A DYNAMIC CAPACITATED FACILITY LOCATION AND NETWORK DESIGN PROBLEM WITH CAPACITY EXPANSION FOR WASTEWATER TREATMENT NETWORKS

by

Seyda Alperen Pehlivan

Submitted to Graduate School of Natural and Applied Sciences

in Partial Fulfillment of the Requirements

for the Degree of Master of Science in

Industrial and System Engineering

Yeditepe University

2022

A DYNAMIC CAPACITATED FACILITY LOCATION AND NETWORK DESIGN PROBLEM WITH CAPACITY EXPANSION FOR WASTEWATER TREATMENT NETWORKS

APPROVED BY:

Prof. Dr. Melek Başak ....
(Thesis Supervisor)
(Yeditepe University)

Assist. Prof. Dr. Duygun Fatih Demirel ....
(Thesis Co-Supervisor)
(İstanbul Kültür University)

Assist. Prof. Dr. Semih Yalçındağ ....
(Yeditepe University)

Prof. Dr. Temel Öncan ....
(Galatasaray University)

DATE OF APPROVAL: ..... / ..... / 20.....

I hereby declare that this thesis is my own work and that all information in this thesis has been obtained and presented in accordance with academic rules and ethical conduct. I have fully cited and referenced all material and results as required by these rules and conduct, and this thesis study does not contain any plagiarism. If any material used in the thesis requires copyright, the necessary permissions have been obtained. No material from this thesis has been used for the award of another degree.

I accept all kinds of legal liability that may arise in case contrary to these situations.

Name, Last Name Seyda Alperen Pehlivan

Signature ....

ACKNOWLEDGEMENTS

It is a pleasure to thank the people who made this thesis possible. First, I would like to express my gratitude to my advisors Prof. Dr. Melek Başak and Assist. Prof. Dr. Duygun Fatih Demirel for their endless support, contributions, and patience. Their knowledge and experience were my greatest guide during both my thesis and my research assistant journey.

I would like to thank Assist. Prof. Dr. Semih Yalçındağ and Assist. Prof. Dr. Eylül Damla Gönül Sezer for their support and contribution during the verification phase of my study.

I am very thankful to all of my colleagues and all the members of Yeditepe University Industrial Engineering Department for their support and friendly attitude.

Finally and for most, I would like to thank my precious family for their limitless love and support, which makes everything more beautiful.

ABSTRACT

A DYNAMIC CAPACITATED FACILITY LOCATION AND NETWORK DESIGN PROBLEM WITH CAPACITY EXPANSION FOR WASTEWATER TREATMENT NETWORKS

In our rapidly changing world, especially the increase in population and industrialization activities results in significant increases in wastewater production, which leads to the problem of treatment of the wastewater produced. Therefore, developing effective wastewater treatment networks continues to be vital to ensure environmental sustainability. In this context, wastewater treatment facilities must be located in the most appropriate places to make the best use of limited resources and the network required for wastewater transportation should be designed to meet future needs with the least cost. Although the facility location and network design problems in the literature are related to this issue, they do not cover some specific constraints for wastewater transport and often have a static structure in which time is not taken into consideration. In this thesis, firstly, a dynamic mixed-integer linear programming model for wastewater treatment facility location selection and network design is proposed in which facilities and links are assumed to be capacitated. Later on, a more complicated model which allows the facilities to be expanded in case of need according to time-dependent demand is proposed and defined on a three-dimensional space.

Several test experiments are conducted for the proposed models and they are solved with Python modeling library for optimization DOcplex and CPLEX version of 20.1. Moreover, extreme test instances are generated and their results are observed to verify the proposed models.

ÖZET

ATIK SU ARITMA AĞLARI İÇİN KAPASİTE GENİŞLETMELİ DİNAMİK KAPASİTELİ TESİS YERLEŞİMİ VE AĞ TASARIMI PROBLEMİ

Hızla değişen dünyamızda özellikle nüfus ve sanayileşme faaliyetlerindeki artış, atık su üretiminde de ciddi artışları beraberinde getirmekte; bu da üretilen atık suların etkin bir şekilde arıtılması problemini doğurmaktadır. Bu nedenle, etkili atık su arıtma ağlarının geliştirilmesi, çevresel sürdürülebilirliği sağlamak adına hayati bir önem arz etmeye devam etmektedir. Bu bağlamda, sınırlı kaynakların en iyi şekilde kullanılması amacıyla, atık su arıtma tesisleri en uygun yerlerde konumlandırılmalı, atık suyun taşınımı için gerekli şebeke de en az maliyetle geleceğe dönük ihtiyacı da karşılayacak şekilde tasarlanmalıdır. Literatürde mevcut olan tesis yer seçimi ve ağ tasarımı problemleri bu konu ile ilgili olsalar da, atık su taşınımına özel birtakım kısıtları kapsamamakta ve çoğunlukla zamanı hesaba katmayan statik bir yapıda bulunmaktadır. Bu tez çalışmasında, ilk olarak, atık su arıtma tesisi yer seçimi ve ağ tasarımına yönelik, tesislerin ve bağlantıların kapasiteli olarak varsayıldıkları dinamik bir karma tam sayılı doğrusal programlama modeli önerilmektedir. Daha sonra, zamana bağlı talebe göre ihtiyaç durumunda mevcut tesislerin genişletilmesine de olanak sağlayan daha karmaşık bir model sunulmuş ve üç boyutlu bir uzay üzerinde tanımlanmıştır.

Model için çeşitli test deneyleri oluşturulmuştur ve bu deneyler, Python DOcplex modelleme kütüphanesinde yer alan CPLEX 20.1 versiyonuyla çözülmüştür. Ayrıca, önerilen modeli doğrulamak için aşırı test örnekleri oluşturulmuş ve sonuçları gözlemlenmiştir.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS/ABBREVIATIONS

cm	centimeter
km	kilometer
kWh	Kilowatt-hour
m	meter
$U[a,b]$	Uniformly distributed random variable between a and b
$	United States dollar
BD	Bender's decomposition
CFLNDP	Capacitated facility location-network design problem
CNDP	Capacitated network design problem
DFLNDP	Dynamic facility location-network design problem
DNDP	Dynamic network design problem
e.g.	exempli gratia (for example)
et al.	et alii (and all others)
etc.	et cetera (and so forth)
FA	Firefly algorithm
FLNDP	Facility location-network design problem
FLP	Facility location problem
GA	Genetic algorithm
GHG	Greenhouse gas
i.e.	id est (that is)
IP	Integer programming
IWO	Invasive Weed Optimization algorithm
MILP	Mixed-integer linear programming
MINLP	Mixed-integer nonlinear programming
NDP	Network design problem
NPV	Net present value
NPW	Net present worth

OECD	Organisation for Economic Co-operation and Development
SA	Simulated Annealing
TURKSTAT	Turkish Statistical Institute
UFLNDP	Uncapacitated facility location-network design problem
UFLP	Uncapacitated facility location problem
UNDP	Uncapacitated network design problem
VNS	Variable Neighborhood Search
WQI	Water quality index
WRI	World Resources Institute

1. INTRODUCTION

Water is vital for the survival of living organisms. From the smallest living organism to the largest living being, water sustains all biological life and human activities. Moreover, water is essential for natural systems, wildlife and environmental resilience and is placed at the center of sustainable development. It has an effect on numerous bio-molecular activities [1]. In addition, water resources have a crucial role in various segments of economies such as agriculture, livestock production, forestry, industrial activities, hydro-power generation and fisheries.

With the rapid disposal of water resources and the critical increase in water consumption, developing effective wastewater treatment solutions is vital to ensure environmental sustainability. In recent years, the increase in the amount of water used in the production, agriculture, and energy sectors, as well as in the residences, reveals the necessity of minimizing the damage of the wastewater to the existing water resources, living things, soil, and the environment in general. In this context, wastewater treatment facilities must be located in the most appropriate places to make the best use of limited resources, and the network required for wastewater transportation should be designed to meet future needs with the least cost. The primary motivation of this thesis is to develop an optimization method in order to establish a wastewater treatment network while minimizing the total cost.

In the remainder of this section, global water issues and the importance of wastewater treatment are discussed. Then, the situation in Turkey is examined with the support of several metadata. Finally, considering all these issues, the purpose and contributions of this thesis are presented, and outlines are introduced.

1.1. GLOBAL WATER ISSUES

Humanity's relationship with water has continued to increase throughout history, and the water demand has increased with the growing population, developing urbanization and changing living standards. Water, which has been one of the most critical inputs of agriculture for a long time, is used intensively in industry, especially in energy production. It has become an input that is used and cannot be substituted. Burek *et al.* [2] projected that the global water use will continue to increase by about one percent annually, resulting in an increase of approximately 30 percent to the current level of water use by the year of 2050. The increase in the usage areas of water has once again revealed the strategic importance of this scarce resource. However, with the increasing consumption and disposal of water resources due to population growth, industrial developments and climate change, water scarcity is becoming an increasingly prominent and widespread problem worldwide [3,4]. Besides, water quality deteriorates rapidly in terms of water quality index (WQI) [5,6].

In the last decade, many research has been made about the threat of water scarcity and very severe consequences are foreseen in the next 50 years [7,8,9,10]. Recently, World Economic Forum 2020 [11] determined the world's top ten global risks and three of them were related to water and environmental sustainability. In addition, according to the United Nations' recent report [12], more than two billion people live in areas subject to water stress. They also stated that around 3.4 billion people (45 percent of the global population) lack access to safely managed sanitation facilities. According to independent assessments, the world will face a global water deficit of 40 percent by 2030. This situation will be worsened by global challenges such as Covid-19 and climate change. Figure 1.1 illustrates the annual baseline water stress of the world, which is provided by the World Resources Institute (WRI) in their water risk atlas [13]. Baseline water stress calculates the ratio of the total water withdrawals to the available renewable water resources.

1.2. WASTEWATER TREATMENT

The water distributed to consumers must meet certain quality standards and quantities. If the water is not pure enough for human consumption, its content and components may negatively affect human health. Globally, two million tons of sewage, industrial and agricultural waste into the world's waterways and at least 1.8 million children under the age of five die yearly from water-related diseases [14]. If infrastructures and water management are not improved, millions of people will continue to die each year and there will be further losses in biodiversity and ecosystem resilience. That is why wastewater treatment is essential. With the increasing threat of water scarcity, developing effective wastewater treatment solutions continues to be

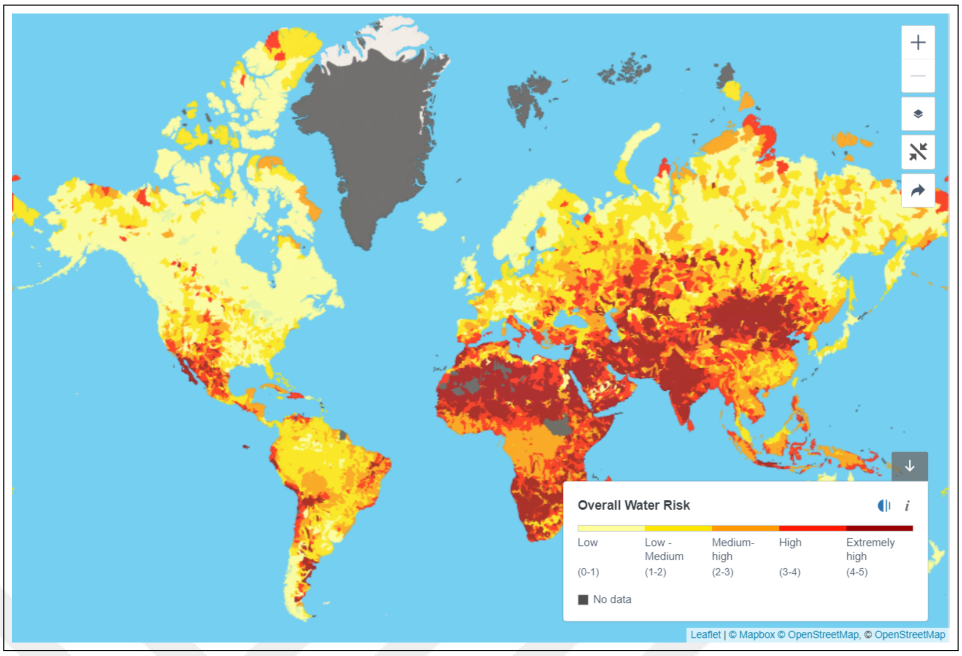


Figure 1.1. Annual baseline water stress per region, WRI

vital to ensuring environmental sustainability.

In fact, the concept of wastewater can mean distinctive things to different people as it has several definitions. Basically, wastewater can be defined as water that has been polluted due to domestic, industrial, agricultural and other activities or has lost its characteristic hygienic properties, accumulated after rain flows in parks, streets and similar areas [15,16]. The most comprehensive definition of wastewater is proposed by Raschid-Sally and Jayakody [17]:

> Wastewater is a combination of one or more of: domestic effluent consisting of blackwater (excreta, urine and faecal sludge) and greywater (kitchen and bathing wastewater); water from commercial establishments and institutions, including hospitals; industrial effluent, stormwater and other urban run-off; agricultural, horticultural and aquaculture effluent, either dissolved or as suspended matter.

On the other hand, wastewater treatment is the physical, chemical and biological processes applied to ensure that the wastewater generated as a result of various uses does not change the typical characteristics (physical, chemical, bacteriological and ecological) of the receiving environment where they are discharged. Briefly, it is the process of returning wastewater to the water cycle. Thus, it becomes directly reusable for numerous purposes. Different degrees of wastewater treatment according to the increasing level of treatment are described in general terms: primary treatment, secondary treatment, tertiary and/or advanced treatment

[18]. Primary treatment includes physical treatment units that involve the removal of floating and settled solids from wastewater. Secondary treatment includes biological and chemical treatment units that remove residual organics and suspended solids. Advanced treatment, on the other hand, is the treatment process used to remove pollutants (nitrogen, phosphorus, heavy metals, toxic organic substances, etc.) that cannot be sufficiently purified or treated with physical or biological treatment methods. Nitrification, denitrification, adsorption, ion exchange, etc., are the main advanced treatment methods.

Wastewater from households and industry creates a significant pressure in receiving environments due to organic matter, nutrient load and hazardous materials. In cities, most of the wastewater is collected by sewage systems connected to the wastewater treatment plants. The level of treated water determines the scale of the impact on aquatic ecosystems. As mentioned before, wastewater treatment is an essential practice to use water more efficiently and protect existing resources and environmental sustainability. Therefore, the management and treatment of wastewater have been an area of concern for humanity for centuries. Especially following the technological developments, effectively functioning wastewater collection and treatment infrastructures have been established in the most developed countries to serve more than 91 percent of the population [19,20]. In order to execute the treatment and discharging processes, the wastewater should be transferred from the source points to the treatment plants in the most appropriate way. Afterward, the treated water should be conveyed to the discharge point [21]. Hence, the location of the wastewater treatment plants and the establishment of the appropriate network are crucial in terms of cost, accessibility and environmental factors.

1.3. WASTEWATER TREATMENT IN TURKEY

Turkey is a country encircled by seas on three sides and has plenty of rivers and lakes as shown in Figure 1.2 which illustrates the topographic map of the country. The Aegean Sea covers the western side of the country, whereas the Black Sea covers the northern side and the Mediterranean Sea covers the southern side. Although its geography seems rich in water, Turkey is not a water-rich country. State Hydraulic Works of Turkey estimated that the annual amount of usable water per capita in Turkey was 1652 m^3 in 2000, 1544 m^3 in 2009, and

1346 m^3 in 2020 [22]. According to their criteria, countries with an annual per capita water amount of more than 8,000 m^3 are water-rich, countries with less than 2,000 m^3 of water are among the countries with water scarcity, and countries with less than 1,000 m^3 of water are among the water-poor countries. Turkish Statistical Institute (TURKSTAT) projected that the population of Turkey will reach 100 million by 2040 [23]. Therefore, considering the current water amount and consumption, water shortage is foreseen in Turkey in the upcoming years. Besides, a significant part of the country is dealing with water stress, as can be seen in Figure 1.1. Organization for Economic Co-operation and Development (OECD) published the water and climate change adaptation report and stated that the water quality and the water quantity are among the primary concerns of Turkey [24]. In particular, some studies have been carried out recently to solve the water problem in metropolises and large cities in the country [25,26,27,28].

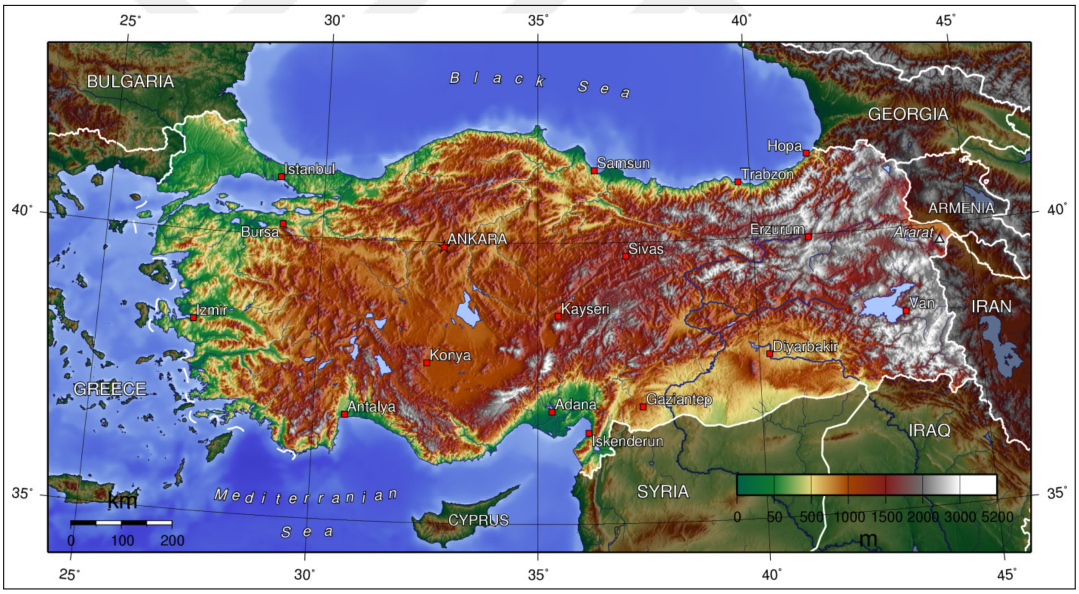


Figure 1.2. Topographic map of Turkey

Although the number of municipalities served by wastewater treatment plants in Turkey has increased significantly in the last 30 years, 755 out of 1399 municipalities in the country do not have any wastewater treatment plants [29]. According to TURKSTAT metadata, 21 percent of the total population live in regions with no wastewater treatment plants [30]. Moreover, 10 percent of the total population does not live in municipalities with sewerage networks. Only 61 percent of the wastewater generated from industrial and domestic water use is treated in water treatment plants. Table 1.1 summarizes Turkey's municipal wastewater statistics every four years from 2002 to 2018. The development mentioned in wastewater management in recent years is clearly observed, especially with the increasing rate of the treated amount of wastewater and the rate of the population living in the region where there is a treatment plant.

Nevertheless, one of the main fundamental points to note is that, according to 2018 data, approximately one-third of the wastewater treatment plant capacities are not used. The ratio of the treated wastewater discharged from the network according to the receiving environments to the total wastewater treatment plant capacity has increased from 2002 to 2018 but reached only 66.54 percent. That level of idleness can be seen as a serious loss in terms of capacity utilization.

Table 1.1. Municipal wastewater statistics of Turkey

Statistic	**Year**				
	2002	**2006**	**2010**	**2014**	**2018**
Total number of municipalities	3,227	3,225	2,950	1,396	1,399
Number of municipalities served by wastewater treatment plants	248	362	438	513	644
Rate of municipalities served by wastewater treatment plants (%)	7.69	11.22	14.85	36.75	46.03
Number of municipalities served by sewerage system	2,115	2,321	2,235	1,309	1,357
Rate of municipalities served by sewerage system (%)	65.54	71.97	75.76	93.77	96.99
Rate of population served by wastewater treatment plants (%)	35.5	50.6	61.7	68.1	79
Rate of population served by sewerage system (%)	83	87	88	90	91
Amount of wastewater discharged from municipal sewerage (million $m^3/year$)	2,498	3,367	3,582	4,297	4,795
Amount of wastewater treated by wastewater treatment plants (million $m^3/year$)	1,312	2,140	2,719	3,484	4,236
Rate of treated amount of wastewater discharged from municipal sewerage (%)	52.52	63.56	75.91	81.08	88.34
Capacity of wastewater treatment plants (million $m^3/year$)	2,359	3,648	5,293	5,549	6,367
Efficiency of wastewater treatment plants (%)	55.61	58.66	51.37	62.78	66.53

According to the same metadata, the total number of wastewater treatment plants in the country, which was 145 in 2002, increased to 991 in 2018. In 2018, 527 of the wastewater treatment plants in Turkey can be categorized as a biological treatment plants, 206 as a natural treatment plants, 55 as a physical treatment plants and 203 as advanced treatment plants. Approximately 6.4 billion m^3 of wastewater was treated in these facilities in 2018. Approximately 1.7 billion m^3 of this quantity is biological treatment, 25 million m^3 is natural treatment, 1.7 billion m^3 is physical treatment, and 2.9 billion m^3 is an advanced treatment. These statistics reveal that the ratio of advanced treatment is not adequate. The details of the wastewater treatment plant statistics are summarized in Table 1.2.

Table 1.2. Wastewater treatment plant statistics of Turkey

Measurement	**Type of Plant**	**Year**				
		2002	**2006**	**2010**	**2014**	**2018**
Number of wastewater treatment plants	Biological treatment	114	135	199	345	527
	Natural treatment	0	0	35	118	206
	Physical treatment	28	26	39	49	55
	Advanced treatment	3	23	53	92	203
	Total	145	184	326	604	991
Capacity of the wastewater treatment plants (million $m^3/year$)	Biological treatment	1,320	1,511	1,733	1,713	1,718
	Natural treatment	0	0	12	58	26
	Physical treatment	771	1,329	1,839	1,823	1,738
	Advanced treatment	267	808	1,709	1,955	2,885
	Total	2,359	3,648	5,293	5,549	6,367

In addition, it is observed that there is an unbalanced distribution across the regions of the country in terms of wastewater treatment. For instance, the the population living in municipalities with wastewater treatment plants in Northeast Anatolia remained below 10 percent until 2018. On the other hand, the fact that this rate is over 90 percent in Istanbul is a prominent indicator of the mentioned unbalanced distribution. Table 1.3 summarizes this rate per every nomenclature of territorial units for statistics (NUTS) regions of the country. It is seen that the rate is still below 60 percent in these Central East Anatolia, North East Anatolia and West Black Sea regions.

Table 1.3. Rate of population served by wastewater treatment plants per NUTS statistical regions of Turkey (%)

Region	Year				
	2002	2006	2010	2014	2018
Aegean	47.6	50.8	64.0	74.0	83.4
Central Anatolia	8.2	40.4	51.9	61.0	74.1
Central East Anatolia	21.3	45.2	43.1	46.7	53.5
East Black Sea	17.4	6.9	25.0	25.7	62.4
East Marmara	48.5	64.2	77.1	81.4	87.3
İstanbul	62.7	80.6	91.0	97.8	99.6
Mediterranean	12.7	34.3	51.4	61.0	68.7
North East Anatolia	9.3	11.2	9.0	8.3	46.7
South East Anatolia	29.3	38.6	42.4	47	60.6
West Anatolia	50.8	63.2	80.9	86.2	90.9
West Black Sea	15.4	19.2	25.8	35	59.8
West Marmara	9.6	17.0	16.2	38.6	74.1

1.4. CONTRIBUTION OF THIS THESIS

In the light of all the facts mentioned above, it is clear that the requirements for the quality and the quantity of wastewater treatment processes are significant and there are severe problems regarding the efficient usage of existing plant capacities. Wastewater treatment and the efficiency of the plants are not at an appropriate level at all, primarily due to cost, geographical conditions, or infrastructure deficiencies. Therefore, it is crucial to determine the location of the wastewater treatment plants and the settlements they will serve. In addition to the plant locations, decisions such as the design of the necessary wastewater network, the capacity of the plants to meet the demand that will change according to the years, or increasing the capacity of the existing facilities are among the important issues to be considered.

The main motivation of this study is to propose an optimization method to establish an effective wastewater treatment network. In this context, firstly, a dynamic mixed-integer linear programming model for wastewater treatment facility location and network design is proposed in which facilities and links are assumed to be capacitated. Considering the varying behavior of the cost and demand over time in real-world occasions, the proposed model addresses the problem in a dynamic or time-varying structure. Later on, a more complicated model which allows the facilities to be expanded (i.e., their capacities can be increased) in case of need according to time-dependent demand is proposed and defined on a three-dimensional space considering the importance of altitudes of the nodes for wastewater treatment networks. Both mathematical models ensure all corresponding demands in a defined network while improving the utilization of the facilities. The facilities are categorized according to their capacities.

To the best of our knowledge, there is a lack of a dynamic facility location-network design model, especially in water distribution networks. It is clear that constructing a mathematical formulation with a dynamic structure is a necessary condition in order to be more practical and applicable to real-world situations. Studies regarding dynamic location models in the literature ignore a fundamental point: the construction times of facilities and links. While developing a dynamic formulation, the main aim is to consider the fluctuations in key parameters. Clearly, it is not practical for a facility or a link to become operational right after the construction decision is made. Therefore, for each facility or a link; defining distinct construction times would modify the system more practically. That structure is considered in the proposed models in this thesis.

In addition, the pumping and pressure factors encountered in water transportation should be evaluated. Although energy calculations in a water distribution network have been studied in the literature, no study handles them simultaneously with the typical facility location-network design problems. In this study, pumping and energy calculations are considered to represent economics of scale for the wastewater treatment networks as realistic as possible.

Another crucial point is the verticality which is a factor that directly affects the costs associated with the water distribution networks. Therefore, instead of defining a network with two dimensions as in the typical facility location problems, including the third dimension as altitude would be a more realistic approach in terms of cost analysis, which is provided in this thesis.

1.5. OUTLINE OF THIS THESIS

The rest of this thesis is organized as follows:

- In Section 2, the related problems in the literature are introduced and the existing literature is re-evaluated.

- In Section 3, a mathematical formulation from the literature is introduced and then the proposed mathematical models are presented in detail.
- In Section 4, computational experiments are conducted and their results are discussed.
- In Section 5, the validation and the verification of the proposed model is provided.
- In Section 6, the applicability of the proposed model is discussed and possible future works related to this thesis are introduced.
- In Section 7, the main conclusions from this thesis are drawn.

