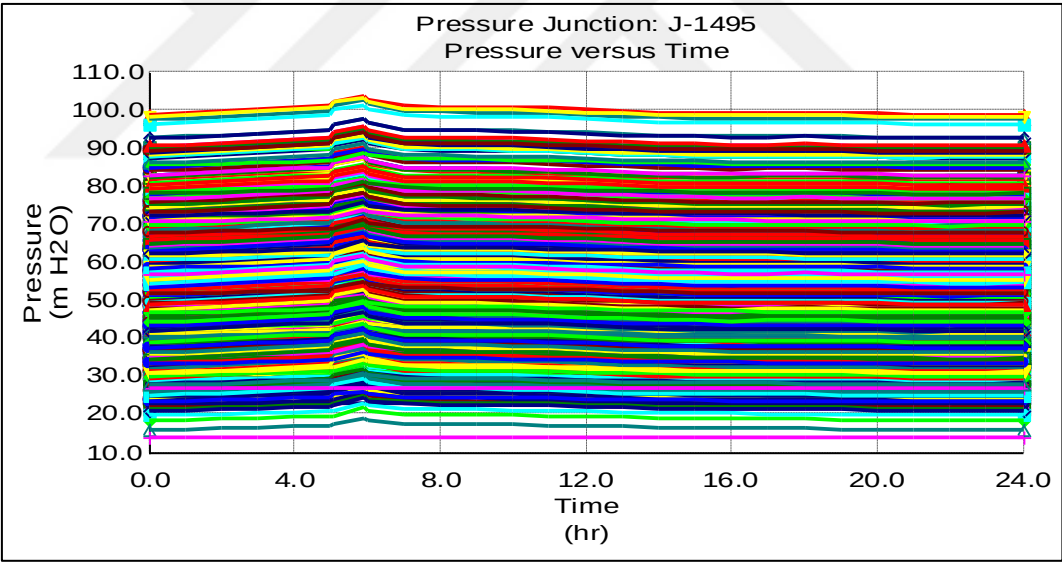


Figure C11. Pressure Variations for Solution N83250P2T/39

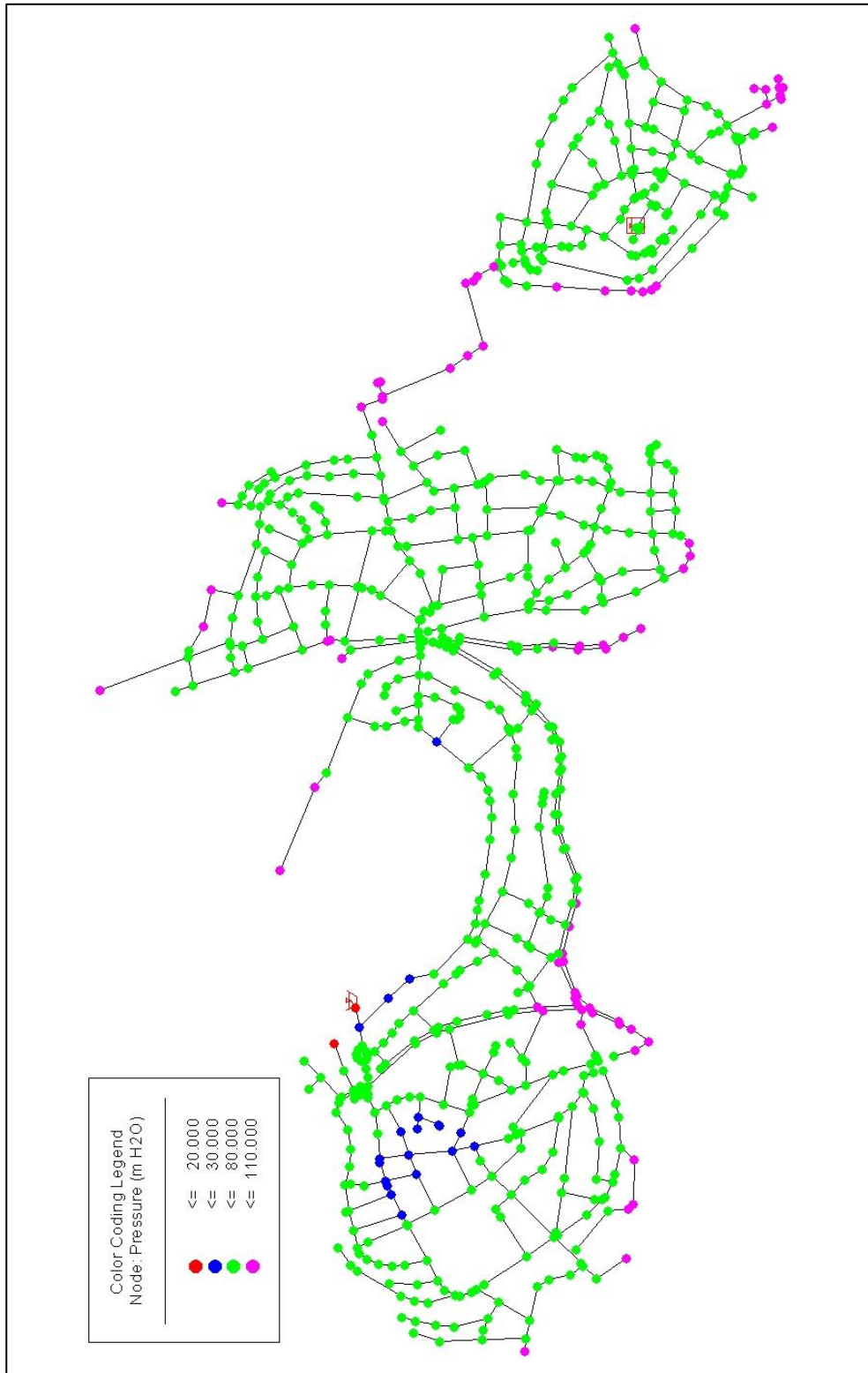


Figure C12. Pressure Distribution for N83250P1T/250 at 03:00

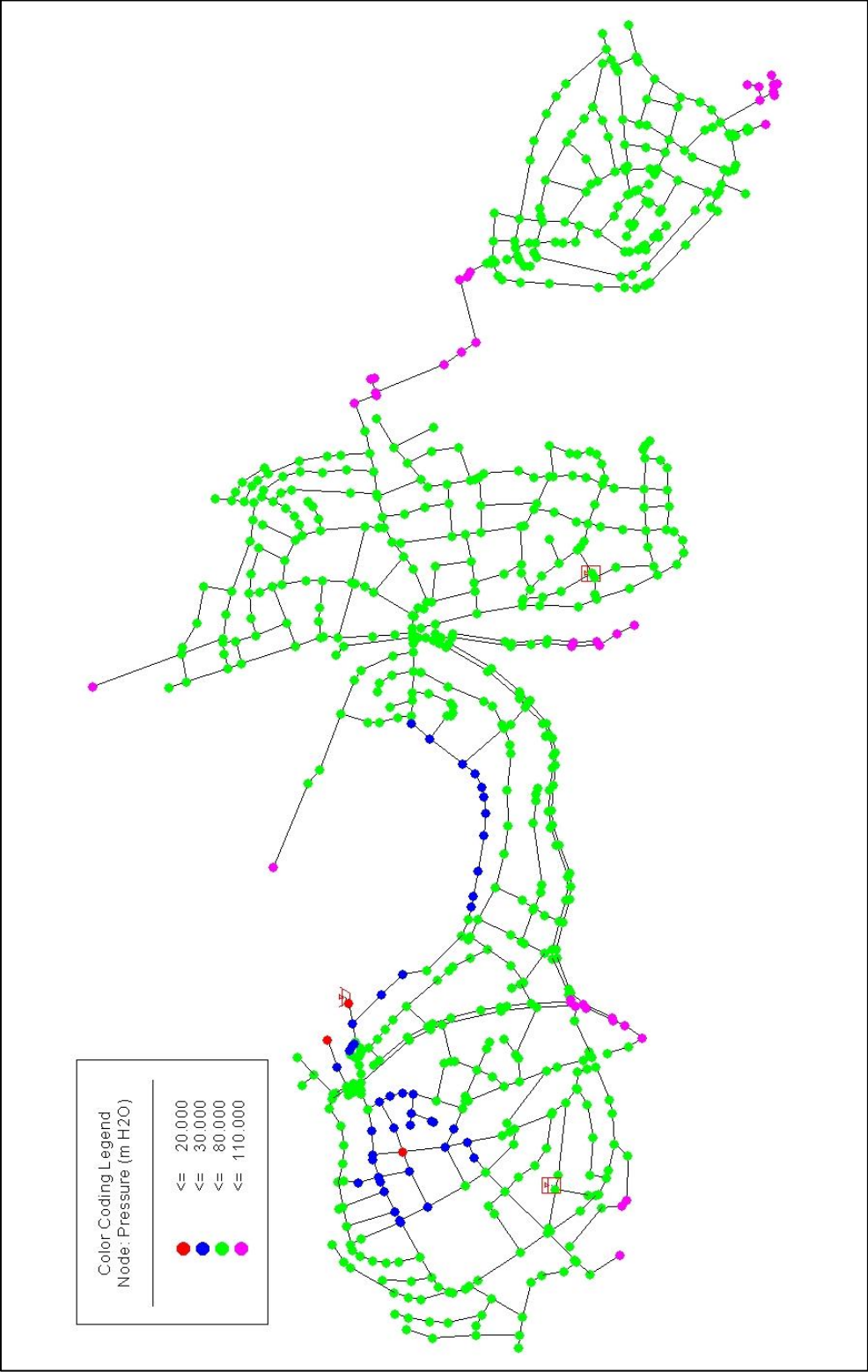


Figure C13. Pressure Distribution for N83250P2T/39 at 03:00