2. LITERATURE REVIEW

In general, facility location problems (FLP) are related to decisions about where to locate facilities and how to allocate these facilities to meet the customers' demand for particular services or products while minimizing the total cost. These types of problems can be used in several different sectors such as health, education, fire, production, storage and in supply chain methods like water distribution networks. On the other hand, the network design problem (NDP) deals with creating network links that connect customers to facilities and determine the traffic flow on those links. Generally, in both problems, there is a condition that the demand is fully satisfied. These problems are often helpful for real-world applications since they consider fixed facility costs, facility set-up costs, link set-up costs, transportation costs and operating costs.

The facility location-network design problem (FLNDP) is a more complex problem obtained by integrating the two problems mentioned above. Typical facility location models have an underlying network and the required number of facilities are selected from candidate locations. However, such a condition is not sought in FLNDP formulations. Customers may not receive service directly from the facilities; the demand or the product may pass through multiple nodes in a specified network for distribution purposes. This is clearly a more cost-effective method, as it provides the opportunity to improve the existing network rather than having to construct a new facility to meet the demand.

Figure 2.1 presents the schematic framework of these decision problems in order to highlight their generic structures. Points represent nodes or demand points, whereas squares represent located facilities over the network. Moreover, solid lines represent assignment decisions; dotted lines represent routing decisions, and dash-dotted lines represent a combination of assignment and routing decisions.

2.1. FACILITY LOCATION PROBLEM

The study of the FLP dates back to the early 20th century. In 1909, Weber [31] proposed the first single facility location formulation by minimizing the distance between the facility

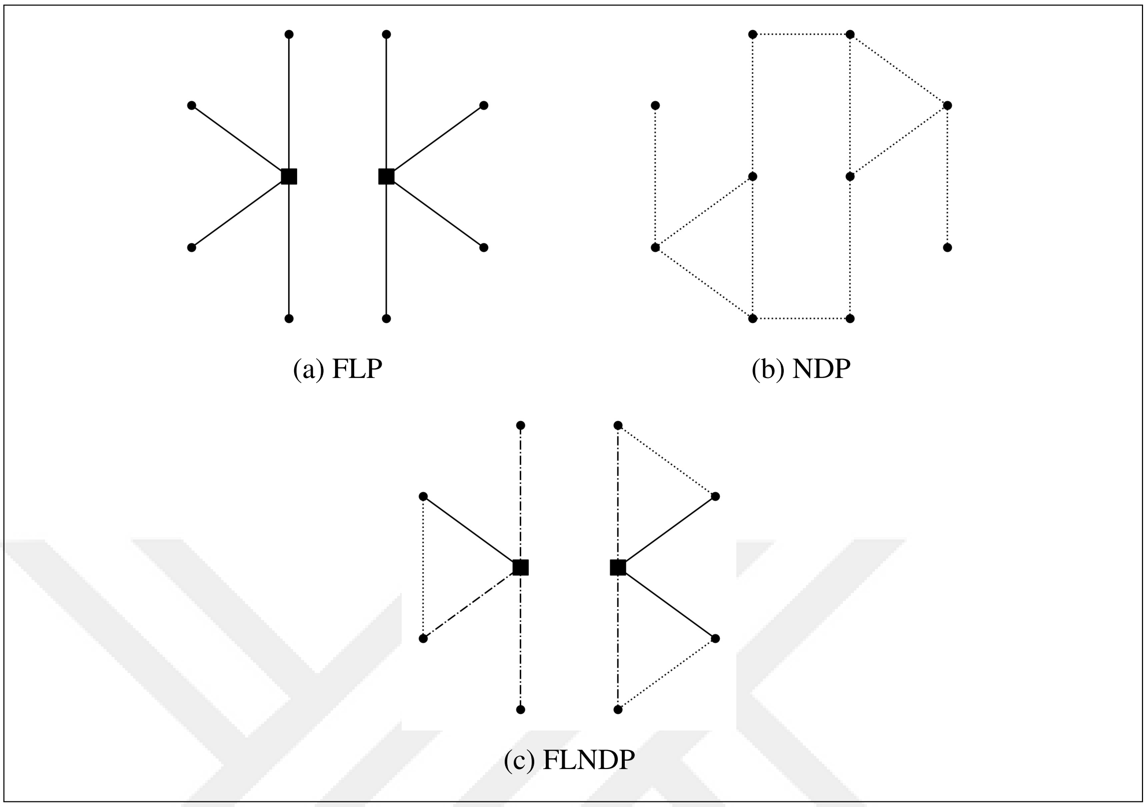


Figure 2.1. Schematic framework of the related decision problems

and the customers. Although Weber seemed to overestimate supply factors and ignore the demand in his theory, it paved the way for later studies. Kuehn and Hamburger [32] proposed the uncapacitated facility location problem (UFLP) while considering the transportation cost and the operating cost in 1963. Later on, Hakimi [33] presented a novel mathematical model that minimizes the total demand-weighted travel distance in order to locate one or more facilities in 1964. His formulation is later known as the "P-median problem". ReVelle [34] modified that problem and proposed the maximum capture problem, which seeks the location of a given number of facilities in order to maximize the captured market share. Another important issue considered in the FLP formulations is the coverage. The coverage concept is used to specify if a customer can be served by a facility in an acceptable travel distance or time. Toregas *et al.* [35] proposed the set covering location problem, which minimizes the total number of facilities to be used. On the other hand, Church *et al.* [36] proposed the maximum covering location problem, which maximizes the number of covered demand points. Elloumi *et al.* [37] presented a different objective that minimizes the maximum travel time. Since the beginning of the 21st century, several comprehensive reviews have been published in this field [38,39,40,41,42].

2.2. NETWORK DESIGN PROBLEM

NDP has been started to be studied more recently than the FLP. In 1984, Magnanti and Wong [43] proposed a solution technique for fixed-charge NDP, detailing the decomposition techniques and algorithms comparatively. Later, Balakrishnan *et al.* [44] showed that the large-scale fixed-charge uncapacitated network design problem (UNDP) can be solved effectively by the dual ascent method and drop-add heuristic. Multi-commodity NDP was studied by Balakrishnan *et al.* [45]. Gendron *et al.* [46] compared and discussed the solution methods (simplex-based cutting plane algorithms, Lagrangean relaxation and heuristics) for capacitated network design problem (CNDP). Drezner and Wesolowsky [47] introduced another class of NDP in which links can either be one-way or two-way; termed as the mixed-link problem. Yang and Bell [48] presented a detailed review of NDP by discussing the solution methods in 1998. These problems are still an active area of research, one may refer to Croxton *et al.* [49], Costa *et al.* [50], Frangioni and Gendron [51], Hewitt *et al.* [52].

2.3. FACILITY LOCATION AND NETWORK DESIGN PROBLEM

The integration of these two main problems was first studied by Daskin *et al.* [53] in 1993. They proposed an uncapacitated facility location-network design problem (UFLNDP) to minimize the total construction and transportation costs in a specified network. In this problem, they defined a set of nodes representing the customers and candidate facility locations; edges represent possible links between those nodes. They ensure that all demands must be routed over the specified network. Later, Melkote [54] developed a solution method for this problem in his doctoral thesis. Even further, Daskin and Melkote [55,56] worked together in order to investigate the capacitated facility location-network design problem (CFLNDP) and the maximum covering location-network design problem in 2001. Subsequently, extensive and growing literature has emerged to handle this problem. Different structures and assumptions have been used and numerous solution methods have been proposed. Ravi and Sinha [57] and Chen and Chen [58] proposed several approximation algorithms to combine facility location and cable installation in which capacity constraints are included on links or cables in their case. In a similar vein, Mortezaei and Jabalameli [59] introduced an extension for the CFLNDP in which a limited amount of demand can be routed over a link. They presented a bi-objective mathematical formulation and introduced a hybrid algorithm to solve the proposed problem. Cocking [60] introduced the budget-constrained UFLNDP and developed several algorithms to solve this problem in his doctoral thesis. Soon after, Cocking *et al.* [61] presented several heuristics and metaheuristics to solve this problem with minisum objective. In their study, simple greedy heuristics, a custom heuristic that separates the problems of facility location and network design, a basic local search, and simulated annealing (SA) heuristics using two different kinds of neighborhoods are presented. Bigotte *et al.* [62] formulated a mixed-integer program that simultaneously promotes urban centers to new levels of hierarchy and improves network links and developed a heuristic method that can find optimum or near-optimum solutions. They assumed that lower-level services may always be obtained from higher-level centers. Contreras *et al.* [63] used a different objective by minimizing the maximum travel time for this problem and proposed two integer programming formulations which they compared empirically. Rahmaniani and Ghaderi [64] proposed one of the most detailed models for the budget-constrained UFLNDP. They considered different type of capacitated links and implemented the firefly algorithm (FA) to solve the problem. In recent years, several studies have been conducted to formulate the combined FLNDP [65,66,67,68,69,70,71].

Most studies addressing the FLNDP have focused on static or single-period models. This approach actually does not resemble real-world occasions. Since the main parameters such as the demand and the cost have a waiving pattern over time, it is more appropriate to conduct a dynamic or time-variant formulation in order to demonstrate the problem more accurately. In such formulations, a time horizon may also be defined as a distinct set, and a different decision can be made regarding each time period that takes place in this specified set. One of the most well-prepared studies in this respect is proposed by Ghaderi and Jabalameli [72]. They presented the budget-constrained dynamic facility location-network design problem (DFLNDP) and applied it for healthcare evaluations in the Illam Province of Iran. They introduced various heuristic methods, such as a greedy heuristic and SA, as well as some exact methods. They split the problem into sub-problems in order to simplify the formulation. Ghaderi [73] implemented the Variable Neighborhood Search (VNS) algorithm, whereas Pearce and Forbes [74] applied Bender's decomposition (BD) into this problem. In 1960s, Manne [75] and Ballou [76] studied the dynamic location problem. In general, their studies handled a single facility and the main issue was capacity expansion. In 1971, Scott [77] introduced the dynamic multi-facility location-allocation model and provided a heuristic based on dynamic programming. Later on, Sweeney and Tatham [78] applied BD to this problem. Dynamic single-facility location problem remained an area of interest, one may refer to Drezner and Wesolowsky [79], Bastian and Volkmer [80], Andreatta and Mason [81]. Moreover, Wesolowsky [82] presented a dynamic multi-facility minisum problem. In the presented problem, opening and closing of facilities may occur in the time horizon. Afterward, the studies considered the dynamic multi-facility location problem had been diversified [83,84,85,86]. In some studies, authors also consider the dynamic multi-echelon facility location model in which they handle different type of commodities [87,88,89]. Dynamic network design problems (DNDP) are also studied in the literature; one may refer to the stochastic formulations proposed by Lee and Dong [90], Ukkusuri and Kaan [91].

One of the main concerns in this thesis is the possibility of extending the capacities of facilities and links. The phenomena of capacity expansion have been studied in dynamic location problems [92,93,94,95,96]. Strategic decisions regarding supply chain networks such as optimal facility expansion capacities, selection of the optimum combination of facilities and their allocations are included in these studies. Julka *et al.* [97] introduced a comprehensive review to investigate multi-factor models for capacity expansion of facilities. More recently, Badri *et al.* [98] developed a mixed-integer programming model for the supply chain network design problem in which they considered capacity expansion. The main objective of their study is to maximize the supply chain net profit. They proposed decisions about supplier selections, production facility locations, the amount of raw materials to be supplied, the amount of each product to be produced and the expansion plans over a time horizon.

Table 2.1 summarizes the important characteristics of the related FLNDP problems in the literature by categorizing them as static or dynamic and according to their capacity considerations. Table 2.2 summarizes the solution approaches of the related problems. First, their optimization problems are classified as IP (integer programming), MILP (mixed-integer linear programming) and MINLP (mixed-integer nonlinear programming). Then, if any, their heuristic approaches such as genetic algorithm (GA), variable Neighborhood Search (VNS), firefly algorithm (FA), Bender's decomposition (BD), Invasive Weed Optimization algorithm (IWO), simulated annealing (SA) and case-specific custom heuristics are presented.

Table 2.1. Literature summary of the FLNDP: Important Characteristics

Study	**Study Period**		**Capacity**			
	Static	**Dynamic**	**No Capacity**	**Facility Capacity**	**Link Capacity**	**Capacity Expansion**
Melkote & Daskin [55]	✓		✓			
Melkote & Daskin [56]	✓			✓		
Ravi & Sinha [57]	✓				✓	
Chen & Chen [58]	✓			✓	✓	
Mortezaei & Jabalameli [59]	✓			✓	✓	
Cocking *et al.* [61]	✓		✓			
Bigotte *et al.* [62]	✓		✓			
Contreras *et al.* [63]	✓		✓			
Rahmaniani & Ghaderi [64]	✓				✓	
Lashine *et al.* [65]	✓			✓	✓	
Shishebori *et al.* [66]	✓		✓			
Wu *et al.* [67]	✓			✓		
Kaya & Ozkok [68]	✓				✓	
Liu & Kwon [69]	✓		✓			
Brahami *et al.* [70]	✓			✓	✓	
Sadat Asl *et al.* [71]	✓			✓	✓	
Ghaderi & Jabalameli [72]		✓	✓			
Ghaderi [73]		✓	✓			
Pearce & Forbes [74]		✓	✓			
Badri *et al.* [98]		✓		✓		✓
Model A (This Thesis)		✓		✓	✓	
Model B (This Thesis)		✓		✓	✓	✓

Table 2.2. Literature summary of the FLNDP: Solution Approaches

Study	**Program**			**Heuristic**						
	IP	**MILP**	**MINLP**	**GA**	**VNS**	**FA**	**BD**	**IWO**	**SA**	**Custom**
Melkote & Daskin [55]		✓								
Melkote & Daskin [56]		✓								
Ravi & Sinha [57]										✓
Chen & Chen [58]	✓									✓
Mortezaei & Jabalameli [59]		✓							✓	✓
Cocking *et al.* [61]								✓		✓
Bigotte *et al.* [62]		✓								✓
Contreras *et al.* [63]		✓								
Rahmaniani & Ghaderi [64]		✓				✓				
Lashine *et al.* [65]		✓								
Shishebori *et al.* [66]			✓							
Wu *et al.* [67]		✓		✓						
Kaya & Ozkok [68]			✓						✓	
Liu & Kwon [69]							✓			
Brahami *et al.* [70]		✓		✓						
Sadat Asl *et al.* [71]						✓		✓		
Ghaderi & Jabalameli [72]		✓							✓	✓
Ghaderi [73]		✓	✓		✓					
Pearce & Forbes [74]			✓				✓			
Badri *et al.* [98]		✓								
Model A (This Thesis)		✓								
Model B (This Thesis)		✓								

If the topics are narrowed to wastewater treatment, it can be seen that the first optimization problem was proposed by Loucks *et al.* [99] in 1967. They presented a linear programming model that determines the minimum total cost of any particular dissolved oxygen control policy in a river basin by ensuring the design efficiencies required for each wastewater treatment facility in order to maintain any specified set of stream dissolved oxygen concentration. In 1973, Converse [100] proposed a dynamic design optimization model that decides the optimum number of plants in a specific region. Jarvis *et al.* [101] proposed an assignment and network design problem that minimizes the expected cost of the wastewater treatment system. Kovacs *et al.* [102] proposed a two-stage approach for sewer management to a specific place called the Lake Balaton resort area. In another study, Zhu *et al.* [103] adapted the UFLP to solve a regional treatment system. The first combined FLNDP for the wastewater treatment system was proposed by Leitão *et al.* [104] in 2005. They used Daskin's formulation in order to optimize the wastewater treatment systems in the rural areas of Portugal. They then developed a greedy algorithm to search for local optimal solutions. Zeferino *et al.* [105] developed a bi-objective optimization model that maximizes the dissolved oxygen concentration in the river water while minimizing the infrastructure cost. Right after, they published another article addressing the same issue, but this time the uncertainty in demand was taken into account [106]. Recently, Rezaei *et al.* [107] proposed a multi-objective deterministic FLP for water reclamation planning. The proposed model aims to minimize the costs and greenhouse gas (GHG) emissions associated with the entire wastewater system and maximize the value of reclaimed water. In 2020, Zhao *et al.* [108] developed a two-stage guarantee rate optimization model that considers uncertainty from source pollutants.

To the best of our knowledge, there is a lack of a dynamic facility location-network design model especially in water distribution networks. As mentioned earlier, constructing a mathematical formulation with a dynamic structure is a necessary condition in order to be more practical and applicable into real-world occasions. Studies regarding dynamic location models in the literature ignore a fundamental point which is the construction time. While developing a dynamic formulation, the main aim is to consider the fluctuations in key parameters. It is clearly not practical for a facility or a link to become operational right after the construction decision is made. Therefore, for each facility or a link, defining distinct construction times modify the system more accurately. Given that the money will also ramify in value within any given time horizon, the costs associated with a dynamic model should be defined as adaptable into ramification.

In addition, the pumping and pressure factors encountered in water transportation should be evaluated. Although energy calculations in a water distribution network have been studied in the literature before in which pumping cost due to pressure is calculated, there is no study that handles them simultaneously with the typical facility location-network design problems. Nevertheless, especially the study of Brand and Ostfeld [109] in this field has been one of the main sources of inspiration for this thesis. Their formulations for pumping pipeline construction cost, gravitational pipeline construction cost, pump construction cost, pump energy cost and treatment plant construction cost are used in this study. Besides, the studies presented by Öztürk [110], Lim *et al.* [111] and Chover *et al.* [112] are also taken as references in the economic scale of this study.

Another crucial point to note is that verticality is a factor that directly affects the costs associated with the water distribution networks. Therefore, instead of defining a network with two dimensions as in the typical facility location problems, including the third dimension as altitude would be a more realistic approach in terms of cost analysis.

3. METHODOLOGY

In this section, it is aimed to establish an optimization methodology that differs from the solution methodologies in the literature in some points, which will be detailed later, in order to solve global or regional water problems and ensure their sustainability. Accordingly, first of all, the mathematical formulation proposed by Melkote and Daskin [56] will be introduced, which was taken as a reference while developing the methodology and discuss the reasons behind why this problem is applicable to wastewater treatment networks in Section 3.1. Then, two novel dynamic MILP models for optimizing wastewater treatment networks will be proposed in Section 3.2 and Section 3.3. The features and assumptions of these models (Model A, Model B), which were developed gradually, are introduced and their formulations are presented.

3.1. CFLNDP

CFLNDP was proposed by Melkote and Daskin [56] and designed to be used in several application areas such as regional planning, distribution, logistics, telecommunications, energy management. They extended the UFLNDP proposed by Daskin [53] by including facility capacities. The authors first presented a MILP model, then derived the model with various valid inequalities to relax the problem. This section introduces their final formulation with the addition of valid inequalities.

The authors defined the problem on a network of nodes and possible connections/links between these nodes. The assumptions used in the model can be listed as follows:

- Each node is a demand point and has a certain level of demand to be met.
- Facilities can only be located at nodes. All the demands of these nodes can only be met from the located facilities.
- A maximum of one facility can be located at each node.
- The designed network has a structure in which the demand flows from demand points to the facilities. This structure is known as a customer-to-server system.

- The traveling costs between nodes are symmetrical.
- Demand is divisible. It is not necessary for a single facility to meet the whole demand of a node.

The notations used in the formulation of CFLNDP are displayed in Table 3.1. Instead of some notations used by Melkote and Daskin in their original formulation, different notations are used to be compatible with Model A and Model B.

Table 3.1. Notations used in CFLNDP

Notation	Description
Sets	
$\mathcal{N}$	Set of nodes; $i, j \in \mathcal{N}$ and set of demand points; $k \in \mathcal{N}$
$\mathcal{U}$	Set of possible directed links; $(i, j) \in \mathcal{U}$
Parameters	
d^k	Demand related to demand point k
$dist_{ij}$	Distance between node i and node j
f_i	Cost of constructing a facility on node i
c_{ij}	Cost of constructing link (i, j)
tc_{ij}	Traveling cost per demand unit on link (i, j)
CF_i	Capacity of a facility on node i
Decision Variables	
Z_i	1 if a facility is located on node i, 0 otherwise
X_{ij}	1 if link (i, j) is constructed, 0 otherwise
Y_{ij}^k	Fraction of the demand related to demand point k that flows on link (i, j)
W_i^k	Fraction of the demand related to demand point k served by a facility on node i

Contrary to previous studies that examined classical FLP, Melkote and Daskin used a distinct index of k to specify the nodes that represent demand points. With the addition of valid inequalities they derived, the authors first included this k index to the decision variables of Y_{ij} and W_i. With this modification, they have constructed a mathematical model that not only decides where to locate facilities and corresponding links to be established but also decides on the fraction of demand of the nodes served by a facility or a link. Although this new structure makes the mathematical model more complex, it allows one to control and observe the network flow and traffic.

Generally, in wastewater management, wastewater produced in settlements is collected and emphasis is placed on bringing it to the treatment facilities. Accordingly, the corresponding demand refers to the wastewater generated in the settlements and waiting to be treated. Therefore, considering its customer-to-server structure and the possibility of controlling the network flow, the combined FLNDP formulation proposed by Melkote and Daskin stands out with its applicability for establishing wastewater treatment networks.

Given the assumptions made and the notations used, the mathematical model they proposed can be given as follows:

$$\textbf{min.} \quad \sum_{i\in\mathcal{N}} f_i Z_i + \sum_{(i,j)\in\mathcal{U}} c_{ij} X_{ij} + \sum_{(i,j)\in\mathcal{U}} \sum_{k\in\mathcal{N}} tc_{ij} Y_{ij}^k d^k \tag{3.1}$$

$$\textbf{subject to;} \quad \sum_{j\in\mathcal{N}} Y_{ji}^k = \sum_{j\in\mathcal{N}} Y_{ij}^k + W_i^k; \qquad \forall i,k \in \mathcal{N} : i \neq k \tag{3.2}$$

$$W_i^i + \sum_{j\in\mathcal{N}} Y_{ij}^i = 1; \qquad \forall i \in \mathcal{N} : d^i > 0 \tag{3.3}$$

$$\sum_{k\in\mathcal{N}} d^k W_i^k \leq CF_i\, Z_i; \qquad \forall i \in \mathcal{N} \tag{3.4}$$

$$\sum_{i\in\mathcal{N}} W_i^k = 1; \qquad \forall k \in \mathcal{N} : d^k > 0 \tag{3.5}$$

$$Y_{ij}^k \leq X_{ij}; \qquad \forall (i,j) \in \mathcal{U}, \forall k \in \mathcal{N} \tag{3.6}$$

$$X_{ij} + X_{ji} \leq 1; \qquad \forall (i,j) \in \mathcal{U} \tag{3.7}$$

$$Z_i \in \{0,1\}; \qquad \forall i \in \mathcal{N} \tag{3.8}$$

$$X_{ij} \in \{0,1\}; \qquad \forall (i,j) \in \mathcal{U} \tag{3.9}$$

$$Y_{ij}^k \geq 0; \qquad \forall (i,j) \in \mathcal{U}, \forall k \in \mathcal{N} \tag{3.10}$$

$$W_i^k \geq 0; \qquad \forall i,k \in \mathcal{N} \tag{3.11}$$

The objective function (3.1) minimizes the total cost in the defined network, which consist of the construction cost of the facilities and links as well as the traveling cost. Constraints (3.2) ensure that each demand flows to node i is either served by node i to supply its demand or transferred to the other nodes. Constraints (3.3) ensure that if node i has a demand, it must be served either from a facility on node i or another facility in the network. Constraints (3.2) - (3.3), combined, provide the conservation of flow. The capacity constraints (3.4) ensure that demand can be served by node i only if there is a facility at node i. Additionally, the total demand served by node i cannot exceed the capacity of the facility constructed at node i. Constraints (3.5) state that each demand must be served. If a node has no demand, no action is necessary for that node. Constraints (3.6) ensure that the link (i, j) should be established in order to serve any demand in the network. Constraints (3.7) preclude the activation of oppositely directed flows on any link (i, j). The flow on any constructed link will be in only one direction. Constraints (3.8 - 3.11) state the non-negativity and integrality constraints for the decision variables.

Although the proposed mathematical model appears to be defined on an empty network, it is possible to integrate an existing network. To achieve this, the existing facilities and links can be detected and the construction costs of these materials can be set to zero in the model parameters.

3.2. MODEL A

Considering the static nature of CFLNDP, a novel mathematical formulation, Model A, in which the demand and the cost may change over time, is presented in this section. In order to be resembling real-world occasions, fixed capacities are not defined as in CFLNDP; on the contrary, the facilities are categorized with respect to their capacities. Thus, the categories of the facilities are decided in Model A, together with their locations. Besides that, all possible links between nodes are defined as capacitated, so they cannot serve an unlimited amount of demand in the network. The dynamic nature of the mathematical model requires a different structure in the construction of facilities and links, especially considering the time-varying cost. That is, in real-world occasions, the construction of a facility cannot be instantaneous. Therefore, Model A considers the construction times of facilities and links.

Just as CFLNDP, Model A is defined on a network of nodes and possible links between these nodes. However, there is a time horizon consisting of multiple time periods. While some assumptions used in CFLNDP are valid, the following assumptions have been used to represent Model A:

- Each node is a demand point and has a certain level of demand to be met for each time period throughout the time horizon.

- Each node is a candidate location for hosting a facility. All the demands of the nodes in a time period can only be met from the operational facilities in that time period.
- Facilities are divided into levels/categories with respect to their capacities and there can only be one facility in a node throughout the time horizon, regardless of its level.
- Once their constructions are completed, facilities and links will be operational until the end of the time horizon.
- Links have a fixed capacity and can serve a limited amount of demand.
- Opening facilities and links are not instantaneous; construction times are considered.
- There is a distance limit while constructing a link which means if the distance between two arbitrary nodes is above this limit, a link between these nodes cannot be activated.
- Links are directed and can only be constructed in one direction. The traveling costs between nodes are symmetrical.
- Demand is divisible. It is not necessary for a single facility to meet the whole demand of a node.

Table 3.2 introduces the notations used in Model A. Besides the decisions such as the locations of the facilities and links or the time periods of their construction, it is also taken into account in which periods the facilities and links are to be operational. For instance, if the construction of a facility has started but is yet to be completed, that facility is assumed to be non-operational. These kinds of facilities cannot serve any demand until their construction period has been completed.

One of the most fundamental properties of dynamic models is the variety of key parameters over time. For instance, the demand of the nodes or all corresponding costs may increase, may decrease or have a fluctuating pattern throughout the time horizon. In wastewater treatment cases, the demand is directly related to the population of the corresponding settlements. Thus, statistical/possibilistic techniques can be applied to forecast the uncertain parameters to model these possible fluctuations or varieties. While considering the construction costs or operating costs, the net present value (NPV) or net present worth (NPW) method could

Table 3.2. Notations used in Model A

Notation	Description
Sets	
$\mathcal{N}$	Set of nodes; $i, j \in \mathcal{N}$ and set of demand points; $k \in \mathcal{N}$
$\mathcal{U}$	Set of possible directed links; $(i, j) \in \mathcal{U}$: $dist_{ij} \leq \upsilon$
$\mathcal{L}$	Set of facility levels; $l \in \mathcal{L}$
$\mathcal{T}$	Set of time periods; $t \in \mathcal{T}$
Parameters	
d_k^t	Demand related to demand point k at time period t
$dist_{ij}$	Distance between node i and node j
f_i^{lt}	Cost of constructing a facility with level l on node i at time period t
c_{ij}^t	Cost of constructing link (i, j) at time period t
tc_{ij}^t	Traveling cost per demand unit on link (i, j) at time period t
g_i^{lt}	Operating and maintenance cost of a facility with level l on node i at time period t
h_{ij}^t	Operating and maintenance cost of link (i, j) at time period t
τ^l	Construction time for a facility with level l
δ	Construction time for a link
CF^l	Capacity of a facility with level l
CL	Capacity of a link
υ	Distance limit for constructing a link
Decision Variables	
Z_i^{lt}	1 if a facility with level l is located on node i at time period t, 0 otherwise
ZOP_i^{lt}	1 if a facility with level l is operational on node i at time period t, 0 otherwise
X_{ij}^t	1 if link (i, j) is constructed at time period t, 0 otherwise
XOP_{ij}^t	1 if link (i, j) is operational at time period t, 0 otherwise
Y_{ij}^{kt}	Fraction of the demand related to demand point k that flows on link (i, j) at time period t
W_i^{kt}	Fraction of the demand related to demand point k served by a facility on node i at time period t

be helpful in evaluating and comparing capital projects or financial products with cash flows spread over time [113]. Therefore, it can be used to calculate the current total value of a future stream of payments. Given the time period t, net cash flow in that period R_t and the discount rate r; the formula for NPV can be given as follows, where n represents the total number of time periods:

$$NPV = \sum_{t=1}^{n} \frac{R_t}{(1+r)^t} \tag{3.12}$$

In accordance with the assumptions made and the notations used in Table 3.2, the formulation of Model A is as follows:

$$\textbf{min.} \quad \sum_{i\in\mathcal{N}}\sum_{l\in\mathcal{L}}\sum_{t\in\mathcal{T}} f_i^{lt} Z_i^{lt} + \sum_{i\in\mathcal{N}}\sum_{l\in\mathcal{L}}\sum_{t\in\mathcal{T}} g_i^{lt} ZOP_i^{lt} + \sum_{(i,j)\in\mathcal{U}}\sum_{t\in\mathcal{T}} c_{ij}^t X_{ij}^t + \sum_{(i,j)\in\mathcal{U}}\sum_{t\in\mathcal{T}} h_{ij}^t XOP_{ij}^t + \sum_{(i,j)\in\mathcal{U}}\sum_{k\in\mathcal{N}}\sum_{t\in\mathcal{T}} tc_{ij}^t Y_{ij}^{kt} d_k^t \tag{3.13}$$

$$\textbf{subject to;} \quad ZOP_i^{lt} = \sum_{q=1}^{t-\tau^l} Z_i^{lq}; \qquad \forall i \in \mathcal{N}, \forall l \in \mathcal{L}, \forall t \in \mathcal{T} \tag{3.14}$$

$$ZOP_i^{lt} \geq ZOP_i^{l,t-1}; \qquad \forall i \in \mathcal{N}, \forall l \in \mathcal{L}, \forall t \in \mathcal{T} \tag{3.15}$$

$$XOP_{ij}^t = \sum_{q=1}^{t-\delta} X_{ij}^q; \qquad \forall (i,j) \in \mathcal{U}, \forall t \in \mathcal{T} \tag{3.16}$$

$$XOP_{ij}^t \geq XOP_{ij}^{t-1}; \qquad \forall (i,j) \in \mathcal{U}, \forall t \in \mathcal{T} \tag{3.17}$$

$$\sum_{j\in\mathcal{N}} Y_{ji}^{kt} = \sum_{j\in\mathcal{N}} Y_{ij}^{kt} + W_i^{kt}; \qquad \forall i,k \in \mathcal{N} : i \neq k, \forall t \in \mathcal{T} \tag{3.18}$$

$$W_i^{it} + \sum_{j\in\mathcal{N}} Y_{ij}^{it} = 1; \qquad \forall i \in \mathcal{N}, \forall t \in \mathcal{T} : d_i^t > 0 \tag{3.19}$$

$$\sum_{k\in\mathcal{N}} d_k^t W_i^{kt} \leq \sum_{l\in\mathcal{L}} CF^l \, ZOP_i^{lt}; \qquad \forall i \in \mathcal{N}, \forall t \in \mathcal{T} \tag{3.20}$$

$$\sum_{k\in\mathcal{N}} d_k^t Y_{ij}^{kt} \leq XOP_{ij}^t \, CL; \qquad \forall (i,j) \in \mathcal{U}, \forall t \in \mathcal{T} \tag{3.21}$$

$$\sum_{i\in\mathcal{N}} W_i^{kt} = 1; \qquad \forall k \in \mathcal{N}, \forall t \in \mathcal{T} : d_k^t > 0 \tag{3.22}$$

$$XOP_{ij}^t + XOP_{ji}^t \leq 1; \qquad \forall (i,j) \in \mathcal{U}, \forall t \in \mathcal{T} \tag{3.23}$$

$$\sum_{l\in\mathcal{L}} ZOP_i^{lt} \leq 1; \qquad \forall i \in \mathcal{N}, \forall t \in \mathcal{T} \tag{3.24}$$

$$Z_i^{lt}, ZOP_i^{lt} \in \{0,1\}; \qquad \forall i \in \mathcal{N}, \forall l \in \mathcal{L}, \forall t \in \mathcal{T} \tag{3.25}$$

$$X_{ij}^t, XOP_{ij}^t \in \{0,1\}; \qquad \forall (i,j) \in \mathcal{U}, \forall t \in \mathcal{T} \tag{3.26}$$

$$Y_{ij}^{kt} \geq 0; \qquad \forall (i,j) \in \mathcal{U}, \forall k \in \mathcal{N}, \forall t \in \mathcal{T} \tag{3.27}$$

$$W_i^{kt} \geq 0; \qquad \forall i,k \in \mathcal{N}, \forall t \in \mathcal{T} \tag{3.28}$$

The objective function (3.13) minimizes the total cost in the defined network, which consist of construction costs of the facilities and links, operating costs of the facilities and links as well as the traveling costs. Since the parameters may change over time in the specified time horizon, all corresponding costs are distinct for each time period. As described earlier, the NPV method can be used to define all corresponding costs. Therefore, the objective function minimizes the net present value of the total cost in the network over the specified time horizon.

Constraints (3.14) ensure that a facility with level l at node i will be operational τ^l periods after the initial investment decision is made. One may refer to the related assumption that facilities are not active and cannot serve any demand during their construction period. Constraints (3.15) ensure that once a facility with level l is located at node i at time period t, it will remain operational until the end of the time horizon. Constraints (3.16) ensure that the link (i, j) will be operational δ periods after the initial construction decision. Constraints (3.17) ensure that once the link (i, j) is constructed at time period t, it will remain operational until the end of the time horizon.

Constraints (3.18) ensure that each demand flows to node i is either served by node i to supply its demand or transferred to the other nodes for each time period t. Constraints (3.19) ensure that if node i has a demand at time period t, it must be served either from a facility at node i or by another facility in the network at time period t. Similar to the formulation of CFLNDP, Constraints (3.18) - (3.19), combined, provide the conservation of flow in each time period.

Constraints (3.20) - (3.21) are the capacity constraints. The facility capacity constraints (3.20) ensure that demand at time period t can be served by node i only if there is an operational facility at node i in that time period. Additionally, the total demand served by node i at time period t cannot exceed the capacity of the operational facility. The link capacity constraints (3.21) state that, in order to use the link (i, j) to serve the demand of demand point k at time period t, the link must be operational in that period and the total demand flows on that link cannot exceed its capacity.

Constraints (3.22) state that each demand must be served for each demand point throughout the time horizon. Constraints (3.23) preclude the activation of oppositely directed flows on any link (i, j), which means flow on any operational link will only be in one direction. Constraints (3.24) ensure that there can be only one facility in a node throughout the time horizon.

Constraints (3.25 - 3.28) correspond to the integrality and non-negativity constraints of the decision variables.

3.3. MODEL B

The idea of a dynamic/time-dependent system, which is ignored in most of the proposed facility location and network design problems, is examined in detail in Model A. Similarly to this situation, some other crucial points are often ignored for the wastewater treatment networks. The importance of distance in water transportation is an irrefutable fact. While considering the distance between two endpoints in water transportation, the vertical difference between these points, which means the difference in their altitudes, can also affect the transportation cost and difficulty. This fact may also contradict the assumption of symmetrical transportation cost. In that context, Model B aims to consider the transportation cost more realistically by defining three-dimensional space.

Moreover, the capacity expansion options for the facilities, which are categorized according to their capacities, are also taken into account. As discussed in Section 2.3, the capacity expansion phenomena is generally considered in continuous cases. In Model B, we consider the capacity expansion in a discrete structure. That is, the capacity expansion means an increase in facility levels which are defined as discrete integer values that affect the capacity.

Model B is defined on a three-dimensional space consisting of nodes and possible links between nodes and has a time horizon, as in Model A. The main assumptions of the model are as follows:

- Each node is a demand point and has a certain level of demand to be met for each time period throughout the time horizon.
- Each node is a candidate location for hosting a facility. All the demands of the nodes in a time period can only be met from the operational facilities in that time period.
- Facilities are divided into levels/categories with respect to their capacities and there can only be one facility in a node for each time period.
- A facility that is constructed at any node can be expanded to higher levels within the

time horizon. However, it cannot be decreased to a lower category or removed from the network.

- Links are divided into categories according to their capacities, just like facilities.
- There is a distance limit while constructing a link which means if the distance between two arbitrary nodes is above this limit, a link between these nodes cannot be activated.
- Opening facilities and links are not instantaneous; construction times are considered. Construction times depend on the categories of the facilities and links.
- Links are directed and can only be constructed in one direction.
- Demand is divisible. It is not necessary for a single facility to meet the whole demand of a node.

The notations used in the formulation of Model B are displayed in Table 3.3.

Table 3.3. Notations used in Model B

Notation	Description
Sets	
$\mathcal{N}$	Set of nodes; $i, j \in \mathcal{N}$ and set of demand points; $k \in \mathcal{N}$
$\mathcal{U}$	Set of possible directed links; $(i, j) \in \mathcal{U}$: $dist_{ij} \leq \upsilon$
$\mathcal{L}$	Set of facility levels; $l, \zeta \in \mathcal{L}$
$\mathcal{P}$	Set of link types; $p \in \mathcal{P}$
$\mathcal{T}$	Set of time periods; $t \in \mathcal{T}$
Parameters	
d_k^t	Demand related to demand point k at time period t
$dist_{ij}$	Distance between node i and node j
Δh_{ij}	Vertical difference between node i and node j
f_i^{lt}	Cost of constructing a facility with level l on node i at time period t
$e_i^{l\zeta t}$	Cost of expanding a facility with level l to level ζ on node i at time period t
c_{ij}^{pt}	Cost of constructing link (i, j) with link type p at time period t
tc_{ij}^{pt}	Traveling cost per demand unit on link (i, j) with link type p at time period t
g_i^{lt}	Operating and maintenance cost of a facility with level l on node i at time period t
h_{ij}^{pt}	Operating and maintenance cost of link (i, j) with link type p at time period t
τ^l	Construction time for a facility with level l
$\theta^{l\zeta}$	Expansion time for a facility with level l to level ζ
δ^p	Construction time for a link with type p
CF^l	Capacity of a facility with level l
CL^p	Capacity of a link with type p
υ	Distance limit for constructing a link
Decision Variables	
Z_i^{lt}	1 if a facility with level l is located on node i at time period t, 0 otherwise
ZOP_i^{lt}	1 if a facility with level l is operational on node i at time period t, 0 otherwise
$EX_i^{l\zeta t}$	1 if a facility with level l is expanded to level ζ on node i at time period t, 0 otherwise
X_{ij}^{pt}	1 if link (i, j) with type p is constructed at time period t, 0 otherwise
XOP_{ij}^{pt}	1 if link (i, j) with type p is operational at time period t, 0 otherwise
Y_{ij}^{kpt}	Fraction of the demand related to demand point k that flows on link (i, j) with type p at time period t
W_i^{kt}	Fraction of the demand related to demand point k served by a facility on node i at time period t

The formulation of Model B, taking into account the notations used and all assumptions made, is given below:

$$\textbf{min.} \quad \sum_{i\in\mathcal{N}}\sum_{l\in\mathcal{L}}\sum_{t\in\mathcal{T}} f_i^{lt} Z_i^{lt} + \sum_{i\in\mathcal{N}}\sum_{l\in\mathcal{L}}\sum_{t\in\mathcal{T}} g_i^{lt} ZOP_i^{lt} + \sum_{(i,j)\in\mathcal{U}}\sum_{p\in\mathcal{P}}\sum_{t\in\mathcal{T}} c_{ij}^{pt} X_{ij}^{pt}$$
$$+ \sum_{(i,j)\in\mathcal{U}}\sum_{p\in\mathcal{P}}\sum_{t\in\mathcal{T}} h_{ij}^{pt} XOP_{ij}^{pt} + \sum_{(i,j)\in\mathcal{U}}\sum_{k\in\mathcal{N}}\sum_{p\in\mathcal{P}}\sum_{t\in\mathcal{T}} tc_{ij}^{pt} Y_{ij}^{kpt} d_k^t \tag{3.29}$$
$$+ \sum_{i\in\mathcal{N}}\sum_{l\in\mathcal{L}}\sum_{\zeta\in\mathcal{L}:\zeta>l}\sum_{t\in\mathcal{T}} e_i^{l\zeta t} EX_i^{l\zeta t}$$

subject to;

$$ZOP_i^{lt} = \sum_{q=1}^{t-\tau^l} Z_i^{lq} + \sum_{q=1}^{t-\theta^{\zeta l}}\sum_{\zeta=1}^{l-1} EX_i^{\zeta lt} - \sum_{q=1}^{t-\theta^{l\zeta}}\sum_{\zeta=l+1}^{|L|} EX_i^{l\zeta t}; \qquad \forall i \in \mathcal{N}, \forall l \in \mathcal{L}, \forall t \in \mathcal{T} \tag{3.30}$$

$$ZOP_i^{lt} \geq \sum_{\zeta=l+1}^{|L|} EX_i^{l\zeta t}; \qquad \forall i \in \mathcal{N}, \forall l \in \mathcal{L}, \forall t \in \mathcal{T} \tag{3.31}$$

$$\sum_{l\in\mathcal{L}} ZOP_i^{lt} \geq \sum_{l\in\mathcal{L}} ZOP_i^{l,t-1}; \qquad \forall i \in \mathcal{N}, \forall t \in \mathcal{T} \tag{3.32}$$

$$XOP_{ij}^{pt} = \sum_{q=1}^{t-\delta^p} X_{ij}^{pq}; \qquad \forall (i,j) \in \mathcal{U}, \forall p \in \mathcal{P}, \forall t \in \mathcal{T} \tag{3.33}$$

$$XOP_{ij}^{pt} \geq XOP_{ij}^{p,t-1}; \qquad \forall (i,j) \in \mathcal{U}, \forall p \in \mathcal{P}, \forall t \in \mathcal{T} \tag{3.34}$$

$$\sum_{j\in\mathcal{N}}\sum_{p\in\mathcal{P}} Y_{ji}^{kpt} = \sum_{j\in\mathcal{N}}\sum_{p\in\mathcal{P}} Y_{ij}^{kpt} + W_i^{kt}; \qquad \forall i,k \in \mathcal{N} : i \neq k, \forall t \in \mathcal{T} \tag{3.35}$$

$$W_i^{it} + \sum_{j\in\mathcal{N}}\sum_{p\in\mathcal{P}} Y_{ij}^{ipt} = 1; \qquad \forall i \in \mathcal{N}, \forall t \in \mathcal{T} : d_i^t > 0 \tag{3.36}$$

$$\sum_{k\in\mathcal{N}} d_k^t W_i^{kt} \leq \sum_{l\in\mathcal{L}} CF^l\, ZOP_i^{lt}; \qquad \forall i \in \mathcal{N}, \forall t \in \mathcal{T} \tag{3.37}$$

$$\sum_{k\in\mathcal{N}} d_k^t Y_{ij}^{kpt} \leq XOP_{ij}^{pt}\, CL^p; \qquad \forall (i,j) \in \mathcal{U}, \forall p \in \mathcal{P}, \forall t \in \mathcal{T} \tag{3.38}$$

$$\sum_{i\in\mathcal{N}} W_i^{kt} = 1; \qquad \forall k \in \mathcal{N}, \forall t \in \mathcal{T} : d_k^t > 0 \tag{3.39}$$

$$\sum_{p\in\mathcal{P}}\sum_{t\in\mathcal{T}} (X_{ij}^{pt} + X_{ji}^{pt}) \leq 1; \qquad \forall (i,j) \in \mathcal{U} \tag{3.40}$$

$$\sum_{l\in\mathcal{L}} ZOP_i^{lt} \leq 1; \qquad \forall i \in \mathcal{N}, \forall t \in \mathcal{T} \tag{3.41}$$

$$Z_i^{lt}, ZOP_i^{lt} \in \{0,1\}; \qquad \forall i \in \mathcal{N}, \forall l \in \mathcal{L}, \forall t \in \mathcal{T} \tag{3.42}$$

$$X_{ij}^{pt}, XOP_{ij}^{pt} \in \{0,1\}; \qquad \forall (i,j) \in \mathcal{U}, \forall p \in \mathcal{P}, \forall t \in \mathcal{T} \tag{3.43}$$

$$EX_i^{l\zeta t} \in \{0,1\}; \qquad \forall i \in \mathcal{N}, \forall l, \zeta \in \mathcal{L} : \zeta > l, \forall t \in \mathcal{T} \tag{3.44}$$

$$Y_{ij}^{kpt} \geq 0; \qquad \forall (i,j) \in \mathcal{U}, \forall k \in \mathcal{N}, \forall p \in \mathcal{P}, \forall t \in \mathcal{T} \tag{3.45}$$

$$W_i^{kt} \geq 0; \qquad \forall i, k \in \mathcal{N}, \forall t \in \mathcal{T} \tag{3.46}$$

The objective function (3.29) minimizes the total cost in the defined network, which consist of construction costs of facilities and links, operating costs of facilities and links, expanding costs of facilities as well as traveling costs throughout the time horizon. Since the parameters may change over time in the specified time horizon, all corresponding costs are distinct for each time period as in Model A.

Constraints (3.30) enforce that, for a facility with level l to be operational in any time period t at node i, either a facility with level l should have been constructed at least τ^l periods ago, or a facility with any lower-level ζ should have been expanded to level l at least $\theta^{l\zeta}$ periods ago. In addition, the facility with level l should not have been expanded to a higher level. Constraints (3.31) ensure that a facility with level l should be operational to expand to a higher level at time period t. Constraints (3.32) ensure that once a facility with level l is located at node i, a facility with level ζ ($\zeta \geq l$) will be operational until the end of the time horizon. Constraints (3.33) ensure that a link (i, j) with type p will be operational δ^p periods after the initial construction decision. Constraints (3.34) ensure that once a link (i, j) with type p is established, it will remain operational until the end of the time horizon.

Constraints (3.35) ensure that each demand flows to node i is either served by node i to supply its demand or transferred to the other nodes for each time period t. These constraints have the same structure as Constraints (3.18) in Model A; only link types are included as a difference. Constraints (3.36) ensure that if node i has a demand at time period t, it must be served either from a facility at node i or by another facility in the network at time period t. As described earlier in Model A, Constraints (3.35) - (3.36), combined, provide the conservation of flow in each time period.

Constraints (3.37) - (3.38) are the capacity constraints and have the same structure as Constraints (3.20) - (3.21) in Model A. The facility capacity constraints (3.37) ensure that demand at time period t can be served by node i only if there is an operational facility at node

i in that time period. Additionally, the total demand served by node i at time period t cannot exceed the capacity of the operational facility. The link capacity constraints (3.38) state that, in order to use a link (i, j) to serve the demand of demand point k at time period t, the link must be operational in that period and the total demand flows on that link cannot exceed its capacity.

Constraints (3.39) state that each demand must be served for each demand point throughout the time horizon. Constraints (3.40) preclude the activation of oppositely directed flows on any link (i, j), which means flow on any operational link will only be in one direction. Constraints (3.41) ensure that there can be only one operational facility in any node for each time period.

Constraint (3.42 - 3.46) corresponds to the integrality and non-negativity constraints of the decision variables.

4. COMPUTATIONAL EXPERIMENTS

In this section, in order to reveal the structures of the proposed mathematical models, the computational experiments are presented and their experimental results are discussed. First, the parameters and conducted design of experiments for both proposed mathematical models are presented and then their results and competencies are discussed.

All test instances or experiments are solved with Python modeling library for optimization DOcplex and CPLEX version of 20.1. Experiments are run under a Windows system with a 2.8GHz processor and 16 GB of RAM.

4.1. EXPERIMENTAL RESULTS OF MODEL A

In this section, an experimental design is conducted for Model A in order to observe how the proposed mathematical model behaves with the change of set levels. Test instances are performed using the same method by Melkote and Daskin [56] using different set levels. A various number of nodes, time periods, facility levels and distance limits are used in the conducted design. The details of the experimental design with the enumeration of all combinations are given in Table 4.1. A total of forty-eight combinations are obtained by using four different number of nodes, three different number of time periods, two different facility levels and two different distance limits.

One of the most important issues to consider is that $|\mathcal{T}|$ refers to the cardinality of set $\mathcal{T}$ which indicates the total number of time periods in which the demand is taken into account. As mentioned earlier, there is a construction period for both facilities and links. Considering that the model ensures all the demand is served within the time horizon, a preliminary period is assumed for the initial investments. The length of this time interval is determined as the maximum construction time. Thus, in the remainder of this study, t^* indicates the starting point of the time horizon (i.e., pre-activation period) and $t = 1$ indicates the starting time of the demand activation. The time difference between t^* and $t = 1$ is the maximum construction time of any link or facility. Therefore, t^* might take on negative values.

Table 4.1. Details of the experimental design for Model A

Experiment No.		$\lvert\mathcal{L}\rvert = 3$		$\lvert\mathcal{L}\rvert = 5$	
		$\upsilon = 25$	$\upsilon = 50$	$\upsilon = 25$	$\upsilon = 50$
$\lvert\mathcal{N}\rvert = 10$	$\lvert\mathcal{T}\rvert = 10$	$E.A1$	$E.A2$	$E.A3$	$E.A4$
	$\lvert\mathcal{T}\rvert = 20$	$E.A5$	$E.A6$	$E.A7$	$E.A8$
	$\lvert\mathcal{T}\rvert = 30$	$E.A9$	$E.A10$	$E.A11$	$E.A12$
$\lvert\mathcal{N}\rvert = 20$	$\lvert\mathcal{T}\rvert = 10$	$E.A13$	$E.A14$	$E.A15$	$E.A16$
	$\lvert\mathcal{T}\rvert = 20$	$E.A17$	$E.A18$	$E.A19$	$E.A20$
	$\lvert\mathcal{T}\rvert = 30$	$E.A21$	$E.A22$	$E.A23$	$E.A24$
$\lvert\mathcal{N}\rvert = 30$	$\lvert\mathcal{T}\rvert = 10$	$E.A25$	$E.A26$	$E.A27$	$E.A28$
	$\lvert\mathcal{T}\rvert = 20$	$E.A29$	$E.A30$	$E.A31$	$E.A32$
	$\lvert\mathcal{T}\rvert = 30$	$E.A33$	$E.A34$	$E.A35$	$E.A36$
$\lvert\mathcal{N}\rvert = 40$	$\lvert\mathcal{T}\rvert = 10$	$E.A37$	$E.A38$	$E.A39$	$E.A40$
	$\lvert\mathcal{T}\rvert = 20$	$E.A41$	$E.A42$	$E.A43$	$E.A44$
	$\lvert\mathcal{T}\rvert = 30$	$E.A45$	$E.A46$	$E.A47$	$E.A48$

The parameters used in the experimental design are presented in Table 4.2. The locations of the nodes are randomly generated on a 50×50 two-dimensional plane by using the Python library NumPy [114]. The same random seed is used to locate the nodes in each experiment to perform a valid comparison. For instance, the locations of the nodes in the first twelve experiments ($E.A1$ - $E.A12$) are the same. For the next twelve experiments ($E.A13$ - $E.A24$), instead of locating twenty nodes from scratch, the locations of the first ten nodes are kept as it stands and then newly generated nodes are included.

The distance between nodes is determined by Euclidean geometry. Let $(locx_i, locy_i)$ represent the coordinates of node i in two-dimensional space. The distance between node i and node j is determined by Equation (4.1) given below:

$$dist_{ij} = \sqrt{(locx_i - locx_j)^2 + (locy_i - locy_j)^2} \tag{4.1}$$

While presenting the parameters, the uniform distribution is used to determine the associated costs and capacities. $U[a, b]$ denotes a uniformly distributed random variable that lies between certain lower bound a and upper bound b.