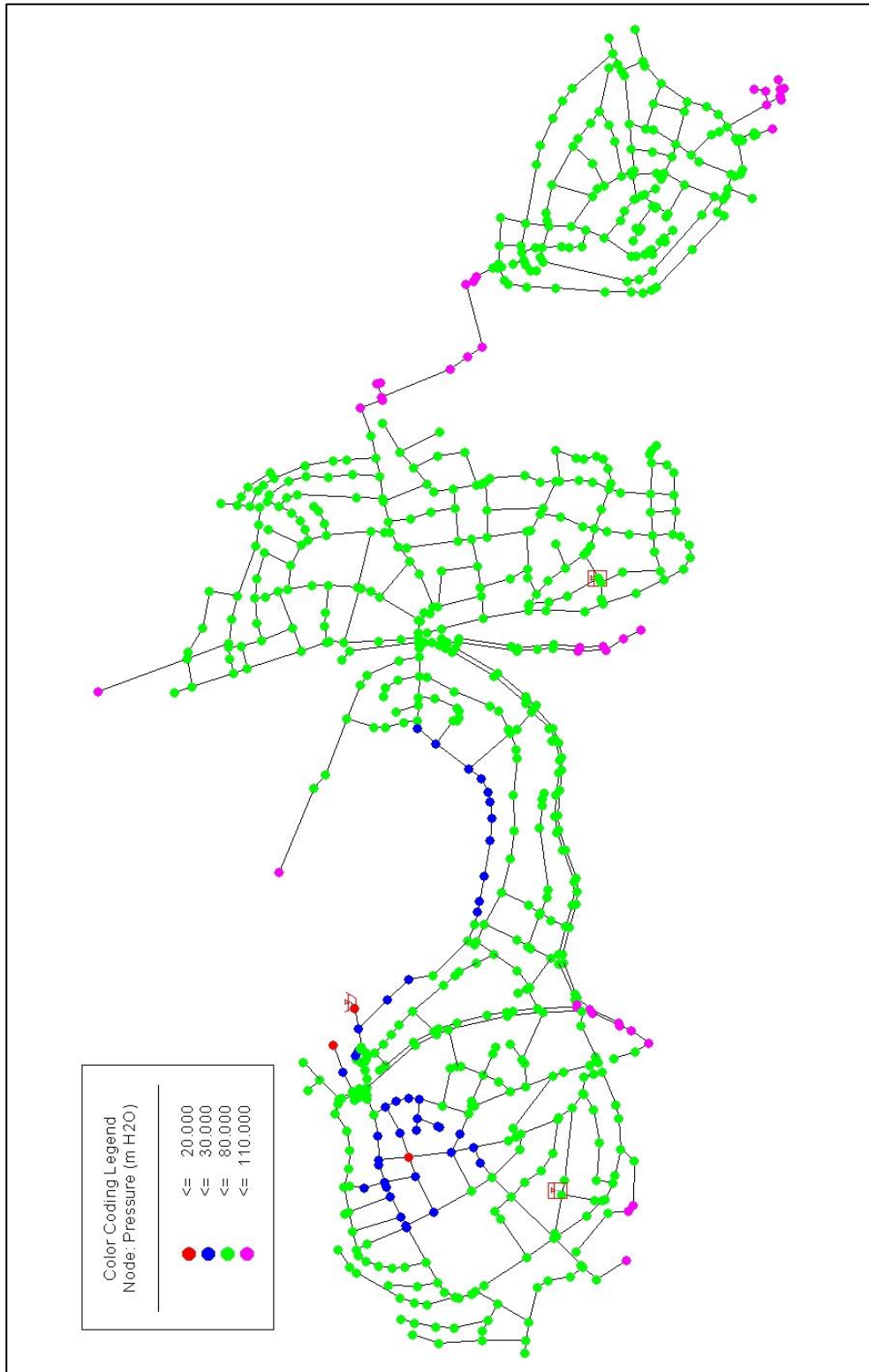


Figure C14. Pressure Distribution for N83250P2T/39 at 12:00

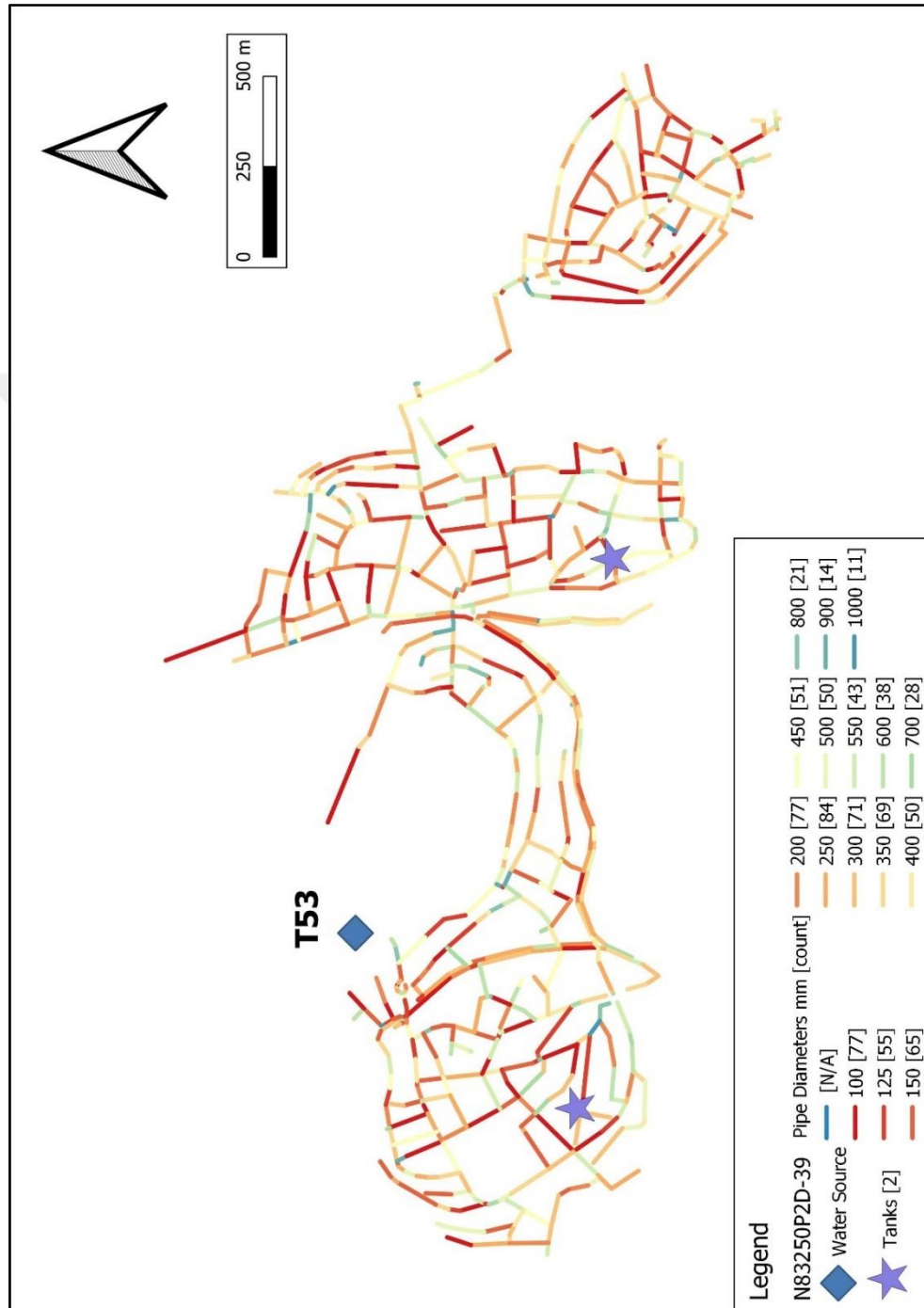


Figure C15. Pipe Diameters for N83250P2T/39

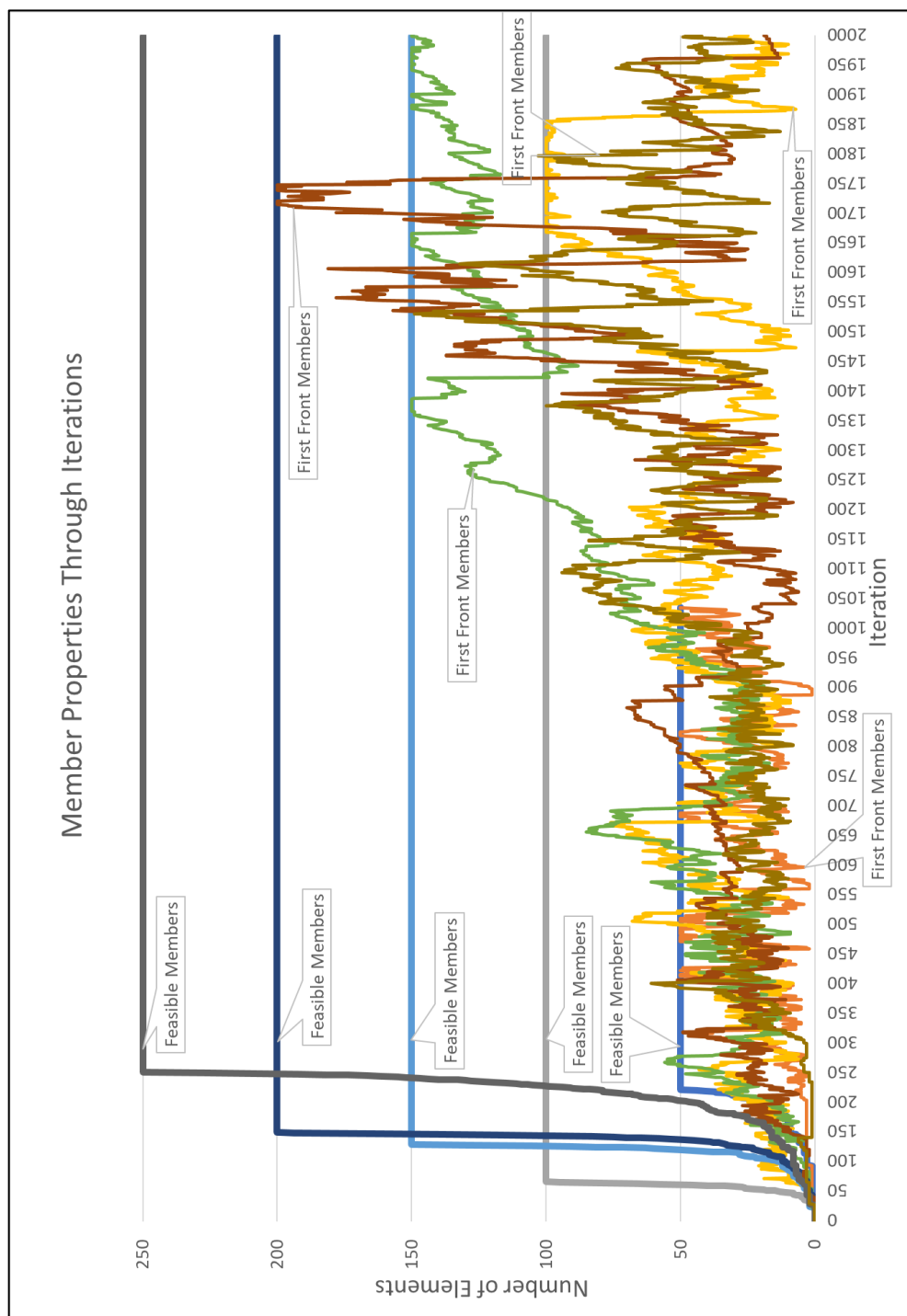


Figure C16. Member Properties for N832T

CURRICULUM VITAE

PERSONAL INFORMATION

Surname, Name: Bektaş, Hayrettin Onur

Nationality: Turkish (TC)

Date and Place of Birth: 6 January 1984, Ankara

Marital Status: Married

Phone: +90 536 276 73 85

E-mail: onur.bektas@metu.edu.tr

EDUCATION

Degree	Institution	Year of Graduation
MSc	METU Geodetical and Geographical Information Technologies	2010
BS	METU Civil Engineering	2007
High School	TED Ankara College	2002

WORK EXPERIENCE

Year	Place	Enrolment
2020 – ...	Ankara Water and Sewerage Administration (ASKİ)	Civil Engineer
2012 – 2020	Alter International Engineering and Consulting Services	Civil Engineer

2011 – 2012	Jandarma İstihkam İnşaat Grup	Third-Lieutenant Civil Engineer
	Komutanlığı	
2007 – 2011	Ankara Water and Sewerage Administration (ASKİ)	Civil Engineer

FOREIGN LANGUAGES

Advanced English, Basic French

HOBBIES

Ankara urban development and water history, old aerial photos and maps, Turkish “Yeşilçam” movie history, Hitchcock cinematography, music history