Table 4.2. Parameters used in experimental design of Model A

Parameter	**Value Generation**
d_k^t	$d_k^1 = U[5, 100]$ $d_k^t = d_k^{t-1} \times U[0.99, 1.05];\ t > 1$
f_i^{lt}	$f_i^{10^*} = U[400, 450]$ $f_i^{20^*} = U[600, 700]$ $f_i^{30^*} = U[900, 1000]$ $f_i^{40^*} = U[1400, 1600]$ $f_i^{50^*} = U[2250, 2500]$ $f_i^{lt} = f_i^{lt-1} \times U[1, 1.05];\ t > 0^*$
c_{ij}^t	$c_{ij}^{0^*} = dist_{ij} \times 0.1$ $c_{ij}^t = c_{ij}^{t-1} \times U[1, 1.05];\ t > 0^*$
tc_{ij}^t	$tc_{ij}^{0^*} = dist_{ij} \times 0.005$ $tc_{ij}^t = tc_{ij}^{t-1} \times U[1, 1.05];\ t > 0^*$
g_i^{lt}	$g_i^{lt} = f_i^{lt} \times U[0.06, 0.1]$
h_{ij}^t	$h_{ij}^t = c_{ij}^t \times U[0.04, 0.06]$
τ^l	$\tau^l = 2;\ l \leq 3$ $\tau^l = 3;\ l \geq 4$
δ	$\delta = 1$
CF^l	$CF^1 = 25$ $CF^2 = 45$ $CF^3 = 80$ $CF^4 = 150$ $CF^5 = 250$
CL	$CL = 120$

In each experiment, a time limit is set on the CPLEX algorithm, and if an optimal solution cannot be obtained within this limit, the algorithm is interrupted and the best solution obtained by that time is recorded. The specified time limit for each experiment is determined by the formula given below:

$$\text{Time Limit} = |\mathcal{N}| \times |\mathcal{T}| \times 30 \tag{4.2}$$

Table 4.3 summarizes the computational results of the experimental design by presenting the objective function value, lower bound, the percentage of optimality gap, CPU time and the weighted average efficiency of the facilities; for each test instance. The weighted average

Table 4.3. Computational results of the experimental design for Model A

Experiment No.	Objective Function	Lower Bound	Optimality Gap(%)	CPU time (sec.)	Weighted Avg. Efficiency(%)
E.A1	10219.91	10219.91	0.00	14.29	98.14
E.A2	10219.91	10219.91	0.00	29.54	98.14
E.A3	10219.91	10219.91	0.00	10.14	98.14
E.A4	10219.91	10219.91	0.00	25.68	98.14
E.A5	20304.72	20304.72	0.00	121.45	98.10
E.A6	20311.08	20311.08	0.00	283.63	98.10
E.A7	20175.77	20175.77	0.00	99.56	98.68
E.A8	20170.48	20170.48	0.00	240.93	98.68
E.A9	37548.81	37548.81	0.00	727.25	99.45
E.A10	37548.81	37548.81	0.00	3078.72	99.45
E.A11	36797.33	36797.33	0.00	189.47	99.94
E.A12	36797.33	36797.33	0.00	1012.86	99.94
E.A13	18628.73	18628.73	0.00	4729.58	99.45
E.A14	18628.73	18564.60	0.34	6000.00	99.45
E.A15	18624.88	18440.12	0.99	6000.00	98.34
E.A16	18629.15	18311.98	1.70	6000.00	98.34
E.A17	36785.94	36299.15	1.32	12000.00	99.47
E.A18	37180.71	36188.16	2.67	12000.00	98.25
E.A19	35979.69	35543.25	1.21	12000.00	99.47
E.A20	36028.12	35544.63	1.34	12000.00	99.47
E.A21	68348.44	66960.40	2.03	18000.00	99.22
E.A22	68460.45	66727.99	2.53	18000.00	99.22
E.A23	66575.36	64980.48	2.40	18000.00	99.22
E.A24	66397.89	64849.13	2.33	18000.00	99.22
E.A25	27475.96	27230.00	0.90	9000.00	99.14
E.A26	27475.49	27222.71	0.92	9000.00	99.14
E.A27	27334.21	26735.78	2.19	9000.00	98.64
E.A28	27097.84	26718.74	1.40	9000.00	99.89
E.A29	54751.95	53485.05	2.31	18000.00	99.69
E.A30	55490.46	53485.80	3.61	18000.00	98.88
E.A31	53761.29	52250.04	2.81	18000.00	98.68
E.A32	51980.52	51088.42	1.75	18000.00	99.13
E.A33	103278.37	98751.75	4.58	27000.00	97.56
E.A34	103997.59	99012.73	5.03	27000.00	98.56
E.A35	100728.14	95474.04	5.50	27000.00	98.06
E.A36	98130.59	94291.63	4.07	27000.00	99.10
E.A37	36531.42	36116.29	1.15	12000.00	98.37
E.A38	36531.42	36078.36	1.26	12000.00	98.37
E.A39	36041.39	35152.42	2.53	12000.00	99.70
E.A40	35783.91	34958.35	2.36	12000.00	99.70
E.A41	74839.26	70894.83	5.56	24000.00	97.87
E.A42	73845.69	69968.93	5.54	24000.00	98.46
E.A43	73627.30	70152.45	4.95	24000.00	98.63
E.A44	72261.44	68118.05	6.08	24000.00	98.14
E.A45	134475.32	130580.91	2.98	36000.00	98.12
E.A46	133569.88	127293.66	5.00	36000.00	98.10
E.A47	129623.25	121343.69	6.82	36000.00	98.75
E.A48	125292.27	117471.73	6.66	36000.00	99.10

efficiency of the facilities is determined by Equation (4.3) given as follows:

$$\text{Weighted Avg. Efficiency (\%)} = \frac{\sum_{k \in \mathcal{N}} \sum_{t \in \mathcal{T}} d_k^t}{\sum_{i \in \mathcal{N}} \sum_{l \in \mathcal{L}} \sum_{t \in \mathcal{T}} CF^l \ ZOP_i^{lt}} \times 100 \tag{4.3}$$

The percentage of optimality gap simply shows the deviation of the incumbent solution, which is the best-known solution, from the best possible solution. The lower bound obtained by the CPLEX algorithm refers to the best possible solution when the algorithm is interrupted. We see that in about one-fourth of the experiments (13 out of 48), an exact optimal solution is found within the specified time limit. Solutions within five percent of the optimality gap are found in 40 out of 48 experiments. Moreover, the weighted average efficiency of the facilities is obtained to be higher than 98 percent in almost all experiments.

Considering the main motivations of this study, the fact that the weighted average efficiency is way higher than 95 percent and the optimality gap is less than 10 percent within the specified time limit in all experiments is a good performance indicator for the proposed model; even though the number of decision variables and constraints has increased considerably due to the dynamic nature of the model. However, as stated in the related studies, the performance of CPLEX might decrease significantly as the scale of the problems increases [55,56,64]. For instance, CPLEX runs out of memory and cannot reach to a lower bound in a test problem conducted by Rahmaniani and Ghaderi [64], in which there are 80 nodes for a CFLND problem. Nevertheless, the CPLEX algorithm has been able to solve all test networks with up to 40 nodes, effectively.

To present an example, the complete network in a single experiment $E.A3$ is illustrated in Figure 4.1. Nodes without facilities are represented by dots and nodes with constructed facilities are represented by squares. In each square, there is a number that denotes the facility level l. Moreover, the arrows represent established links along with their directions. It is observed that, the demand of node 5 is divided partially between the facilities on nodes 3 and 4; one may recall the demand divisibility structure of Model A. The fraction of this division, W_i^{kt} for demand point $k = 5$ throughout the time horizon, is summarized in Table 4.4. Another interesting point of the experiment is the establishment of link (3,4), which is between two nodes with facilities. The capacity of the facility at node 9 is lower than the total demand of nodes 1, 9 and 10 combined. Therefore, part of this combined demand flows over the link (3,4) to be served at the facility at node 7. The details of this particular experiment are provided in Appendix A by displaying the network and decision variables for each time period.

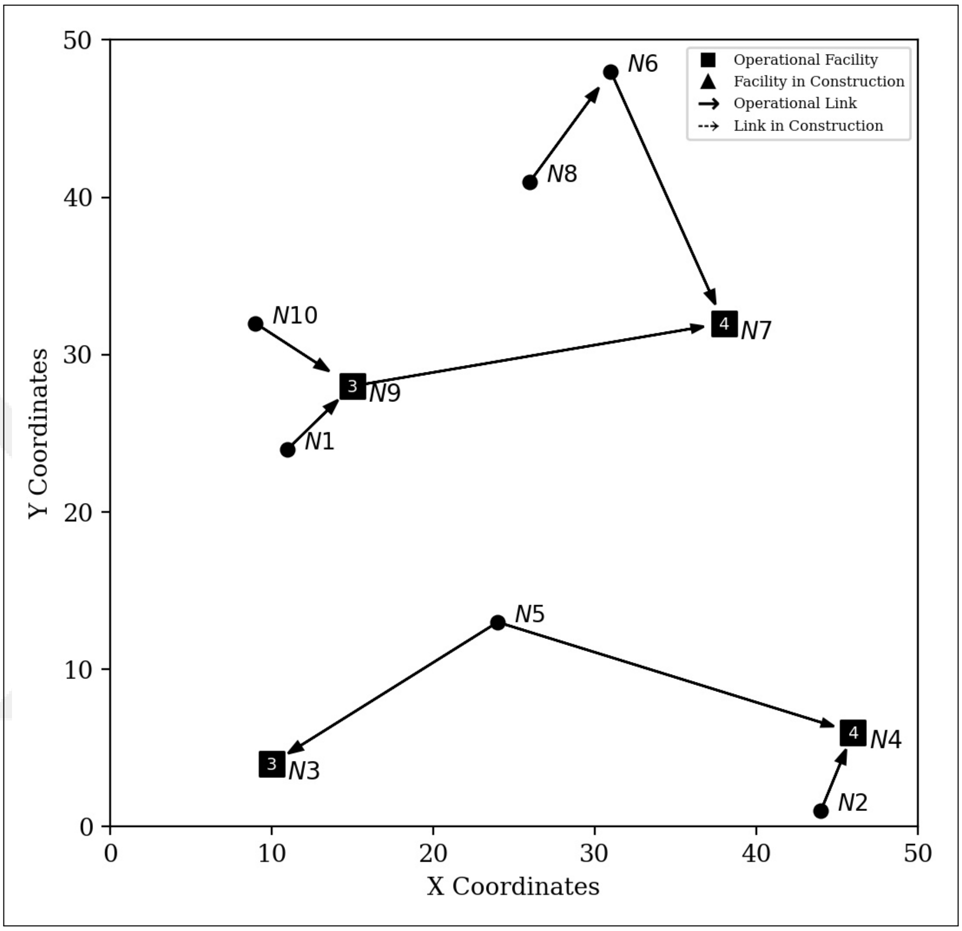


Figure 4.1. Complete network in $E.A3$ (t = 10)

Table 4.4. W_i^{kt} for demand point k=5 in $E.A3$

W_i^{kt}		**Time Period (t)**									
Source (i)	**Demand Point (k)**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
3	5	1.00	1.00	1.00	1.00	1.00	0.91	0.83	0.79	0.78	0.79
4	5						0.09	0.17	0.21	0.22	0.21

Figures 4.2 - 4.9 contain bar plots for the output of the experimental design for each set level combination. More specifically, Figure 4.2 presents the objective function values (i.e., total cost) for each experiment. It is clear that, the increase in $|\mathcal{N}|$ and $|\mathcal{T}|$ triggers an increase in the total cost, which is an expected behaviour in line with the expansion of whole network.

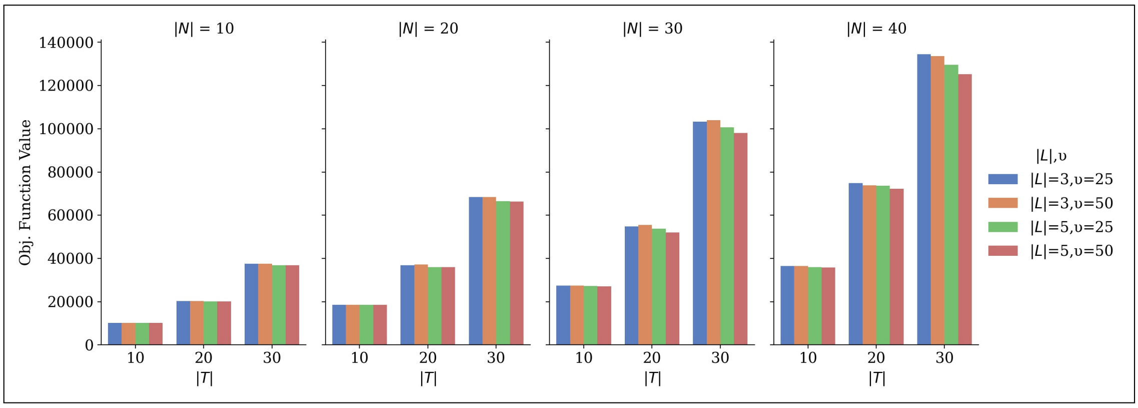


Figure 4.2. Effect of different set levels on total cost for the experimental design of Model A

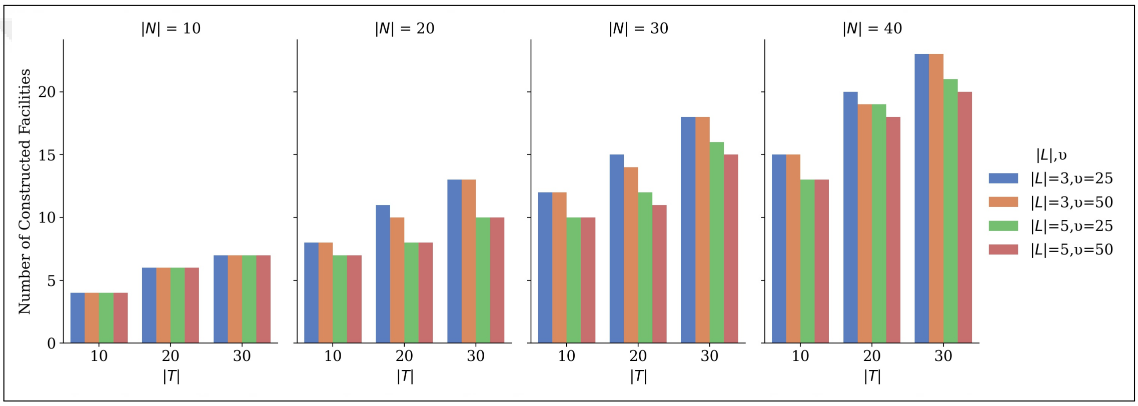


Figure 4.3. Effect of different set levels on number of constructed facilities for the experimental design of Model A

The expected positive trend that defined for demand generation often causes more constructed facilities and links as the time horizon expands (see Figure 4.3 and Figure 4.6); thus more investment is made (see Figure 4.4, Figure 4.5, Figure 4.7, Figure 4.8).

It is also observed that, as $|\mathcal{L}|$ and υ decreases, facility-related costs tend to increase (see Figure 4.4, Figure 4.5), whereas link investments and traveling costs tend to decrease (see Figure 4.7, Figure 4.8, Figure 4.9). The main reason for this situation is the restriction on the possible combinations of decision variables. For example, further restrictions on the distance limit will reduce the number of candidate links to be established. Hence, part of the links involved in the solution of experiments where υ is bigger cannot be established. Therefore, a facility may be constructed instead of establishing that type of link or a group of links. The same holds true while reducing the number of possible facility levels, as it will affect the capacity of the whole network.

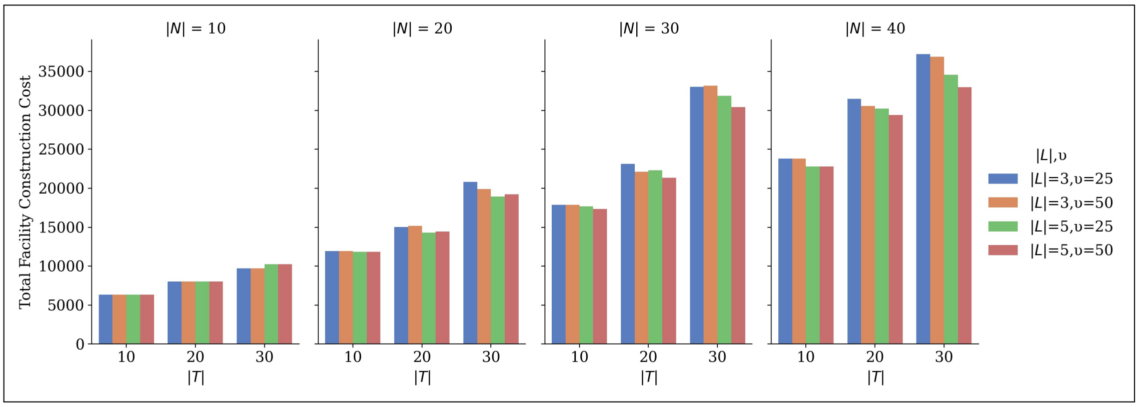


Figure 4.4. Effect of different set levels on total facility construction cost for the experimental design of Model A

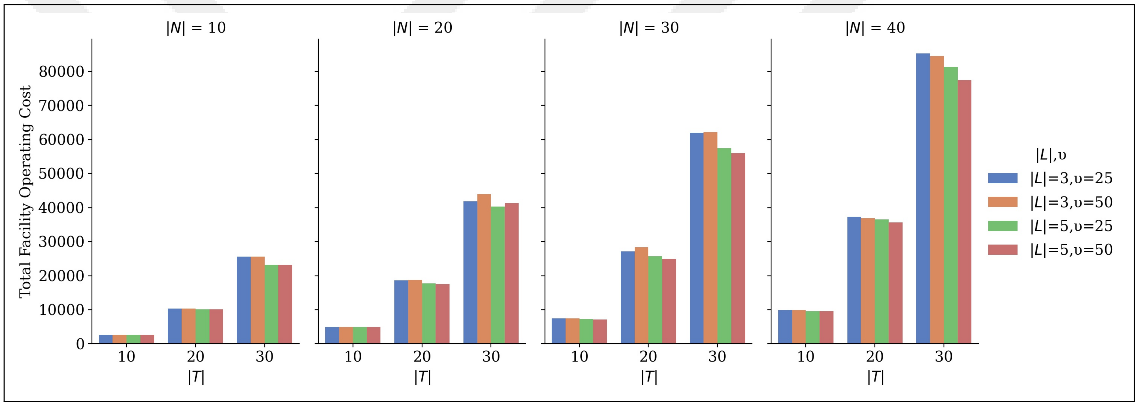


Figure 4.5. Effect of different set levels on total facility operating cost for the experimental design of Model A

Figure 4.5 clearly shows the trade-off between $|\mathcal{T}|$ and the facility operating costs. It is observed that, growth in the time horizon increases facility operating costs significantly more than construction costs. The same holds true for the link investments (see Figure 4.7, Figure 4.8). This situation directly depends on the demand growth rate over time. Since all the demand in the network must be met, the total capacity must be at least as much as the combined demand in the whole network. In scenarios where demand remains stable over the years, the number of facilities built is not expected to increase dramatically. On the other hand, since operational costs are considered in each time period, they are expected to increase as the time horizon expands. The same applies to the total traveling costs as they are considered when transferring and flow occurs in the network.

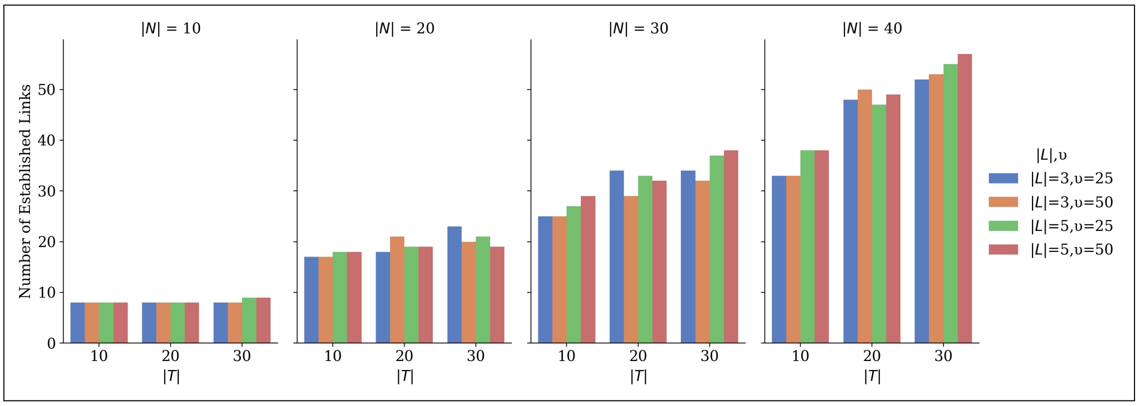


Figure 4.6. Effect of different set levels on number of established links for the experimental design of Model A

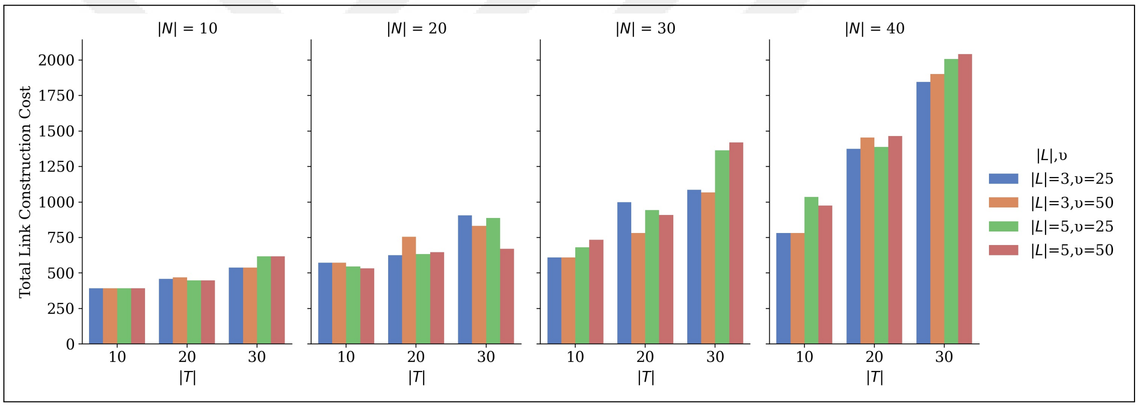


Figure 4.7. Effect of different set levels on total link construction cost for the experimental design of Model A

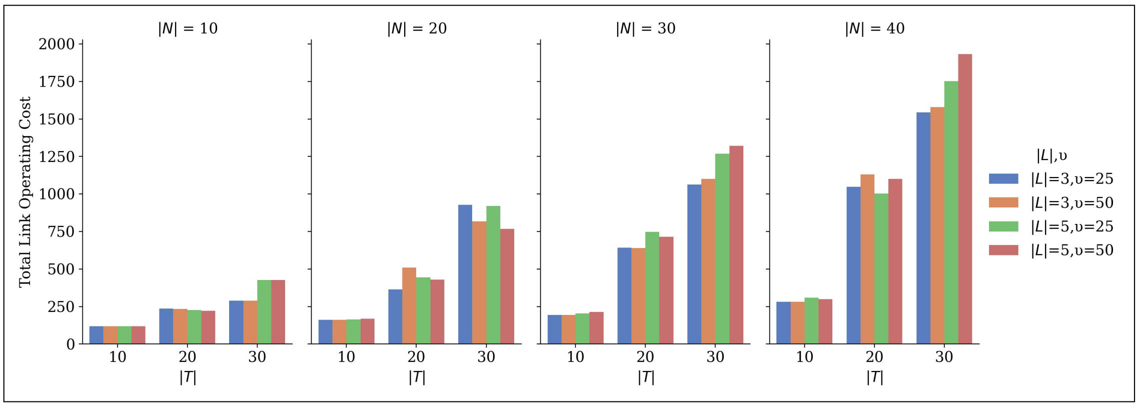


Figure 4.8. Effect of different set levels on total link operating cost for the experimental design of Model A

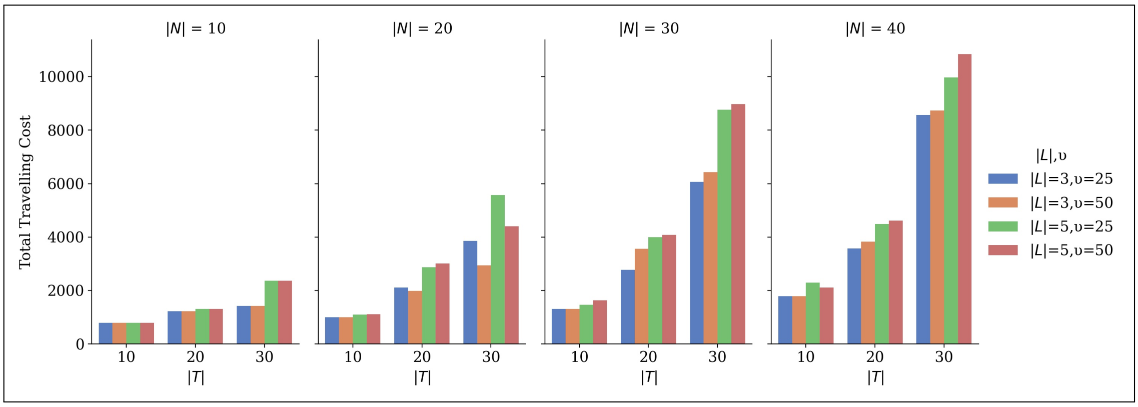


Figure 4.9. Effect of different set levels on total traveling cost for the experimental design of Model A

4.2. EXPERIMENTAL RESULTS OF MODEL B

4.2.1. A Guideline for Parameter Generation

The proposed model is designed to establish a wastewater treatment network; thus, the cost and energy formulas derived in the literature, especially those derived by Brand and Osfeld [109] can be used for parameter generation to obtain more realistic results. For instance, the formula for the facility construction cost can be given as follows:

$$C_1 = 85825\ (Q_T)^{0.71} + 1000\ Q_T \tag{4.4}$$

In Equation (4.4), C_1 represents the cost for constructing a wastewater treatment facility in dollars ($). Q_T variable represents the amount of wastewater treated in that facility in $m^3/hour$. In other words, it represents the capacity of the facility to be constructed. This equation can be used as the cost of opening a facility within the determined capacities of the categorized facilities.

$$C_2 = 382.5\ D_p^{1.455}\ L \tag{4.5}$$

Equation (4.5) is also derived by Brand and Osfeld [109] in which the term C_2 expresses the cost in $ required to establish a wastewater pipeline. The variable D_p represents pipeline diameters in cm, whereas the variable L represents the pipeline length in km. This equation can also be used as the cost of establishing categorized links. In this context, the following formula can be used in order to establish a relation between the diameters and capacities of the pipelines for a wastewater treatment network:

$$Q_W = 3600\,\pi \cdot V(\frac{d}{2})^2 \tag{4.6}$$

In Equation (4.6), the variable Q_W gives the flow rate of water in $m^3/hour$; the variable V represents the velocity of water in $m/second$ and the variable d represents the diameter of the pipeline through the water will be transported in m.

$$C_3 = \frac{EC \times HR}{1000}\,[3.454\,\Delta_h\,Q + 6409\,(Q^{2.852}\,D_p^{-4.87}\,L)] \tag{4.7}$$

Equation (4.7) is derived to obtain the annual energy cost for pipelines in a wastewater treatment network. C_3 represents the annual pumping cost in dollars ($/year), whereas Δ_h represents the vertical distance difference between the endpoints of a pumping line in m. EC is the energy cost in $\$/kWh$ and HR is the number of operational hours of a pumping line in a year. As a base value, EC and HR are determined to be 0.1 $\$/kWh$ and 8760 $hour/year$, respectively. The variable Q denotes the flow rate in a pumping line in $m^3/hour$. Equation (4.7) can be used as traveling cost in Model B.

One may recall that, the total demand that flows on any link is a decision variable, which is represented by Q in Equation (4.7). Linear programming problems assume that each decision variable has a power of one (i.e., they cannot be raised to any power different than one). Therefore, because of the $Q^{2.852}$ term in the formula, using it directly would violate the linearity assumption. This term can be transformed into a piecewise linear function to avoid that situation.

The maximum capacity of a link is determined as 4250 $m^3/hour$, as will be revealed later in this section. By using the Python library "pwlf 2.2.1", which is a library for applying piecewise linear regression to a given data, the function $g(Q)$ is fitted to the function $Q^{2.852}$ in a given interval [1,4250] of Q. While applying this fit, $g(Q)$ is forced to pass through the point (1,1) and the slopes of this function are forced to be a set of simplified integers to reduce the complexity. With an R^2 value of 0.994, which indicates a good fit, the piecewise linear function $g(Q)$ is given below:

$$g(Q) = \begin{cases} 0; & Q < 1 \\ 500,000\,Q - 499,999; & 1 \leq Q \leq 1500 \\ 5,000,000\,Q - 67,505 \times 10^5; & 1500 \leq Q \leq 3000 \\ 10,000,000\,Q - 217,505 \times 10^5; & 3000 \leq Q \leq 4250 \\ 0; & Q > 4250 \end{cases} \tag{4.8}$$

Thus, $Q^{2.852}$ is transformed to $g(Q)$ in Equation (4.7) that involves in the objective function of Model B. The visual representation of this fit is presented in Figure 4.10.

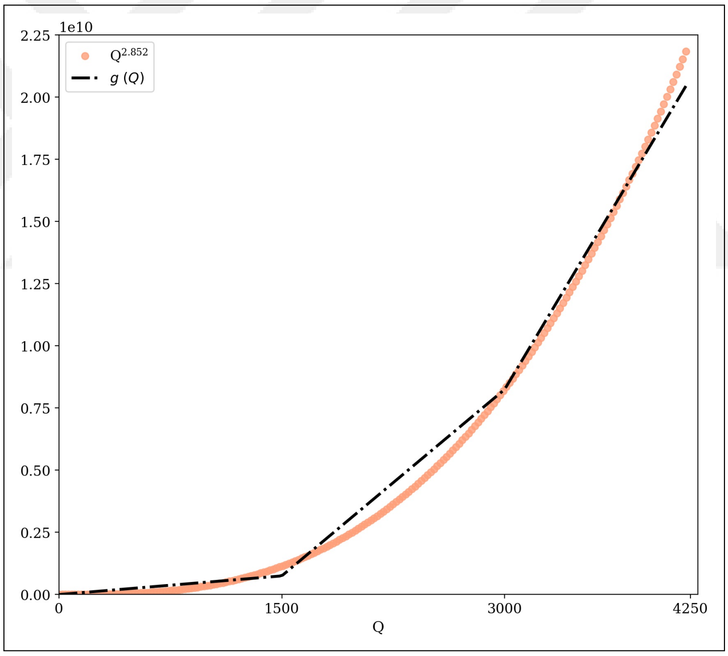


Figure 4.10. Piecewise linear fit of $g(Q)$

In this perspective, Equation (4.7) with the linear transformation can be included in the objective function by using the method that Winston and Goldberg [115] provided for using these functions in linear programming models. To briefly introduce their method with their notations; consider a piecewise linear function $h(x)$ with breakpoints $b_1, b_2, ..., b_n$. Let a number x is between the breakpoints b_k and b_{k+1} and a coefficient z_k is on the interval $[0, 1]$.

Then the value of x may be written as follows:

$$x = z_k b_k + (1 - z_k) b_{k+1} \tag{4.9}$$

Given that the function $h(x)$ is linear on the interval $[b_k, b_{k+1}]$, the following equation is valid:

$$h(x) = z_k \; h(b_k) + (1 - z_k) \; h(b_{k+1}) \tag{4.10}$$

From this point of view, if the function $h(x)$ is replaced by the term $z_1 \; h(b_1) + z_2 \; h(b_2) + \ldots + z_n \; h(b_n)$ and the following constraints are included to the problem; $h(x)$ can be used in the objective function:

$$z_1 \leq y_1, z_2 \leq y_1 + y_2, z_3 \leq y_2 + y_3, \ldots, z_{n-1} \leq y_{n-2} + y_{n-1}, z_n \leq y_{n-1} \tag{4.11}$$

$$y_1 + y_2 + \ldots + y_{n-1} = 1 \tag{4.12}$$

$$z_1 + z_2 + \ldots + z_n = 1 \tag{4.13}$$

$$x = z_1 b_1 + z_2 b_2 + \ldots + z_n b_n \tag{4.14}$$

$$y_i \in \{0, 1\} \; (i = 1, 2, \ldots, n - 1); \; z_i \geq 0 \; (i = 1, 2, \ldots, n) \tag{4.15}$$

As will be revealed in Section 5, a group of experts is consulted to validate the model. In these consultations, it is assumed that the flow of wastewater from a higher altitude location to a lower one by gravity flow with its own energy is not the same as a flow in the opposite direction. Wastewater can be transported from a higher altitude location to a lower one without any interruption, simply by providing the necessary slope for the water flow. However, it is necessary to pressurize the water through pumping systems in the opposite direction. That is an issue that brings extra costs to the wastewater treatment plants. Therefore, while defining the vertical difference between node i and node j, Δh_{ij}; if the altitude of node i is higher than the altitude of node j, the difference is omitted. Let $loch_i$ represents the altitude level of node i; Δh_{ij} is formulated as follows:

$$\Delta h_{ij} = \max\{(loch_j - loch_i), \; 0\} \tag{4.16}$$

Considering the third dimension, the distance between nodes, which is provided in Equation (4.1), is modified. The Euclidean distance in three-dimensional space is given as follows:

$$dist_{ij} = \sqrt{(locx_i - locx_j)^2 + (locy_i - locy_j)^2 + (loch_i - loch_j)^2} \tag{4.17}$$

As described in Section 3.3, the facilities are categorized with respect to their capacities. In this study, we determine five levels of facilities. In order to generate a capacity for each facility level as reliable and realistic as possible, the data on urban wastewater treatment plant capacities of Turkey are collected from a comprehensive report published by the Turkish Ministry of Environment, Urbanization and Climate Change [29]. All the capacities are converted to $m^3/hour$ and then sorted in ascending order to specify their percentiles. Then, the data are divided into five equal intervals with respect to their percentiles and the median value of each interval (i.e., 10^{th} percentile, 30^{th} percentile, 50^{th} percentile, 70^{th} percentile, 90^{th} percentile, respectively) is determined as the capacity of each facility level. The median values are rounded to close integer values to simplify the categorization. As demonstrated in Figure 4.11, which illustrates the histogram of capacities of the urban wastewater treatment plants in Turkey, the capacity of each facility level is determined as 200, 600, 1400, 3500 and 10000 in $m^3/hour$, respectively. One may assume that, if the number of facility levels is changed in different experiments, the number of intervals and their median values can also be obtained by the same logic. For instance, if the number of facility levels is three, the median values of three equal intervals should be the values on the 17^{th} percentile, 50^{th} percentile and 83^{rd} percentile. In that case, the capacity of each level must be around 300, 1400 and 6000 $m^3/hour$, respectively.

While implementing the capacity expansion structure, a concave cost function that represents economies of scales of large expansion sizes summarized by Luss [116] can be used. In another study, a similar cost function, which is named "Power-Sizing Technique", is presented by Sullivan *et al.* [113]. Accordingly, the capacity expansion cost of a facility with a predetermined construction cost can be calculated as follows:

$$C_4 = C_I(\frac{S_E}{S_I})^{\alpha} \tag{4.18}$$

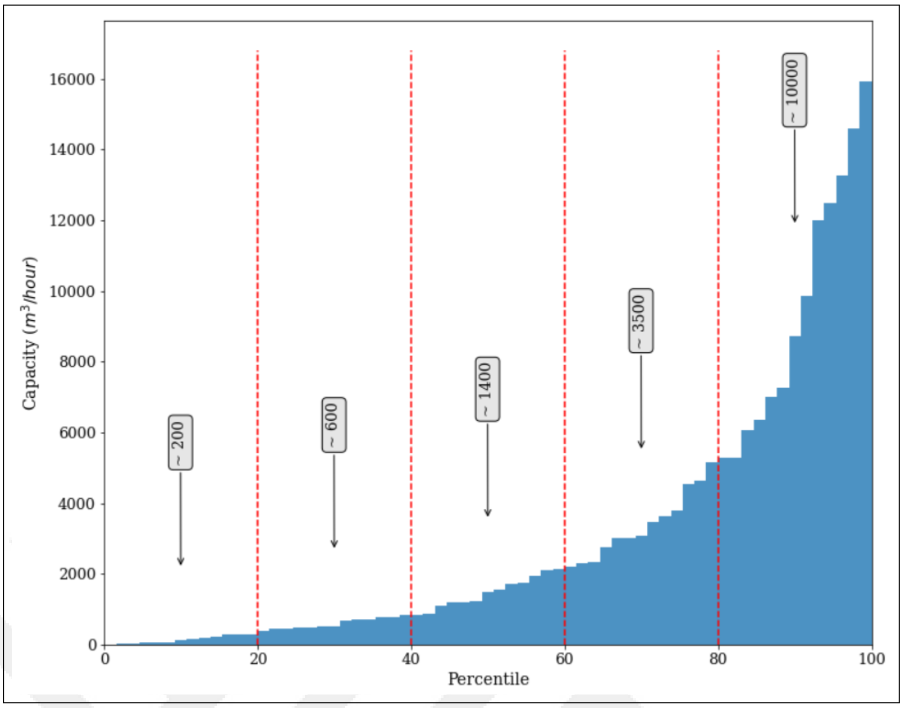


Figure 4.11. Histogram of capacities of the urban wastewater treatment plants in Turkey

C_4 denotes the cost of the capacity expansion, whereas C_I denotes the initial construction cost for the expanded facility. S_E represents the size of the expansion and S_I represents the size of the initial facility. In this method, the coefficient α represents the cost-capacity factor for economies of scale and may vary with respect to the type of facility or the sector it belongs to. In accordance with this, Hendrickson *et al.* [117] stated in their project management book that the coefficient α can be taken as 0.67 for wastewater treatment.

Annual facility operating costs are determined via the formulation provided by Öztürk [110], which is presented as a financial framework for wastewater treatment networks in his book about the details of wastewater engineering. For a treatment facility that constitutes advanced treatment, his formulation is given as follows:

$$C_5 = [0.0488\,(\frac{Q_f}{1000})^{0.8623}] \times 1,000,000 \tag{4.19}$$

C_5 denotes the annual operational cost of a wastewater treatment facility in dollars ($/year), whereas Q_f represents the amount of wastewater to be treated in that facility in m^3/day. His book also states that the annual operating cost of a pipeline is around 0.6 percent of the construction cost.

Initially, three pipeline types are determined in this study. In fact, different pipeline types are assumed to differ only in their diameters to specify their capacities; other characteristic structures are not considered. The diameter of each pipeline type are determined as 600 mm, 800 mm and 1000 mm, respectively. By using Equation (4.6) with a velocity of 1.5 $m/second$, which is the upper limitation for water-like fluids with settleable solids [118], the capacity of each type is obtained to be 1500, 2750 and 4250 $m^3/hour$, respectively.

The construction times of the facilities and links are the same as in the experimental design for Model A, which are described in Table 4.2. The expansion time from a facility with level l to level ζ ($\zeta > l$) has a structure that is defined in Algorithm 4.1.

Algorithm 4.1. Expansion time of the facilities

```
Require: l, ζ
  Exp. Time ← 0
  if l < ζ then
      if l + 2 < ζ then
          Exp. Time ← 2
      else
          Exp. Time ← 1
      end if
  else
      No Expansion!
  end if
  return Exp. Time
```

The population of each demand point is independently obtained by $U[50000, 500000]$. The average daily wastewater produced by a person, which is practically the demand of a person in this network, is taken as 0.15 m^3/day. As in the experiments of Model A, the fluctuating structure throughout the time horizon is included in the parameters.

4.2.2. Experimental Design

After specifying all the parameters, the third dimension is considered while generating the network since Model B considers the altitudes of the nodes. In order to define a more realistic network, instead of locating the nodes on an empty network and assigning an altitude value to them, nodes are assigned to a network with a predetermined vertical topography. More explicitly, nodes are located through two-dimensional coordinates which already have an altitude level. Thus, a more visually presentable network is created. In the remainder of this study, contour lines (i.e., isoline, isopleth) are used to represent the vertical topography of a network. They are used in drawing topographic maps to indicate ground elevation or depression. After locating 10 nodes with contour lines, a network is visualized in Figure 4.12.

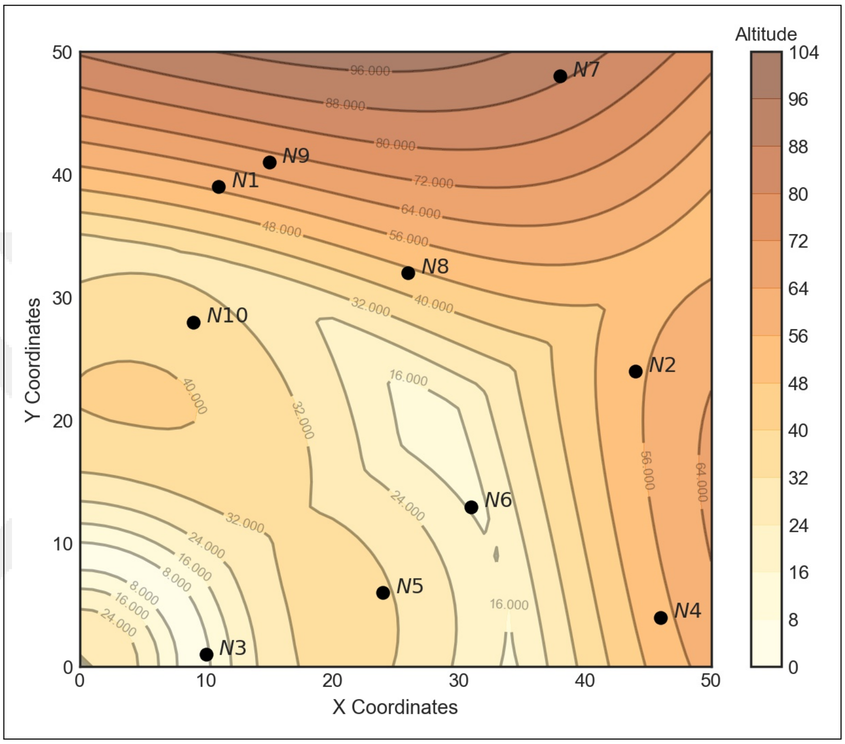


Figure 4.12. Illustrative example of a contour map with 10 nodes

As an alternative way, three-dimensional surface plots may be used. However, they are much more visually confusing than contour maps, as seen in Figure 4.13.

As in the experimental design for Model A, test instances are performed with the different number of set levels. Thus, a total of forty-eight combinations are obtained by using four different number of nodes, three different number of time periods, two different facility levels and two different distance limits. The details of the experimental design with the enumeration of all combinations are given in Table 4.5.

In each experiment, a time limit is set on the CPLEX algorithm, which is determined by Equation (4.2) in the previous section.

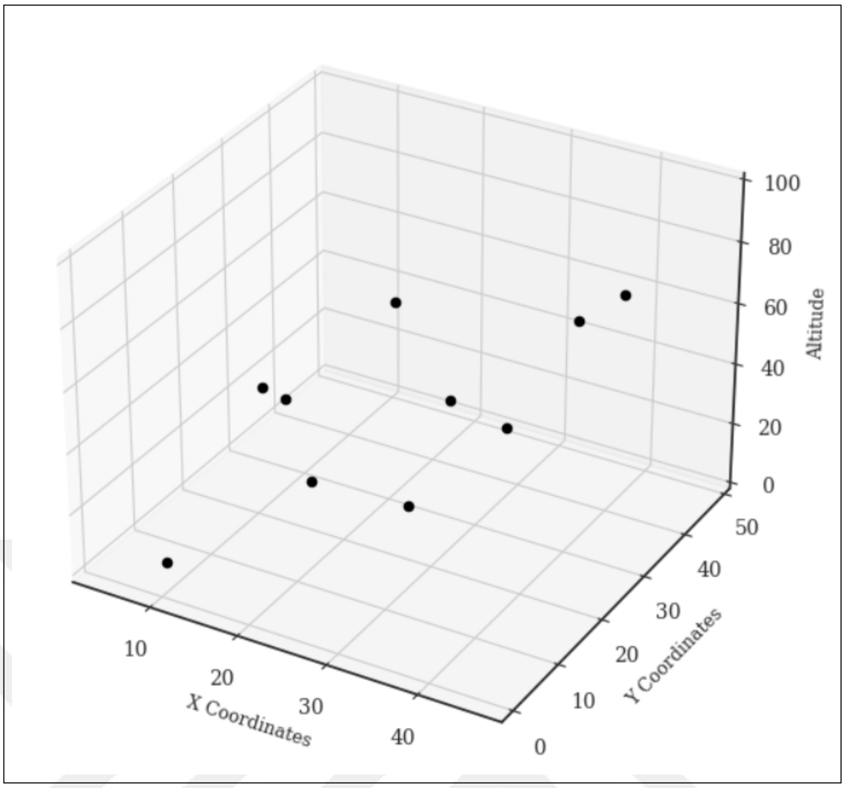


Figure 4.13. Illustrative example of a three-dimensional surface plot with 10 nodes

Table 4.5. Details of the experimental design for Model B

Experiment No.		$\lvert\mathcal{L}\rvert = 3$		$\lvert\mathcal{L}\rvert = 5$	
		$\upsilon = 25$	$\upsilon = 50$	$\upsilon = 25$	$\upsilon = 50$
$\lvert\mathcal{N}\rvert = 10$	$\lvert\mathcal{T}\rvert = 10$	$E.B1$	$E.B2$	$E.B3$	$E.B4$
	$\lvert\mathcal{T}\rvert = 20$	$E.B5$	$E.B6$	$E.B7$	$E.B8$
	$\lvert\mathcal{T}\rvert = 30$	$E.B9$	$E.B10$	$E.B11$	$E.B12$
$\lvert\mathcal{N}\rvert = 20$	$\lvert\mathcal{T}\rvert = 10$	$E.B13$	$E.B14$	$E.B15$	$E.B16$
	$\lvert\mathcal{T}\rvert = 20$	$E.B17$	$E.B18$	$E.B19$	$E.B20$
	$\lvert\mathcal{T}\rvert = 30$	$E.B21$	$E.B22$	$E.B23$	$E.B24$
$\lvert\mathcal{N}\rvert = 30$	$\lvert\mathcal{T}\rvert = 10$	$E.B25$	$E.B26$	$E.B27$	$E.B28$
	$\lvert\mathcal{T}\rvert = 20$	$E.B29$	$E.B30$	$E.B31$	$E.B32$
	$\lvert\mathcal{T}\rvert = 30$	$E.B33$	$E.B34$	$E.B35$	$E.B36$
$\lvert\mathcal{N}\rvert = 40$	$\lvert\mathcal{T}\rvert = 10$	$E.B37$	$E.B38$	$E.B39$	$E.B40$
	$\lvert\mathcal{T}\rvert = 20$	$E.B41$	$E.B42$	$E.B43$	$E.B44$
	$\lvert\mathcal{T}\rvert = 30$	$E.B45$	$E.B46$	$E.B47$	$E.B48$

Table 4.6 summarizes the computational results of the experimental design by presenting the objective function value, lower bound, the percentage of optimality gap, CPU time and the weighted average efficiency of the facilities; for each test instance.

Table 4.6. Computational results of the experimental design for Model B

Experiment No.	Objective Function (million)	Lower Bound (million)	Optimality Gap(%)	CPU time (sec.)	Weighted Avg. Efficiency(%)
E.B1	409.02	409.02	0.00	8.75	96.46
E.B2	329.95	329.95	0.00	2218.16	98.74
E.B3	335.51	335.51	0.00	13.30	95.53
E.B4	327.23	319.75	2.28	3000.00	99.7
E.B5	624.96	624.96	0.00	42.69	96.83
E.B6	557.71	542.44	2.74	6000.00	98.69
E.B7	628.66	628.66	0.00	48.33	96.09
E.B8	542.92	525.52	3.21	6000.00	99.87
E.B9	3924.32	924.32	0.00	248.14	96.04
E.B10	908.02	877.69	3.34	9000.00	99.51
E.B11	1031.91	1031.91	0.00	266.23	93.33
E.B12	871.84	848.03	2.73	9000.00	99.51
E.B13	605.69	583.49	3.67	6000.00	94.04
E.B14	602.57	583.47	3.17	6000.00	94.04
E.B15	570.09	558.27	2.07	6000.00	98.75
E.B16	575.30	557.47	3.10	6000.00	98.75
E.B17	999.27	949.48	4.98	12000.00	99.19
E.B18	1002.86	943.07	5.96	12000.00	99.19
E.B19	987.48	936.43	5.17	12000.00	95.22
E.B20	993.26	928.18	6.55	12000.00	95.22
E.B21	1615.88	1539.03	4.76	18000.00	96.43
E.B22	1629.97	1535.42	5.80	18000.00	96.43
E.B23	1590.81	1464.06	7.97	18000.00	98.46
E.B24	1606.75	1453.91	9.51	18000.00	98.55
E.B25	803.96	775.71	3.51	9000.00	96.12
E.B26	800.22	767.89	4.04	9000.00	96.12
E.B27	785.30	755.37	3.81	9000.00	97.55
E.B28	777.49	744.39	4.26	9000.00	97.55
E.B29	1424.66	1315.49	7.66	18000.00	97.48
E.B30	1419.77	1307.93	7.88	18000.00	98.1
E.B31	1358.44	1289.31	5.09	18000.00	97.56
E.B32	1353.04	1276.40	5.66	18000.00	99.12
E.B33	2513.63	2266.51	9.83	27000.00	97.36
E.B34	2483.45	2217.18	10.72	27000.00	97.36
E.B35	2399.31	2149.26	10.42	27000.00	96.53
E.B36	2380.09	2110.53	11.33	27000.00	95.98
E.B37	1156.29	1070.48	7.42	12000.00	95.1
E.B38	1113.31	1019.83	8.40	12000.00	95.27
E.B39	1098.62	992.36	9.67	12000.00	97.26
E.B40	1079.09	963.13	10.75	12000.00	98.03
E.B41	2273.21	1900.13	16.41	24000.00	96.47
E.B42	2268.41	1889.31	16.71	24000.00	96.65
E.B43	2109.31	1856.02	12.01	24000.00	97.83
E.B44	2083.96	1821.32	12.6	24000.00	98.6
E.B45	3518.92	3001.78	14.7	36000.00	94.12
E.B46	3415.05	2976.19	12.85	36000.00	94.49
E.B47	3270.98	2854.51	12.73	36000.00	96.38
E.B48	3284.91	2765.23	15.82	36000.00	97.13

It is observed that in seven out of 48 experiments, an exact optimal solution is found within the specified time limit. Solutions within five percent of the optimality gap are found in about half of the experiments (22 out of 48) and solutions within 10 percent of the optimality gap are found in 36 out of 48 experiments. It is clear that the performance of the CPLEX algorithm is inferior compared to the result of Model A, which is expected given the higher complexity of Model B. Nevertheless, the algorithm has been able to solve each experiment and the weighted average efficiency is higher than 90 percent in all combinations.

To illustrate a single experiment, $E.B8$ is chosen and the complete network in this particular experiment is presented in Figure 4.14. Besides all other decisions that can be seen in the figure, it is observed that, a facility with level 1 is expanded to level 4 on node 7 time period $t = 11$. There are particular links that connect two nodes with facilities, such as the links (3,6) and (4,6). Moreover, even though a facility is located on node 7, part of its demand is served by the facility on node 8. That connection is conducted indirectly by the links (7,9) and (9,8), respectively. The details of this particular experiment are provided in Appendix B, by displaying the network and decision variables for each time period.

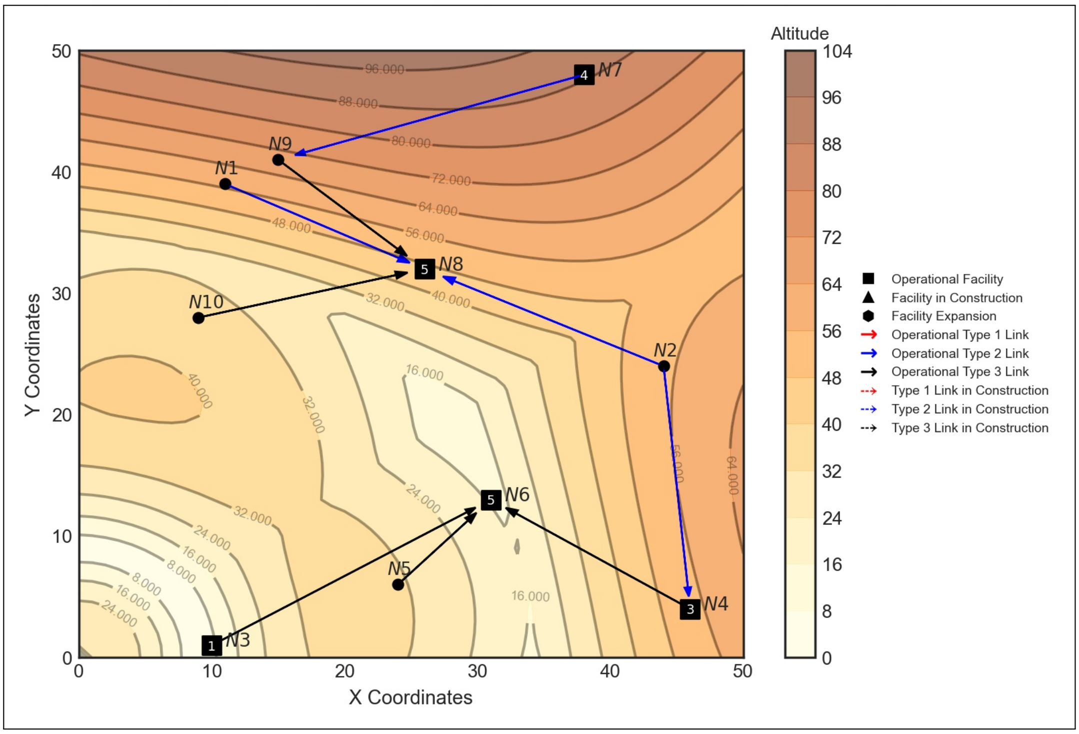


Figure 4.14. Complete network in $E.B8$ ($t = 20$)

Figures 4.15 - 4.24 present bar plots for the output of the experimental design for each set level combination. As in the experimental results of Model A, it is clear that the increase in $|\mathcal{N}|$ and $|\mathcal{T}|$ triggers an increase in the total cost as presented in Figure 4.15 that illustrates the effect of different set levels on total cost, which is an expected behavior in line with the expansion of the whole network. In fact, facility costs appear to grow linearly with the increase in both $|\mathcal{N}|$ and $|\mathcal{T}|$. Another important point observed from Figure 4.15 is that the total cost increases slightly in experiments where the number of possible facility levels $|\mathcal{L}| = 3$. This behavior appears more clearly in experiments where the network is bigger (i.e., experiments with more nodes). On the other hand, in terms of the total cost, the distance limit υ is found to have a minor effect.

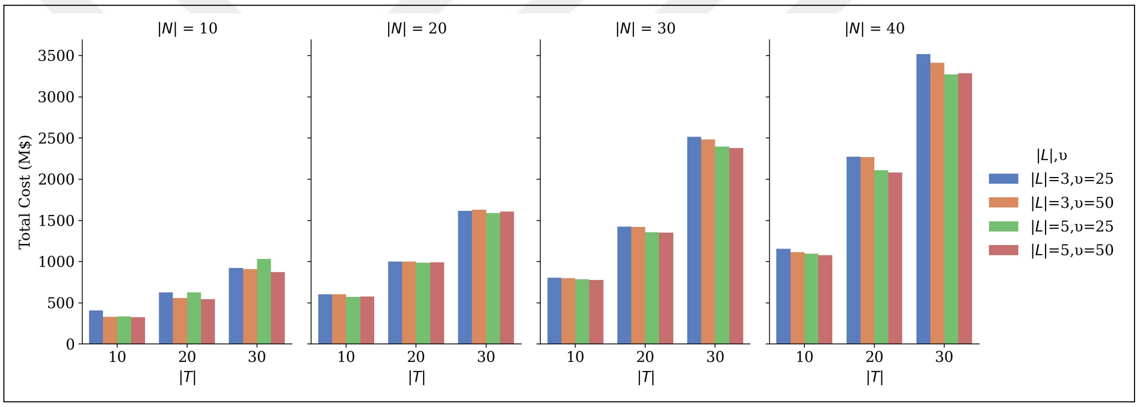


Figure 4.15. Effect of different set levels on total cost for the experimental design of Model B

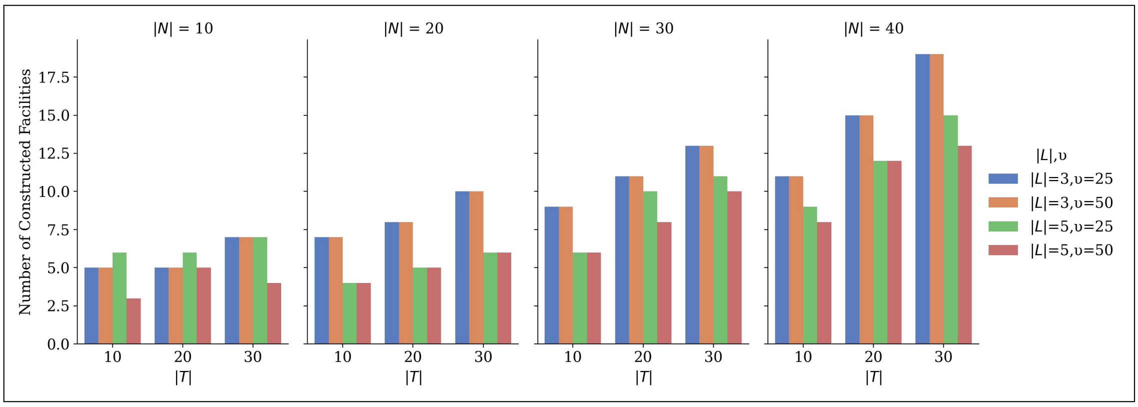


Figure 4.16. Effect of different set levels on number of constructed facilities for the experimental design of Model B

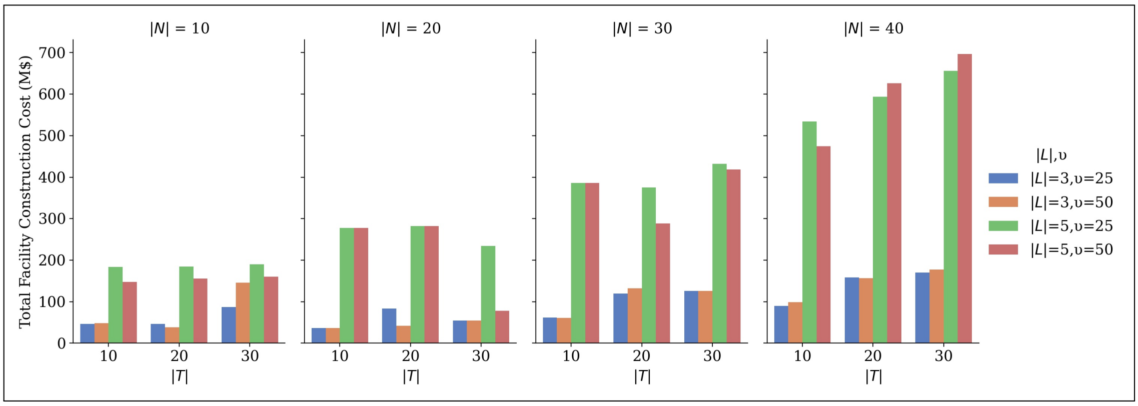


Figure 4.17. Effect of different set levels on total facility construction cost for the experimental design of Model B

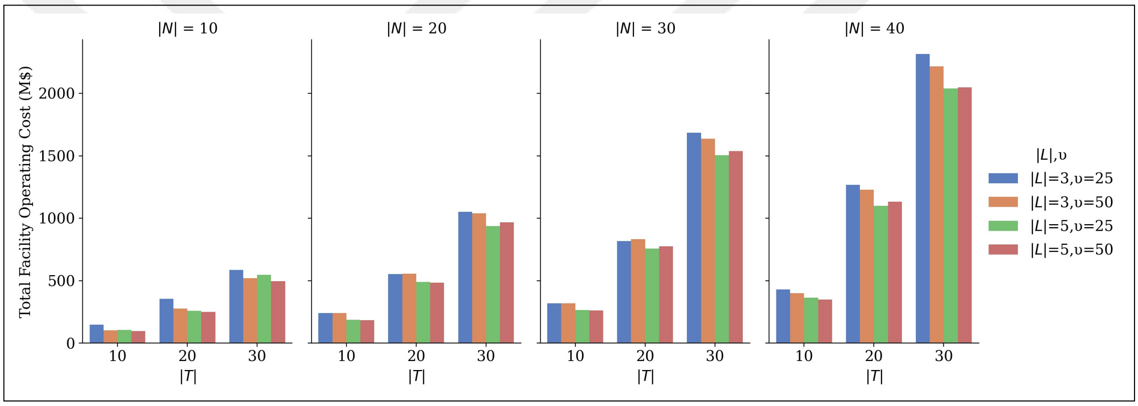


Figure 4.18. Effect of different set levels on total facility operating cost for the experimental design of Model B

Moreover, it is also observed that, the number of constructed facilities is very sensitive to the number of possible facility levels, $|\mathcal{L}|$ (see Figure 4.16). One may recall that, the maximum facility capacity depends on $|\mathcal{L}|$, since the data on urban wastewater treatment plant capacities are divided into $|\mathcal{L}|$ equal intervals in order to specify facility capacities. The maximum facility capacity is 10000 $m^3/hour$ when $|\mathcal{L}| = 5$, whereas it is 6000 $m^3/hour$ when $|\mathcal{L}| = 3$. Therefore, to ensure all the demand in the network, more constructed facilities are expected when $|\mathcal{L}| = 3$. However, on the contrary, total facility construction cost is found to be significantly greater in experiments where $|\mathcal{L}| = 5$ (see Figure 4.17). This behavior is due to far more number of capacity expansions in experiments where $|\mathcal{L}| = 3$ (see Figure 4.23), which also affects the total expansion cost in the network (see Figure 4.24). One may recall that, expansion costs are not included while considering construction costs. Moreover, the total facility operating cost, when all these situations are evaluated together, is in parallel

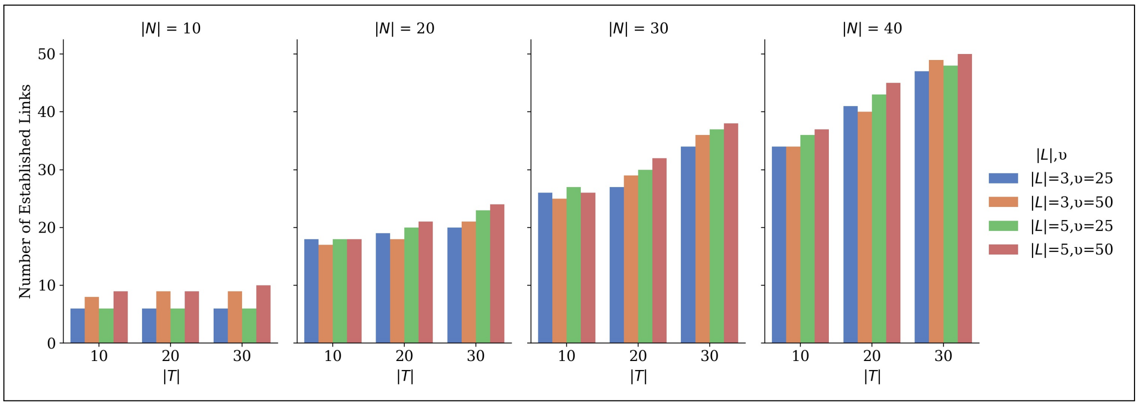


Figure 4.19. Effect of different set levels on number of established links for the experimental design of Model B

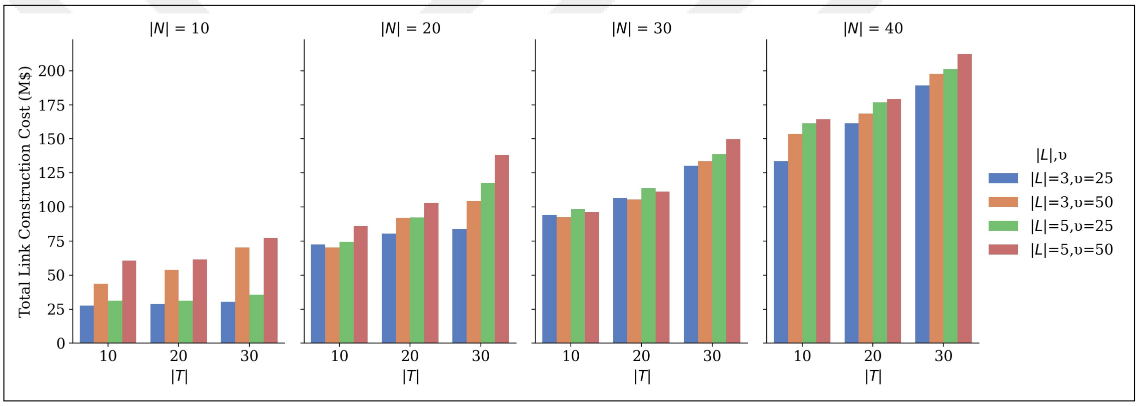


Figure 4.20. Effect of different set levels on total link construction cost for the experimental design of Model B

with the number of constructed facilities (see Figure 4.18).

As distance limit υ and number of possible facility levels $|\mathcal{L}|$ increases, number of established links tend to increase slightly (see Figure 4.19). The main reason for this situation is the restriction on the possible combinations of decision variables, as explained in the results of Model A. The same situation holds true for Model B. With a similar example, further restriction on the distance limit υ will reduce the number of candidate links to be established and, consequently, part of the links that are involved in the solution of experiments where υ is bigger, cannot be established. The link related expenditures are in parallel with the number of established links (see Figure 4.20, Figure 4.21). Moreover, total traveling cost is also affected by $|\mathcal{L}|$, especially for the smaller networks (see Figure 4.22).

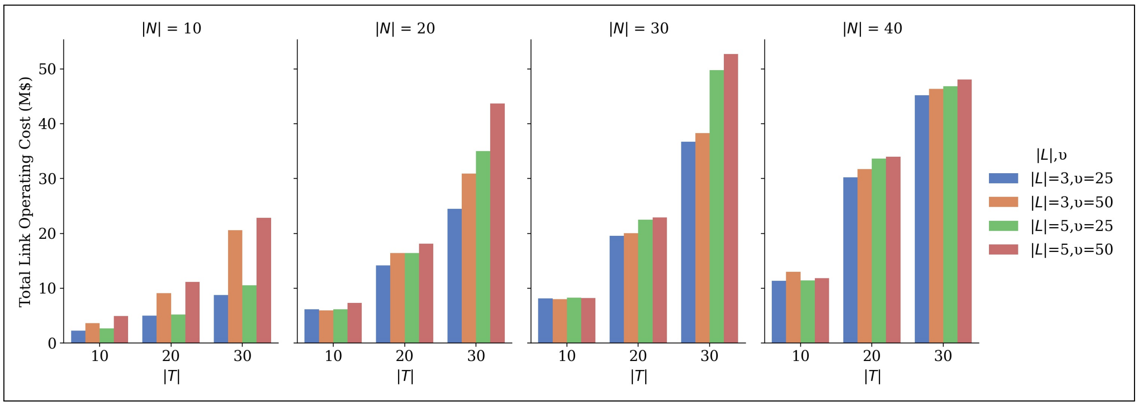


Figure 4.21. Effect of different set levels on total link operating cost for the experimental design of Model B

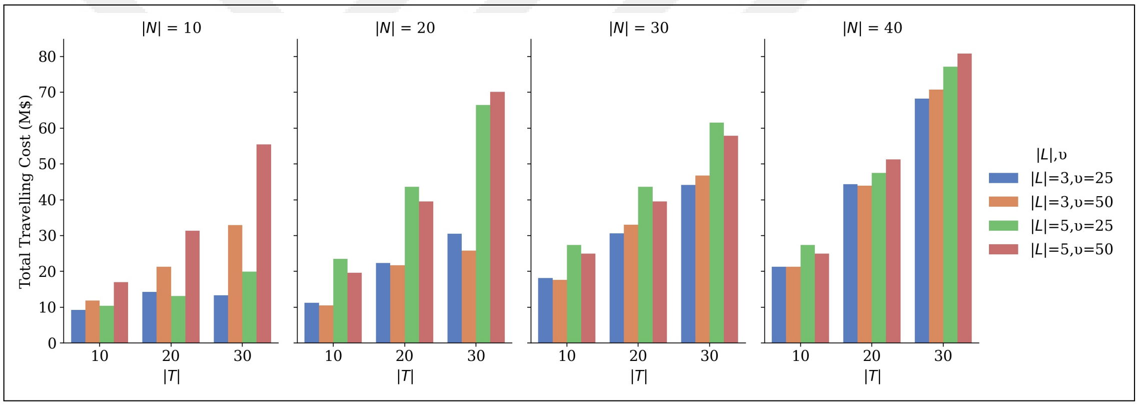


Figure 4.22. Effect of different set levels on total traveling cost for the experimental design of Model B

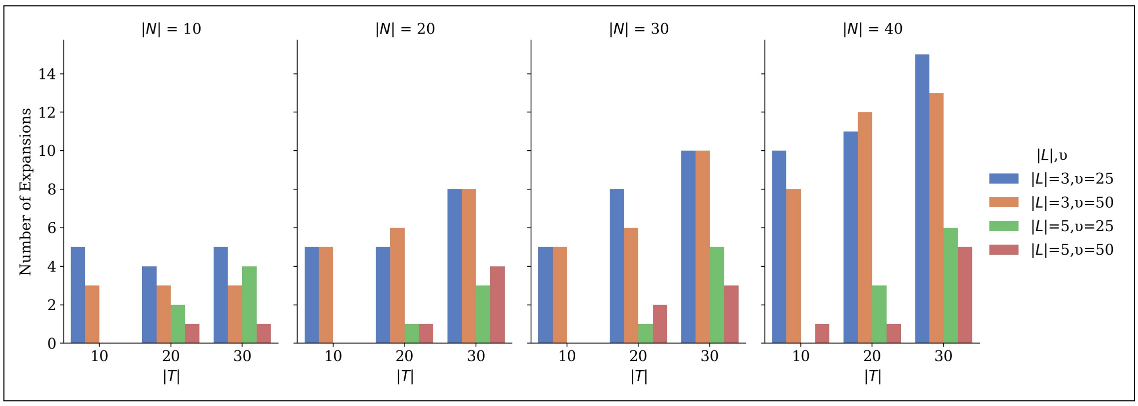


Figure 4.23. Effect of different set levels on number of facility expansions for the experimental design of Model B

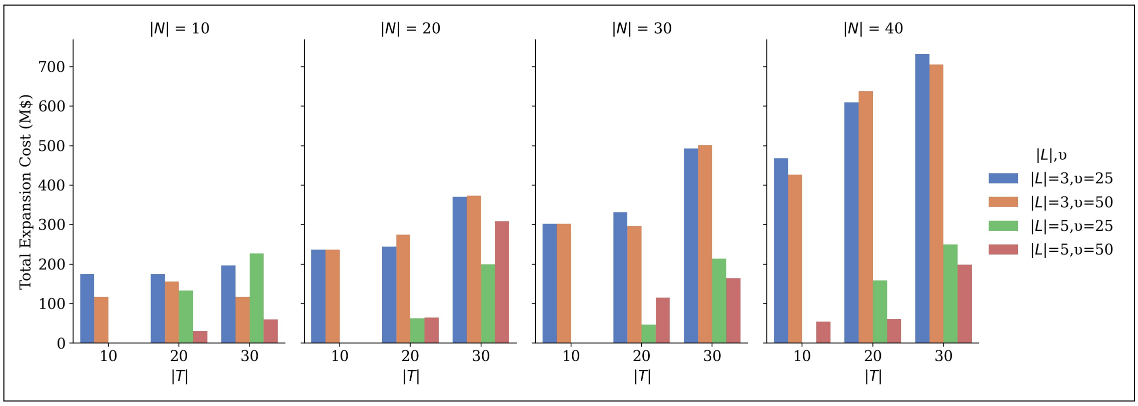


Figure 4.24. Effect of different set levels on total expansion cost for the experimental design of Model B

5. VERIFICATION AND VALIDATION

It is necessary to check that, the mathematical modeling method or solution technique represents the actual system in a good and accurate way. In this matter, what constitutes a good or accurate representation can be subjective. However, the main criteria to judge the goodness of a model representation in terms of performance issues will mainly depend on the accuracy level of measurements which is extracted from the model solutions. In general, there are two steps that judge the goodness of a model: verification and validation. The former concerns with "building the model right" and the latter concerns with "building the right model" [119]. In other words, verification deals with implementing the assumptions, the constraints and the outputs of the model correctly. On the other hand, validation deals with whether all the aspects of the model are reasonable with respect to the existing system.

Several methods exist to apply verification and validation, especially in simulation models [120]. On the other hand, in optimization models, the most extensive method to validate is comparing the solutions with the actual system measurements and/or the best-known model solutions from the literature. However, as mentioned earlier, there is neither any conjugate model in the literature to conduct a benchmark test nor historical data from real-world scenarios yet for the proposed problems in this thesis. Therefore, a group of experts that involves people from both academia and related sectors are consulted and their opinions are taken as an expert intuition and opinion to validate the model. All of the assumptions and formulations of mathematical models are discussed and some assumptions are included as an outcome of these discussions.

While verifying a mathematical model, some useful methods can be listed as follows [121]:

- Checking the model results on small examples that can be solved by hand
- Checking the model results against known extreme cases
- Checking the model results with respect to the real-world cases
- Checking that the model handles invalid cases robustly
- Checking the consistency and the plausibility that model results change in the proper

direction as model inputs are changed

In practice, verification and validation are often blended. For example, comparing model results with an existing system is a method for both validation and verification. As explained in the previous paragraph, this particular method is not applicable.

To check the consistency of the models, one may consider the experimental results, which are presented in Section 4.1 and Section 4.2. It is clear that, as set levels are increased and the problem becomes more complex, the objective function values (i.e., the total cost) increase and the performance of the CPLEX algorithm decreases. The fact that the average efficiency is higher than 90 percent in all experiments indicates a consistent framework.

To verify Model B appropriately, several test instances with simplified parameters that involve some extreme or invalid cases have been generated, without loss of generality. In the remainder of this section, these generated instances are presented and their results are discussed with respect to the logical consistency and plausibility. Three example nodes are located over an 8×8 unit area, whose x and y coordinates are $(2, 2)$, $(2, 5)$ and $(6, 5)$, respectively. Initially, the altitude of the whole generated network is set to be 0. The generated network is illustrated in Figure 5.1. The parameters are determined initially, but each test instance has a distinct structure. Therefore, some parameters will vary between test instances. In each test instance, these varying parameters will be revealed.

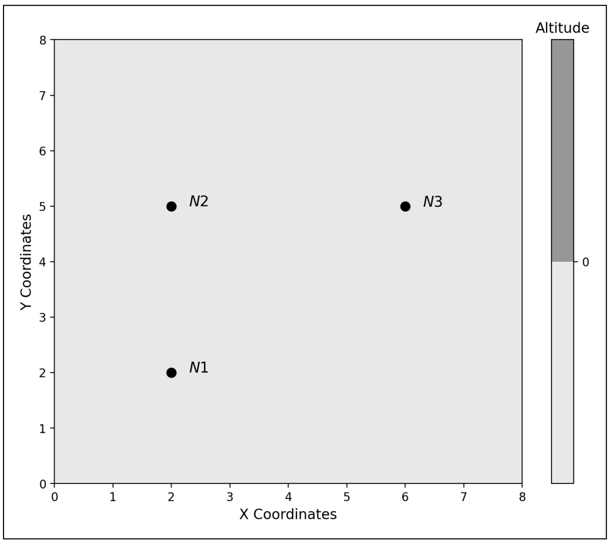


Figure 5.1. Generated network for test instances

By default, the total time period, $|\mathcal{T}|$, is set to be 5. Number of facility levels and number of link types are determined as 2 and 1, respectively($|\mathcal{L}| = 2, |\mathcal{P}| = 1$). All construction times are set to be 1, regardless of their types or locations. The remaining default parameters are summarized in Table 5.1.

Table 5.1. Default parameters for test instances

Parameter	**Value Generation**
d_k^t	$d_k^t = 10;\ \forall k \in \mathcal{I}, \forall t \in \mathcal{T}$
f_i^{lt}	$f_i^{1t} = 20;\ \forall i \in \mathcal{I}, \forall t \in \mathcal{T}$ $f_i^{2t} = 50;\ \forall i \in \mathcal{I}, \forall t \in \mathcal{T}$
$e_i^{l\zeta t}$	$e_i^{1,2,t} = 30;\ \forall i \in \mathcal{I}, \forall t \in \mathcal{T}$
c_{ij}^{pt}	$c_{ij}^{1t} = dist_{ij};\ \forall t \in \mathcal{T}$
tc_{ij}^{pt}	$tc_{ij}^{1t} = 1 + \Delta h_{ij} \times 0.1;\ \forall t \in \mathcal{T}$
g_i^{lt}	$g_i^{1t} = 1;\ \forall i \in \mathcal{I}, \forall t \in \mathcal{T}$ $g_i^{2t} = 2;\ \forall i \in \mathcal{I}, \forall t \in \mathcal{T}$
CF^l	$CF^1 = 20$ $CF^2 = 50$
CL	$CL = 20$

Test Instance 1:

All demands are set to 0, while all of the remaining parameters are set to their default values throughout the network in order to observe how the model behaves when no action is required. Theoretically, no expenditure is expected if there is no demand in demand-based systems. Accordingly, the model behaved as expected by giving the objective function value as 0.

Test Instance 2:

All parameters are set as in Test Instance 1, except for the demand of node 1, which is set to its default value. The main aim is to observe how the model behaves when a limited number of actions are required. Due to the structure of the model, it is expected that a facility will be located only on node 1, to supply its demand and no additional action is required. Thus, a facility with level 1, which is adequate to supply the demand, becomes operational as of the time period $t = 1$. The objective function value is 25 (Facility Construction Cost = 20, Facility Operating Cost = 5). The illustration of this action is given in Figure 5.2.

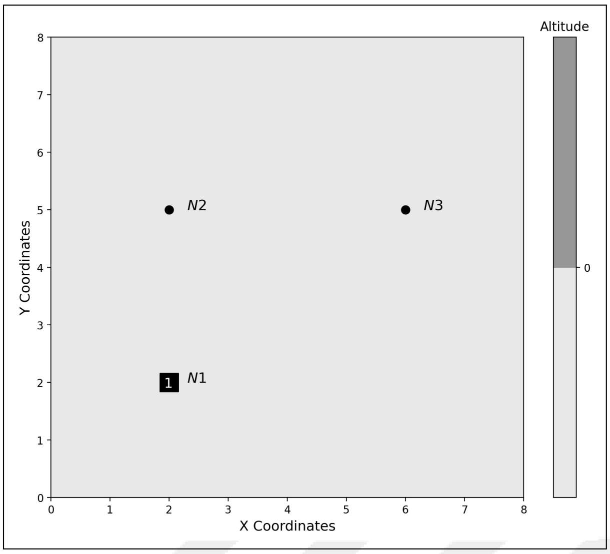


Figure 5.2. Illustration of Test Instance 2

Test Instance 3:

All parameters are set as in Test Instance 1, except for the demand of node 1, which is set to 45 instead of 10 for all time periods. Theoretically, a facility with level 2 must be located on node 1, since the demand is higher than the capacity of a level 1 facility. The model behaves as expected, with an objective function value of 60 (Facility Construction Cost = 50, Facility Operating Cost = 10). The illustration of this particular instance is shown in Figure 5.3.

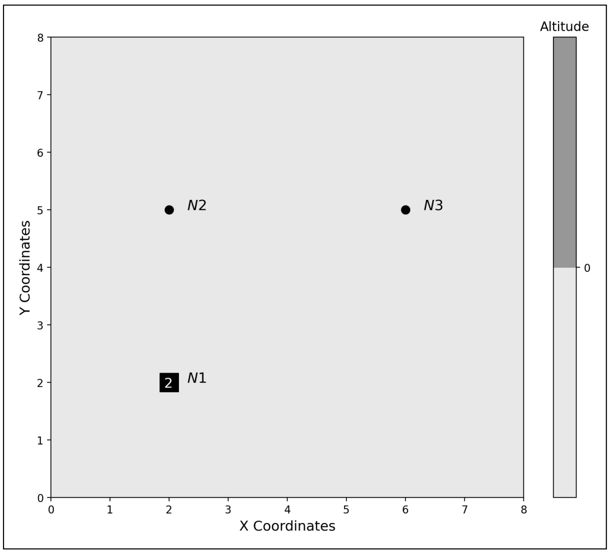
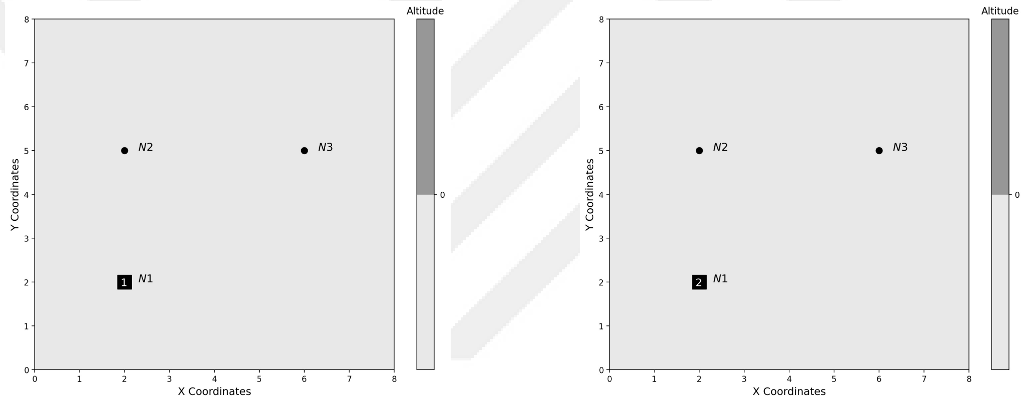


Figure 5.3. Illustration of Test Instance 3

Test Instance 4:

All parameters are set as in Test Instance 2, except for the demand of node 1 at time period $t = 5$, which is set as 45. More explicitly, the demand of node 1 is 10 until time period $t = 5$. Then, it increases to 45, when time period $t = 5$. In this particular instance, the best solution is locating a facility with level 1 as of time period $t = 1$ on node 1 and expanding that facility to level 2 at time period $t = 5$. The model behaves as expected with an objective function value of 56 (Facility Construction Cost = 20, Facility Expanding Cost = 30, Facility Operating Cost = 6). By this means, it is observed that the model has a flexible structure among decision variables. The illustration of this instance is shown in Figure 5.4.



(a) $t = 1$ (b) $t = 5$

Figure 5.4. Illustration of Test Instance 4

Test Instance 5:

All parameters are set to their default values except facility construction costs of node 2 and node 3, whose costs are increased by 10 times. It is expected to observe that, nodes 2 and 3 meet their demand from a facility located on node 1 due to the increase in facility construction cost. With an objective function value of 78 (Facility Construction Cost = 50, Facility Operating Cost = 10, Link Construction Cost = 8, Traveling Cost = 10), the model decides to locate a facility with level 2 on node 1 at time period $t = 1$ and establish links (2,1) and (3,1) at time period $t = 1$. Figure 5.5 presents the details of this test instance.

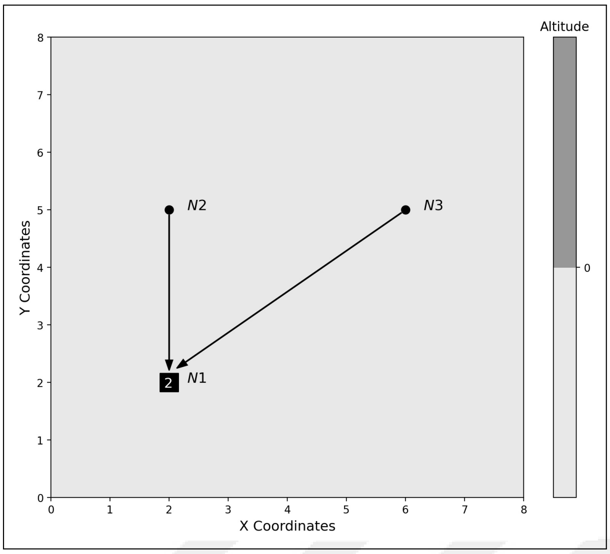


Figure 5.5. Illustration of Test Instance 5

Test Instance 6:

All parameters are set as in Test Instance 5, except the cost of constructing the link (3,1) is 10 times higher than its default value. Therefore, instead of using link (3,1) directly, constructing the link (3,2) in order to transfer the demand of node 3 to node 1 with the assistance of node 2 would be a better solution. One may recall that, each demand flows to an arbitrary node can be transferred to the other nodes. Consequently, the model decides to locate a facility with level 2 on node 1 at time period $t = 1$ and establish links (2,1) and (3,2) at time period $t = 1$, with an objective function value of 82 (Facility Construction Cost = 50, Facility Operating Cost = 10, Link Construction Cost = 7, Traveling Cost =15). The illustration of this instance is shown inFigure 5.6.

Test Instance 7:

All parameters are set to their default values except the capacities of facilities with level 1 and level 2, which are determined as 2 and 5, respectively. One may recall that, each demand must be served for each demand point. However, if the total demand is higher than the maximum allowable capacity throughout the network, which is 15 in this case, there is no feasible region for this model. Therefore, the model returns an infeasibility error as expected. This particular instance is an example of an invalid case.

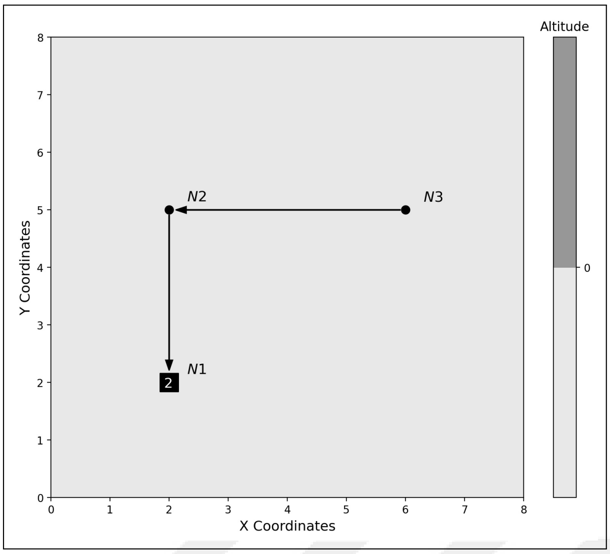


Figure 5.6. Illustration of Test Instance 6

Test Instance 8:

The capacities of links are determined as 2 and the cost of constructing a link is set to 0. All the remaining parameters are set to their default values. Even though transfer costs are eliminated, links are useless if they do not have sufficient capacity to flow over the network. In these cases, each node has to meet its own demand. With an objective function value of 75 (Facility Construction Cost = 60, Facility Operating Cost = 15), facilities with level 1 are located on each node at time period $t = 1$. This particular instance is illustrated in Figure 5.7.

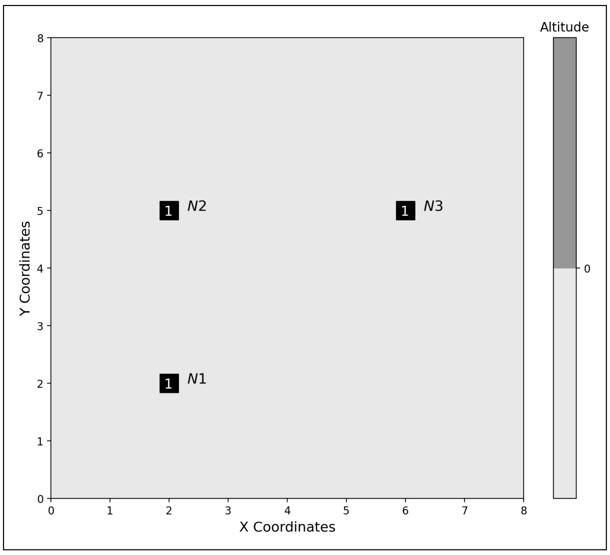


Figure 5.7. Illustration of Test Instance 8

Test Instance 9:

All parameters are set as in Test Instance 8, except the construction time of the level 1 facilities are set to 3. More explicitly, a facility with level 1 cannot be operational until time period $t = 3$. Therefore, since they have to meet their demand at time periods $t = 1$ and $t = 2$, instead of locating facilities with level 1 on each node, level 2 facilities must be located even though they constitute a higher level of costs. In line with this situation, the model decides to locate level 2 facilities on each node at time period $t = 1$, with an objective function value of 180 (Facility Construction Cost = 150, Facility Operating Cost = 30). These actions are illustrated in Figure 5.8.

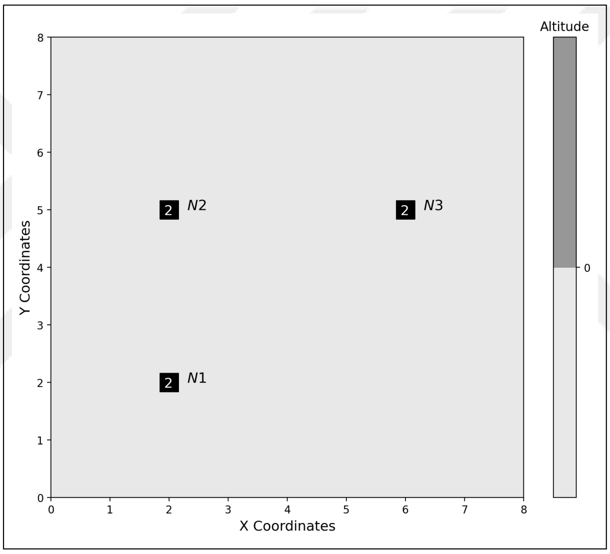


Figure 5.8. Illustration of Test Instance 9

Test Instance 10:

The demand of node 1 is set to 50 for all time periods and the cost of constructing a facility on node 1 is increased by 10 times. The capacity of any link is determined as 30 and the remaining parameters are set to their default values. In this particular instance, the demand of node 1 is higher than the other nodes, but it has to meet its demand from nodes 2 and 3 due to the construction costs. However, it cannot meet its demand from a single node since the capacity of a link is 30, whereas the demand of node 1 is 50. Therefore, the demand has to be divided. One may recall that, the demand divisibility is allowed in Model B. The main aim of this instance is to observe how the demand divisibility structure works. The fundamental point to note is that, a serving node, which can be node 2 or node 3, has to meet its own demand, along with the demand of node 1. The total demand in the network is 70 for all time periods. Thus, the total capacity of the facilities in the network has to be at least 70, which is the total capacity of a level 1 facility and a level 2 facility. In line with all these situations, the model decides to locate a facility with level 1 on node 2 and a facility with level 2 on node 3 at time period $t = 1$. Furthermore, links (1,2), (1,3) and (2,3) are established. Due to the capacity restrictions, direct links between node 1 and serving nodes, (1,2) and (1,3), are not adequate. Some fraction of demand of node 1 has to flow over the link (2,3), since the facility at node 3 has to serve 40 units of demand whereas the link capacity is 30. More explicitly, $W_3^{1t} = 0.8$ and $W_2^{1t} = 0.2$ for all time periods, which means 80 percent of the demand of node 1 is served by node 3, whereas 20 percent of it is served by node 2. Besides, $Y_{12}^{11t} = 0.4$, $Y_{13}^{11t} = 0.6$ and $Y_{23}^{11t} = 0.2$ for all time periods; which means 40 percent of the demand of node 1 flows over the link (1,2), 60 percent of it flows over the link (1,3) and 20 percent of it flows over the link (2,3) for all time periods. Half of the demand flows over the link (1,2) are transferred to the link (2,3) with the assistance of node 2. To conclude, the model returns an objective function value of 103 (Facility Construction Cost = 70, Facility Operating Cost = 15, Link Construction Cost = 12, Traveling Cost = 6). The simplified illustration of this test instance is given in Figure 5.9.

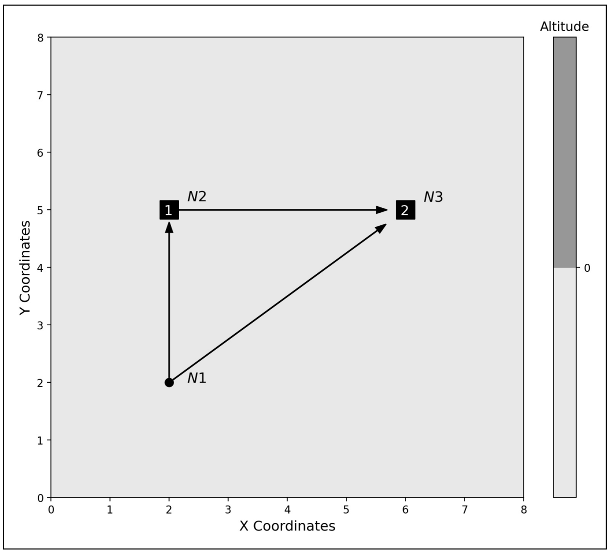


Figure 5.9. Illustration of Test Instance 10

Test Instance 11:

The distance limit, υ, is set to be 3, which is the distance between node 1 and node 2. The cost of constructing a link and traveling cost are set to 0, meaning no transfer cost is constituted. However, due to the distance limit, node 3 cannot establish a link with other nodes. Considering that all the remaining parameters are set to their default values, the best solution would be to locate a facility with level 1 on node 1 or node 2 and establish a link between these nodes. Node 3 has to meet its own demand. The model behaves as expected, with the objective function value of 50 (Facility Construction Cost = 40, Facility Operating Cost = 10). The illustration of this particular instance is shown in Figure 5.10.

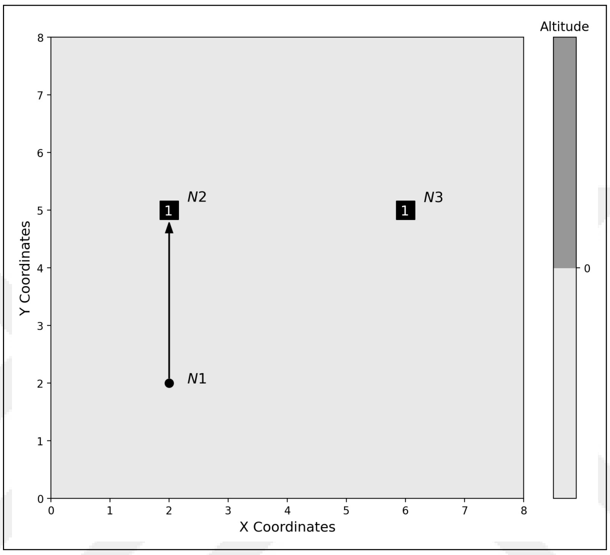


Figure 5.10. Illustration of Test Instance 11

Test Instance 12:

The cost of constructing a facility is increased by 10 times. The altitudes of nodes 1,2 and 3 are determined as 4, 10 and 30, respectively. The capacity of a level 1 facility is set to 30, which is the total demand in the network. All of the remaining parameters are set to their default values. Due to the higher construction cost, a single facility has to be located in the network and the other nodes have to transfer their demands. One may recall that, traveling cost from node i to node j depends on the vertical difference between these nodes, Δh_{ij}. Therefore, a transfer from a higher level of altitude to a lower one constitutes a lower level of cost. In line with this, the model decides to locate a facility on node 1 and establish links (2,1) and (3,1) at time period $t = 1$, with an objective function value of 223 (Facility Construction Cost = 200, Facility Operating Cost = 5, Link Construction Cost = 8, Traveling Cost = 10). These actions are illustrated in Figure 5.11.

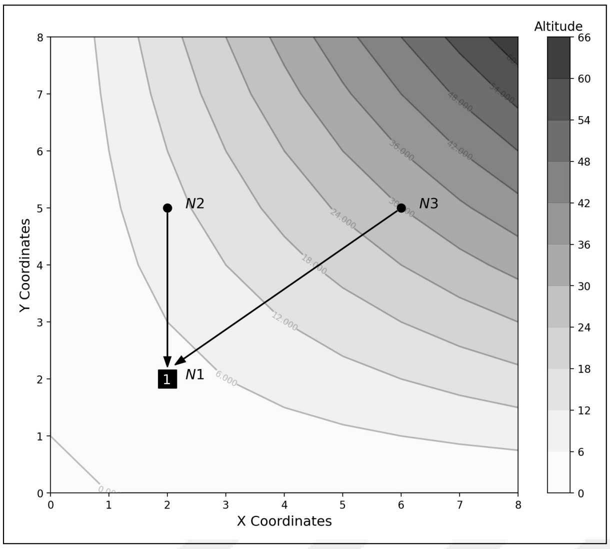


Figure 5.11. Illustration of Test Instance 12

6. DISCUSSIONS AND FUTURE WORK

As stated by Melkote and Daskin [56], CFLNDP is an NP-hard problem in terms of computational complexity. One of the proposed mathematical models in this thesis, Model A, is an extended dynamic version of the original CFLNDP formulation. Consequently, CFLNDP is a special case of Model A, where there is only a single time period and facility level. Furthermore, construction times are ignored in CFLNDP formulation. Therefore, Model A is an NP-hard problem in terms of computational complexity. Accordingly, Model B is an extended version of Model A, in which capacity expansion and link types are included in the network. Thus, Model B is also an NP-hard problem.

It is observed that, the performance of the CPLEX algorithm is decreased significantly as the scale of the problems increases in the proposed mathematical models. In the experimental results of some particular 40-node networks for Model A, the optimality gap exceeds five percent within the specified time limit. Moreover, the optimality gap exceeds 10 percent in some particular 40-node experiments for Model B. As mentioned in Section 4.1, CPLEX runs out of memory in an 80-node test problem conducted by Rahmaniani and Ghaderi [64] for CFLNDP. Considering the higher complexity of Model A and Model B, it is clear that the CPLEX algorithm is effective up to certain problem size. Therefore, in order to deal with bigger networks, different solution techniques such as case-specific heuristics or metaheuristics can be developed to solve the proposed problems in a future direction. As mentioned in Section 2.3, numerous solution methods have been presented for both FLNDP and CFLNDP problems in the literature. Since the structure of the proposed problems in this thesis is different, especially the solution methods used in dynamic formulations are worth noting. As an alternative way, when a huge network is encountered, the network can be split into smaller segments to reduce the complexity of the problem by using several extensive clustering algorithms such as K-means, DBSCAN (Density-Based Spatial Clustering of Applications with Noise), GMM (Gaussian Mixture Model) and BIRCH (The Balance Iterative Reducing and Clustering using Hierarchies) [122,123,124,125]. In addition to all these, certain tolerance levels can be imposed, or penalty costs can be included for some constraints, and consequently, the model assumptions can be relaxed to allow for more modeling flexibility. For instance, the assumption to ensure all demand in the network can be relaxed by including certain penalty costs into the unmet demand of each node.

As a future work, the proposed mathematical models, especially Model B for wastewater treatment networks, can be applied to real-world scenarios. However, detailed micro-scale (i.e., districts, neighborhood) data are necessary, especially for appropriate consideration of demand and the topography of the network. One may recall that the proposed models are applicable to existing networks; that is, it is not necessary to have an empty network; there may be constructed facilities or established links. This structure can be practical since it provides a broad area of concern.

In order to enhance its applicability to real-life wastewater transportation scenarios, several extensions are possible for Model B. One may recall that, the network in Model B is a customer-to-server system; that is, in terms of wastewater treatment networks, wastewater collected from the demand points flows to the treatment facilities. As an alternative structure, reusable water treated at the facilities can be returned to the demand points. Thus, the mathematical model can be formulated as a two-way system involving both customer-to-server and server-to-customer structures. Obviously, in order to conduct such a two-way structure, different pipelines or pipeline types should be established. Therefore, the assumption that the links can only be constructed in one direction between two nodes should be violated. Moreover, some alternative sewage systems (e.g., transfer of wastewater through tankers, collection of wastewater on certain underground reservoirs before it is transferred to a treatment plant, etc.) can be included, especially for topographically rough and patchy nodes with low demand or some isolated nodes in the network. Their transportation structures can be solved by network routing problems.

Considering the dynamic nature of the models, high uncertainties may emerge on some key parameters such as demand and corresponding costs that may ramify over time. In fact, this situation can be considered one of the main fundamental aspects of dynamic formulations. Therefore, if appropriate historical data are available, several statistical/possibilistic techniques can be applied to forecast uncertain parameters. If these forecasts are embedded into the models together with their prediction intervals, stochasticity will be introduced; thus, the problem becomes even harder to solve. Furthermore, taking into this stochastic environment, dynamic stochastic programming models may be generated for a future study.

It is clear that, finding an optimal solution in programming models that consist of nonlinear aspects within their constraints is almost impossible in most cases. Therefore, in order to preserve the linearity, a piecewise linear function is fitted to the term $Q^{2.852}$ in Equation (4.7). It should be noted that the smoothing in the fit may affect the accuracy to a point. Nevertheless, as mentioned earlier, the deviations can be ignored with a high R^2 value of 0.994.

The proposed mathematical models can be used in several other application areas that involve pipeline transportation, such as natural gas transportation systems or, more generally, energy management systems, by adding case-specific constraints. For instance, the specified network can be transformed into a server-to-customer system with minor modifications in the formulation.

7. CONCLUSIONS

This thesis proposes two novel mixed-integer programming models for developing effective wastewater treatment networks. While formulating such models, dynamic structures are used, considering the varying behavior of the key parameters over time. The proposed models decide where to locate facilities and corresponding links, along with the opportunity for controlling and observing the flow and traffic in a specified network and time horizon. In order to enhance its applicability to real-life cases, one of the mathematical models is defined on a three-dimensional space with allowable capacity expansion, which is not considered in the literature to the best of our knowledge.

An experimental design is conducted for both proposed mathematical models and their results are discussed to reveal their sensitivity to set levels. To verify the mathematical models, a group of experts that involves people from both academia and related sectors is consulted. Moreover, some extreme test instances are generated to observe the consistency and plausibility of the models.

In summary, considering the importance of water for environmental resilience and sustainable development, decision methods are presented to develop an effective wastewater treatment network with the least cost. Furthermore, it is expected that these methods are also applicable to several areas other than wastewater treatment. Finally, a guideline is presented for the future directions of this study, including possible methodologies to be used.

The outcomes of this study have been published in an international proceedings paper [126] and there is an ongoing research article process.

REFERENCES

1. Finney JL. Water? What's so special about it? *Philosophical Transactions of the Royal Society B: Biological Sciences*. 2004;359(1448):1145-65.

2. Burek P, Satoh Y, Fischer G, Kahil MT, Scherzer A, Tramberend S, et al. Water futures and solution - fast track initiative (Final Report). IIASA, Laxenburg, Austria; 2016. Available from: http://pure.iiasa.ac.at/id/eprint/13008/.

3. Singh VP, Khedun CP, Mishra AK. Water, environment, energy, and population growth: implications for water sustainability under climate change. *Journal of Hydrologic Engineering*. 2014;19(4):667-73.

4. Dascher ED, Kang J, Hustvedt G. Water sustainability: environmental attitude, drought attitude and motivation. *International Journal of Consumer Studies*. 2014;38(5):467-74.

5. Tyagi S, Sharma B, Singh P, Dobhal R. Water quality assessment in terms of water quality index. *American Journal of water resources*. 2013;1(3):34-8.

6. Wolfram J, Stehle S, Bub S, Petschick LL, Schulz R. Water quality and ecological risks in European surface waters–Monitoring improves while water quality decreases. *Environment International*. 2021;152:106479.

7. Hoekstra AY, Wiedmann TO. Humanity's unsustainable environmental footprint. *Science*. 2014;344(6188):1114-7.

8. Huynen MMTE, Martens P, Akin SM. Climate change: an amplifier of existing health risks in developing countries. *Environment, Development and Sustainability*. 2013;15(6):1425-42.

9. Kummu M, Guillaume JHA, de Moel H, Eisner S, Flörke M, Porkka M, et al. The world's road to water scarcity: shortage and stress in the 20th century and pathways towards sustainability. *Scientific Reports*. 2016;6(1):38495.

10. Boretti A, Rosa L. Reassessing the projections of the World Water Development Report. *Clean Water*. 2019;2(1):1-6.

11. The World Economic Forum. *The Global Risks Report 2020*. vol. 15. Cologny/Geneva, Switzerland: The World Economic Forum; 2020.

12. UNESCO World Water Assessment Programme. *The United Nations world water development report 2021: valuing water*. Paris, France: UNESCO; 2021.

13. WRI. Annual Baseline Water Stress. Accessed: 2022-06-21. https://www.wri.org/applications/aqueduct/water-risk-atlas/#/?advanced=false&basemap=terrain&indicator=w_awr_def_tot_cat&lat=41.35041305068143&lng=-559.6875107288362&mapMode=analysis&month=1&opacity=0.79&ponderation=DEF&predefined=false&projection=absolute&scenario=optimistic&scope=baseline&timeScale=annual&year=baseline&zoom=2.

14. United Nations Environment Programme. Sick water?: the central role of wastewater management in sustainable development: a rapid response assessment. 2010. Available from: https://wedocs.unep.org/20.500.11822/9156.

15. Amoatey P, Bani R. Wastewater management. *Waste Water—Evaluation and Management*. Croatia: InTech. 2011:379-98.

16. Tilley E, Ulrich L, Luthi C, Reymond P, Zurbrügg C. *Compendium of sanitation systems and technologies*. 2nd ed. Dübendorf, Switzerland: Swiss Federal Institute of Aquatic Science and technology (Eawag); 2014.

17. Raschid-Sally L, Jayakody P. Drivers and characteristics of wastewater agriculture in developing countries: Results from a global assessment. Colombo, Sri Lanka: International Water Management Institute; 2008.

18. Food and Agriculture Organization of the United Nations. Wastewater treatment. Accessed: 2021-10-12. https://www.fao.org/3/T0551E/t0551e05.htm.

19. UNICEF, WHO. *Progress on drinking water, sanitation and hygiene*. Genève, Switzerland: UNICEF: United Nations Children's Fund; 2017.

20. Gallego-Schmid A, Tarpani RRZ. Life cycle assessment of wastewater treatment in developing countries: a review. *Water research*. 2019;153:63-79.

21. Mara D. Domestic Wastewater Treatment in Developing Countries. Earthscan Publications, London. 2003:.

22. Devlet Su İşleri. Toprak Su Kaynakları. Accessed: 2021-10-12. https://www.dsi.gov.tr/Sayfa/Detay/754#.

23. TÜİK. Nüfus Projeksiyonları, 2018-2080. Accessed: 2021-10-12. https://data.tuik.gov.tr/Bulten/Index?p=Nufus-Projeksiyonlari-2018-2080-30567.

24. OECD. *Water and Climate Change Adaptation: Policies to Navigate Uncharted Waters*. Paris: OECD Publishing; 2013.

25. Saatci E, Akpinar E. Assessing poverty and related factors in Turkey. *Croatian Medical Journal*. 2007;48(5):628.

26. Saatci AM. Solving water problems of a metropolis. *Journal of Water Resource and Protection*. 2013;5(4A):7-10.

27. van Leeuwen CJ, Chandy PC. The city blueprint: experiences with the implementation of 24 indicators to assess the sustainability of the urban water cycle. *Water Supply*. 2013;13(3):769-81.

28. Bobat A. Water scarcity in Kocaeli/Turkey: problems and measures. *The Journal of Scientific and Engineering Research*. 2019;6:39-52.

29. Türkiye Cumhuriyeti Çevre ve Şehircilik Bakanlığı. Atıksu Arıtma Tesisi ile Hizmet Verilen Belediyeler. Accessed: 2021-10-12. https://cevreselgostergeler.csb.gov.tr/atiksu-aritma-tesisi-ile-hizmet-verilen-belediyeler-i-85746.

30. TÜİK. Municipal Wastewater Statistics. Accessed: 2021-10-12. https://biruni.tuik.gov.tr/medas/?kn=120&locale=en.

31. Fetter FA. [Review of Theory of Location of Industries, by A. Weber C. J. Friedrich]. *Weltwirtschaftliches Archiv*. 1930;32:32-6. Available from: http://www.jstor.org/stable/40414003.

32. Kuehn AA, Hamburger MJ. A heuristic program for locating warehouses. *Management science*. 1963;9(4):643-66.

33. Hakimi SL. Optimum locations of switching centers and the absolute centers and medians of a graph. *Operations Research*. 1964;12(3):450-9.

34. ReVelle C. The maximum capture or “sphere of influence” location problem: Hotelling revisited on a network. *Journal of Regional Science*. 1986;26(2):343-58.

35. Toregas C, Swain R, ReVelle C, Bergman L. The location of emergency service facilities. *Operations Research*. 1971;19(6):1363-73.

36. Church R, ReVelle C. The maximal covering location problem. *Papers of the Regional Science Association*. 1974;32(1):101-18.

37. Elloumi S, Labbé M, Pochet Y. A new formulation and resolution method for the p-center problem. *INFORMS Journal on Computing*. 2004;16(1):84-94.

38. Klose A, Drexl A. Facility location models for distribution system design. *European Journal of Operational Research*. 2005;162(1):4-29.

39. Nagy G, Salhi S. Location-routing: Issues, models and methods. *European Journal of Operational Research*. 2007;177(2):649–672.

40. Alumur S, Kara BY. Network hub location problems: The state of the art. *European Journal of Operational Research*. 2008;190(1):1–21.

41. Revelle CS, Eiselt HA, Daskin MS. A bibliography for some fundamental problem categories in discrete location science. *European Journal of Operational Research*. 2008;184(3):817-48.

42. Melo MT, Nickel S, Saldanha-Da-Gama F. Facility location and supply chain management–A review. *European Journal of Operational Research*. 2009;196(2):401-12.

43. Magnanti TL, Wong RT. Network design and transportation planning: Models and algorithms. *Transportation Science*. 1984;18(1):1-55.

44. Balakrishnan A, Magnanti TL, Wong RT. A dual-ascent procedure for large-scale uncapacitated network design. *Operations Research*. 1989;37(5):716-40.

45. Balakrishnan A, Magnanti TL, Mirchandani P. A dual-based algorithm for multi-level network design. *Management Science*. 1994;40(5):567-81.

46. Gendron B, Crainic TG, Frangioni A. Multicommodity capacitated network design. *Telecommunications Network Planning*. Springer. 1999:1-19.

47. Drezner Z, Wesolowsky GO. Network design: selection and design of links and facility location. *Transportation Research Part A: Policy and Practice*. 2003;37(3):241-56.

48. Yang H, H Bell MG. Models and algorithms for road network design: a review and some new developments. *Transport Reviews*. 1998;18(3):257-78.

49. Croxton KL, Gendron B, Magnanti TL. Variable disaggregation in network flow problems with piecewise linear costs. *Operations Research*. 2007;55(1):146-57.

50. Costa AM, Cordeau JF, Gendron B. Benders, metric and cutset inequalities for multicommodity capacitated network design. *Computational Optimization and Applications*. 2009;42(3):371-92.

51. Frangioni A, Gendron B. 0–1 reformulations of the multicommodity capacitated network design problem. *Discrete Applied Mathematics*. 2009;157(6):1229-41.

52. Hewitt M, Nemhauser GL, Savelsbergh MW. Combining exact and heuristic approaches for the capacitated fixed-charge network flow problem. *INFORMS Journal on Computing*. 2010;22(2):314-25.

53. Daskin MS, Hurter AP, VanBuer MG. Toward an integrated model of facility location and transportation network design. Transportation Center, Northwestern University; 1993.

54. Melkote S. Integrated models of facility location and network design. Ph.D. Thesis, , 1996.

55. Melkote S, Daskin MS. An integrated model of facility location and transportation network design. *Transportation Research Part A: Policy and Practice*. 2001;35(6):515-38.

56. Melkote S, Daskin MS. Capacitated facility location/network design problems.

European Journal of Operational Research. 2001;129(3):481–495.

57. Ravi R, Sinha A. Approximation algorithms for problems combining facility location and network design. *Operations Research*. 2006;54(1):73-81.

58. Chen X, Chen B. Approximation algorithms for soft-capacitated facility location in capacitated network design. *Algorithmica*. 2009;53(3):263-97.

59. Mortezaei M, JabalAmeli M. A hybrid model for multi-objective capacitated facility location network design problem. *International Journal of Industrial Engineering Computations*. 2011;2(3):509-24.

60. Cocking C. Solutions to facility location-network design problems. Ph.D. Thesis, , 2008.

61. Cocking C, Reinelt G, Pfetsch ME. Heuristics for budget facility location-network design problems with minisum objective. *Operations Research Proceedings 2008*. Springer. 2009:563-8.

62. Bigotte JF, Krass D, Antunes AP, Berman O. Integrated modeling of urban hierarchy and transportation network planning. *Transportation Research Part A: Policy and Practice*. 2010;44(7):506–522.

63. Contreras I, Fernández E, Reinelt G. Minimizing the maximum travel time in a combined model of facility location and network design. *Omega*. 2012;40(6):847-60.

64. Rahmaniani R, Ghaderi A. A combined facility location and network design problem with multi-type of capacitated links. *Applied Mathematical Modelling*. 2013;37(9):6400-14.

65. Lashine SH, Fattouh M, Issa A. Location/allocation and routing decisions in supply chain network design. *Journal of Modelling in Management*. 2006;1(2):173-83.

66. Shishebori D, Jabalameli MS, Jabbarzadeh A. Facility location-network design problem: reliability and investment budget constraint. *Journal of Urban Planning and Development*. 2014;140(3):04014005.

67. Wu J, Liu M, Guo X. An Integrated Model of Capacitated Facility Location and

Network Structure. *LISS 2013*. Berlin, Heidelberg: Springer. 2015:1347-53.

68. Kaya O, Ozkok D. A blood bank network design problem with integrated facility location, inventory and routing decisions. *Networks and Spatial Economics*. 2020;20(3):757-83.

69. Liu X, Kwon C. Exact robust solutions for the combined facility location and network design problem in hazardous materials transportation. *IISE Transactions*. 2020;52(10):1156-72.

70. Brahami MA, Dahane M, Souier M, Sahnoun M. Sustainable capacitated facility location/network design problem: a non-dominated sorting genetic algorithm based multiobjective approach. *Annals of Operations Research*. 2020:1-32.

71. Sadat Asl A, Fazel Zarandi M, Sotudian S, Amini A. A fuzzy capacitated facility location-network design model: A hybrid firefly and invasive weed optimization (FIWO) solution. *Iranian Journal of Fuzzy Systems*. 2020;17(2):79-95.

72. Ghaderi A, Jabalameli MS. Modeling the budget-constrained dynamic uncapacitated facility location–network design problem and solving it via two efficient heuristics: a case study of health care. *Mathematical and Computer Modelling*. 2013;57(3-4):382-400.

73. Ghaderi A. Heuristic algorithms for solving an integrated dynamic center facility location-network design model. *Networks and Spatial Economics*. 2015;15(1):43-69.

74. Pearce RH, Forbes M. Disaggregated Benders decomposition and branch-and-cut for solving the budget-constrained dynamic uncapacitated facility location and network design problem. *European Journal of Operational Research*. 2018;270(1):78-88.

75. Manne AS. Capacity expansion and probabilistic growth. *Econometrica: Journal of the Econometric Society*. 1961;29(4):632.

76. Ballou RH. Dynamic warehouse location analysis. *Journal of Marketing Research*. 1968;5(3):271-6.

77. Scott AJ. Dynamic location-allocation systems: some basic planning strategies.

Environment and Planning A: Economy and Space. 1971;3(1):73-82.

78. Sweeney DJ, Tatham RL. An improved long-run model for multiple warehouse location. *Management Science*. 1976;22(7):748-58.

79. Drezner Z, Wesolowsky G. Facility location when demand is time dependent. *Naval Research Logistics (NRL)*. 1991;38(5):763-77.

80. Bastian M, Volkmer M. A perfect forward procedure for a single facility dynamic location/relocation problem. *Operations Research Letters*. 1992;12(1):11-6.

81. Andreatta G, Mason F. A note on “A perfect forward procedure for a single facility dynamic location/relocation problem”. *Operations Research Letters*. 1994;15(2):81-3.

82. Wesolowsky GO. Dynamic facility location. *Management Science*. 1973;19(11):1241-8.

83. Wesolowsky GO, Truscott WG. The multiperiod location-allocation problem with relocation of facilities. *Management Science*. 1975;22(1):57-65.

84. Shulman A. An algorithm for solving dynamic capacitated plant location problems with discrete expansion sizes. *Operations Research*. 1991;39(3):423-36.

85. Chardaire P, Sutter A, Costa MC. Solving the dynamic facility location problem. *Networks: An International Journal*. 1996;28(2):117-24.

86. Dias J, Captivo ME, Clímaco J. Efficient primal-dual heuristic for a dynamic location problem. *Computers Operations Research*. 2007;34(6):1800–1823.

87. Hinojosa Y, Puerto J, Fernández Fr. A multiperiod two-echelon multicommodity capacitated plant location problem. *European Journal of Operational Research*. 2000;123(2):271–291.

88. Melo T, Nickel S, da Gama FS. An LP-rounding heuristic to solve a multi-period facility relocation problem. Fraunhofer (ITWM); 2009. 168.

89. Thanh PN, Péton O, Bostel N. A linear relaxation-based heuristic approach for logistics network design. *Computers & Industrial Engineering*. 2010;59(4):964-75.

90. Lee DH, Dong M. Dynamic network design for reverse logistics operations under uncertainty. *Transportation Research Part E: Logistics and Transportation Review*. 2009;45(1):61–71.

91. Ukkusuri S, Ozbay K. *Advances in Dynamic Network Modeling in Complex Transportation Systems*. vol. 2; 2013.

92. Hugo A, Pistikopoulos EN. Environmentally conscious long-range planning and design of supply chain networks. *Journal of Cleaner Production*. 2005;13(15):1471-91.

93. Troncoso JJ, Garrido RA. Forestry production and logistics planning: an analysis using mixed-integer programming. *Forest Policy and Economics*. 2005;7(4):625-33.

94. Aghezzaf E. Capacity planning and warehouse location in supply chains with uncertain demands. *Journal of the Operational Research Society*. 2005;56(4):453-62.

95. Fleischmann B, Ferber S, Henrich P. Strategic planning of BMW's global production network. *Interfaces*. 2006;36(3):194-208.

96. Ko HJ, Evans GW. A genetic algorithm-based heuristic for the dynamic integrated forward/reverse logistics network for 3PLs. *Computers & Operations Research*. 2007;34(2):346-66.

97. Julka N, Baines T, Tjahjono B, Lendermann P, Vitanov V. A review of multi-factor capacity expansion models for manufacturing plants: Searching for a holistic decision aid. *International Journal of Production Economics*. 2007;106(2):607-21.

98. Badri H, Bashiri M, Hejazi TH. Integrated strategic and tactical planning in a supply chain network design with a heuristic solution method. *Computers & Operations Research*. 2013;40(4):1143-54.

99. Loucks DP, Revelle CS, Lynn WR. Linear progamming models for water pollution control. *Management Science*. 1967;14(4):B-166.

100. Converse A. Optimum number and location of treatment plants. *Journal (Water Pollution Control Federation)*. 1972;44(8):1629-36.

101. Jarvis JJ, Rardin RL, Unger V, Moore RW, Schimpeler CC. Optimal design of

regional wastewater systems: A fixed-charge network flow model. *Operations Research*. 1978;26(4):538-50.

102. Kovács L, Boros E, Inotay F. A two-stage approach for large-scale sewer systems design with application to the Lake Balaton resort area. *European Journal of Operational Research*. 1986;23(2):169–178.

103. Zhu ZP, Revelle C, Rosing K. Adaptation of the plant location model for regional environmental facilities and cost allocation strategy. *Annals of Operations Research*. 1989;18(1):345–366.

104. Leitão J, Matos J, Goncalves AB, Matos J. Contribution of geographic information systems and location models to planning of wastewater systems. *Water Science and Technology*. 2005;52(3):1-8.

105. Zeferino JA, Antunes AP, Cunha MC. Multi-objective model for regional wastewater systems planning. *Civil Engineering and Environmental Systems*. 2010;27(2):95-106.

106. Zeferino JA, Antunes AP, Cunha MC. Regional wastewater system planning under population dynamics uncertainty. *Journal of Water Resources Planning and Management*. 2014;140(3):322-31.

107. Rezaei N, Sierra-Altamiranda A, Diaz-Elsayed N, Charkhgard H, Zhang Q. A multi-objective optimization model for decision support in water reclamation system planning. *Journal of Cleaner Production*. 2019;240:118227.

108. Zhao K, Ng TSA, Liu X. A guarantee rate optimization model for wastewater treatment system design under uncertainty. *Naval Research Logistics (NRL)*. 2020;67(6):420–437.

109. Brand N, Ostfeld A. Optimal design of regional wastewater pipelines and treatment plant systems. *Water Environment Research*. 2011;83(1):53-64.

110. Ozturk I. *Atıksu Mühendisliği*. İstanbul, Türkiye: İSKİ Teknik Kitaplar Serisi; 2017.

111. Lim SR, Park D, Park JM. Environmental and economic feasibility study of a total wastewater treatment network system. *Journal of Environmental Management*. 2008;88(3):564–575.

112. Hernández-Chover V, Bellver-Domingo Hernández-Sancho F. Efficiency of wastewater treatment facilities: The influence of scale economies. *Journal of Environmental Management*. 2018;228:77–84.

113. Sullivan WG, Wicks EM, Koelling CP. *Engineering economy*. Pearson; 2015.

114. Harris CR, et al. Array programming with NumPy. *Nature*. 2020;585(7825):357-62.

115. Winston WL, Goldberg JB. *Operations research: applications and algorithms*. vol. 3. Thomson Brooks/Cole Belmont; 2004.

116. Luss H. Operations research and capacity expansion problems: A survey. *Operations Research*. 1982;30(5):907-47.

117. Hendrickson C, Au T. Fundamental concepts for owners, engineers, architects, and builders. Prentice Hall, Englewood Cliffs, NJ. 2008.

118. Moran S. Chapter 10 - Clean water hydraulics. *An Applied Guide to Water and Effluent Treatment Plant Design*. Butterworth-Heinemann. 2018:117-35.

119. Roache PJ. *Verification and validation in computational science and engineering*. vol. 895. Hermosa Albuquerque, NM; 1998.

120. Banks J, Carson J, Nelson B, Nicol D. *Discrete-Event System Simulation*. 5th ed. Prentice Hall; 2010.

121. Schrage LE. *Optimization modeling with LINGO*. Lindo System; 2006.

122. MacQueen J. Classification and analysis of multivariate observations. *5th Berkeley Symp. Math. Statist. Probability*. 1967:281-97.

123. Ester M, Kriegel HP, Sander J, Xu X, et al. A density-based algorithm for discovering clusters in large spatial databases with noise. *KDD*. vol. 96. 1996:226-31.

124. Duda RO, Hart PE, Stork DG. *Pattern classification and scene analysis*. vol. 3. Wiley New York; 1973.

125. Zhang T, Ramakrishnan R, Livny M. BIRCH: an efficient data clustering method for very large databases. *ACM Sigmod Record*. 1996;25(2):103-14.

126. Pehlivan SA, Demirel DF, Başak SM. A time-variant capacitated facility location and network design problem for wastewater treatment systems. *19th International Logistics and Supply Chain Congress (LMSCM2021)*. 2021:323.

APPENDIX A: DETAILS OF E.A3

Table A.1. Constructed facilities (Z_i^{lt}) for $E.A3$

Node (i)	**Level** (l)	**Time Period** (t)
4	4	-2
7	4	-2
3	3	-1
9	3	-1

Table A.2. Established links (X_{ij}^t) for $E.A3$

Link (i, j)	**Time Period** (t)
(1,9)	0
(2,4)	0
(5,3)	0
(6,7)	0
(8,6)	0
(9,7)	0
(10,9)	0
(5,4)	1

Table A.3. W_i^{kt} for $E.A3$

W_i^{kt}		Time Period (t)									
Source (i)	**Demand Point (k)**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
7	1							0.26	0.28	0.29	0.29
9	1	1.00	1.00	1.00	1.00	1.00	1.00	0.74	0.72	0.71	0.71
4	2	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
3	3	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
4	4	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
3	5	1.00	1.00	1.00	1.00	1.00	0.91	0.83	0.79	0.78	0.79
4	5						0.09	0.17	0.21	0.22	0.21
7	6	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
7	7	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
7	8	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
7	9	0.12	0.11	0.29	0.42	0.51	0.54				
9	9	0.88	0.89	0.71	0.58	0.49	0.46	1.00	1.00	1.00	1.00
9	10	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

Table A.4. Y_{ij}^{kt} for $E.A3$

Y_{ij}^{kt}		Time Period (t)									
Link (i, j)	**Demand Point (k)**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
(1,9)	1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
(2,4)	2	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
(5,3)	5	1.00	1.00	1.00	1.00	1.00	0.91	0.83	0.79	0.78	0.79
(5,4)	5						0.09	0.17	0.21	0.22	0.21
(6,7)	6	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
(6,7)	8	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
(8,6)	8	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
(9,7)	1							0.26	0.28	0.29	0.29
(9,7)	9	0.12	0.11	0.29	0.42	0.51	0.54				
(10,9)	10	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

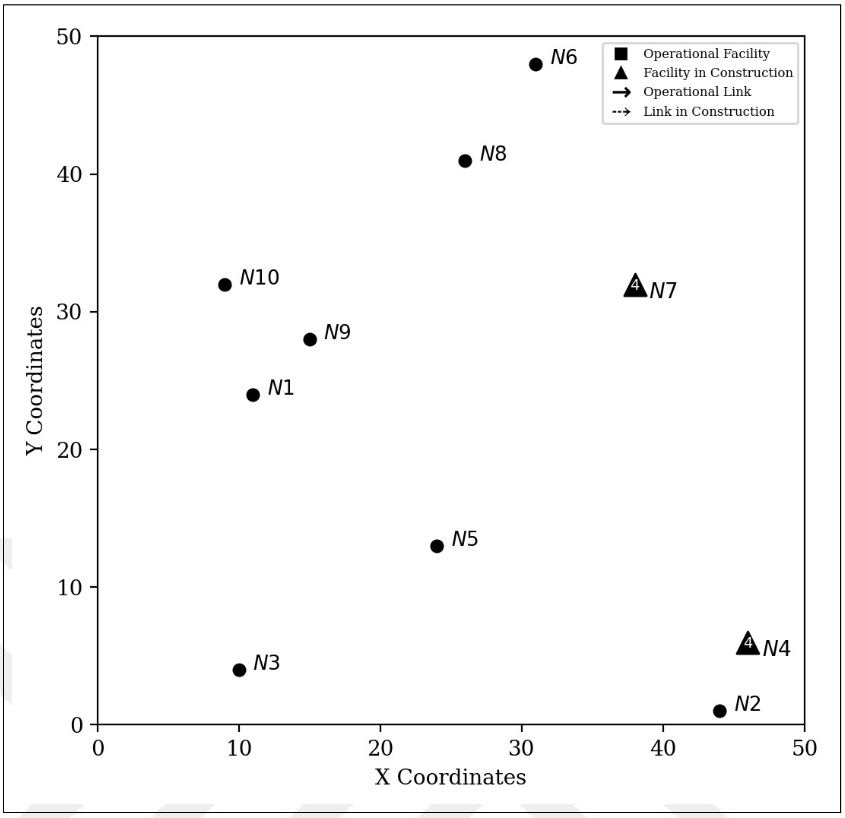


Figure A.1. Complete network in $E.A3$ ($t = -2$)

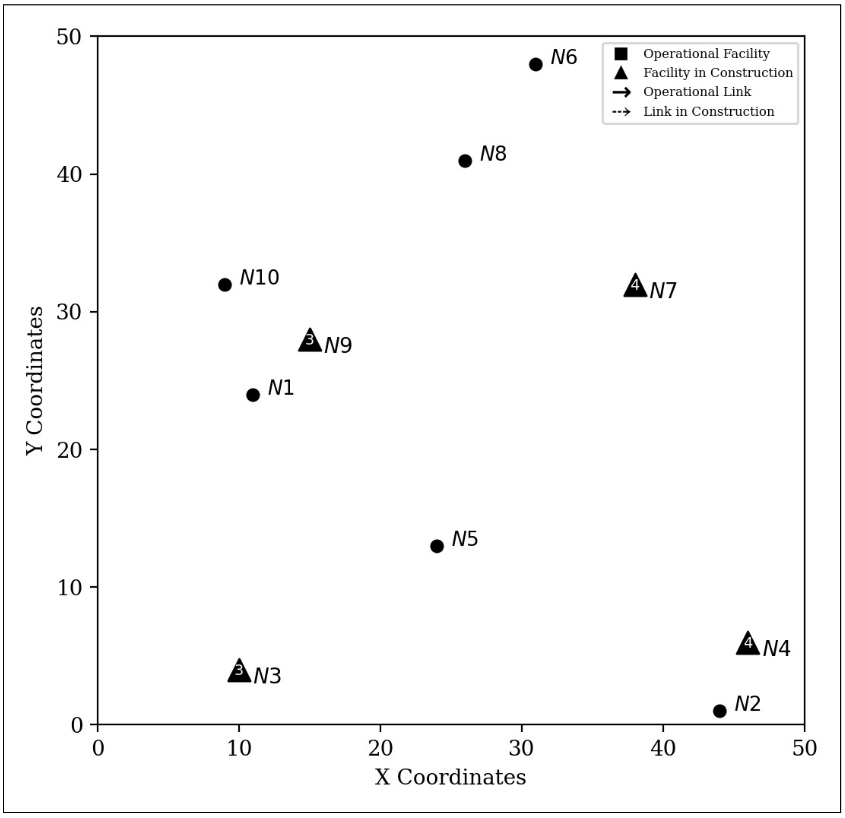


Figure A.2. Complete network in $E.A3$ ($t = -1$)

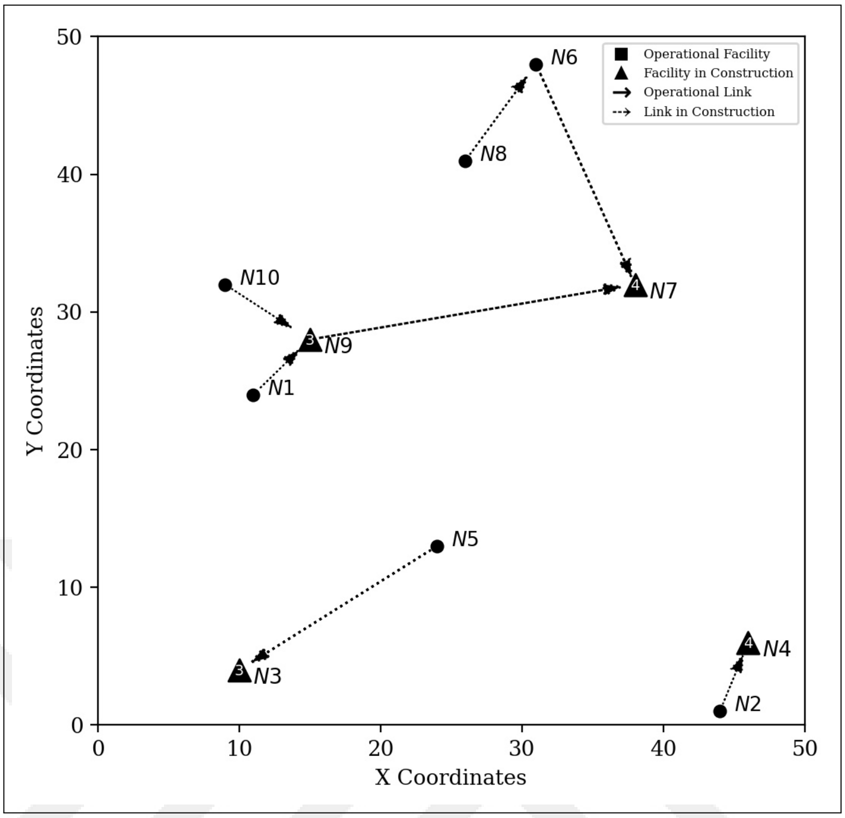


Figure A.3. Complete network in $E.A3$ ($t = 0$)

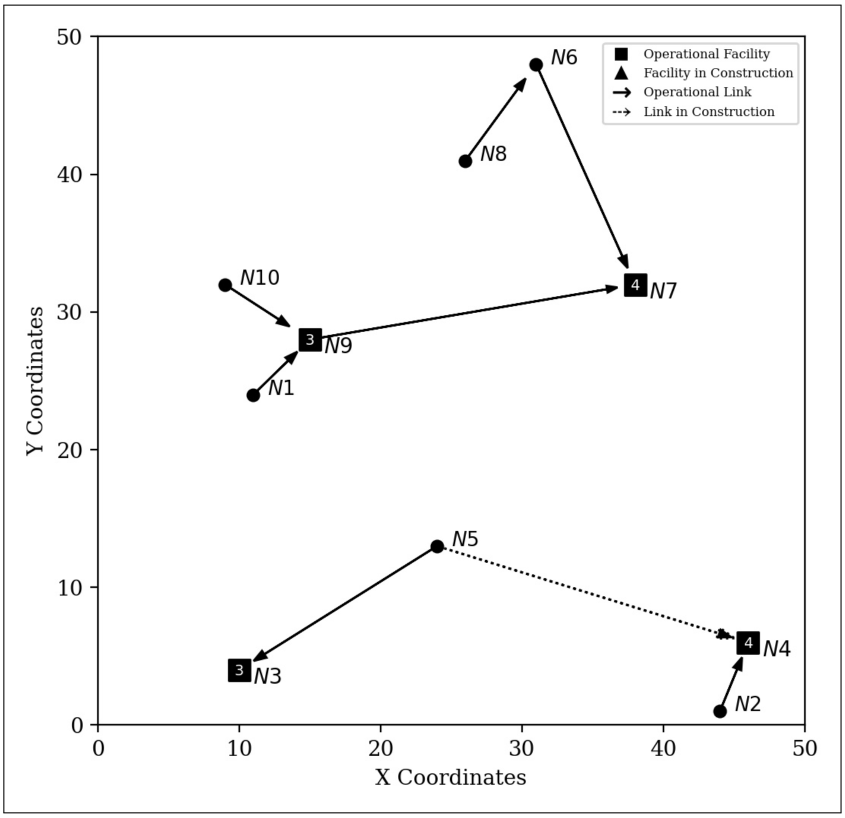


Figure A.4. Complete network in $E.A3$ ($t = 1$)

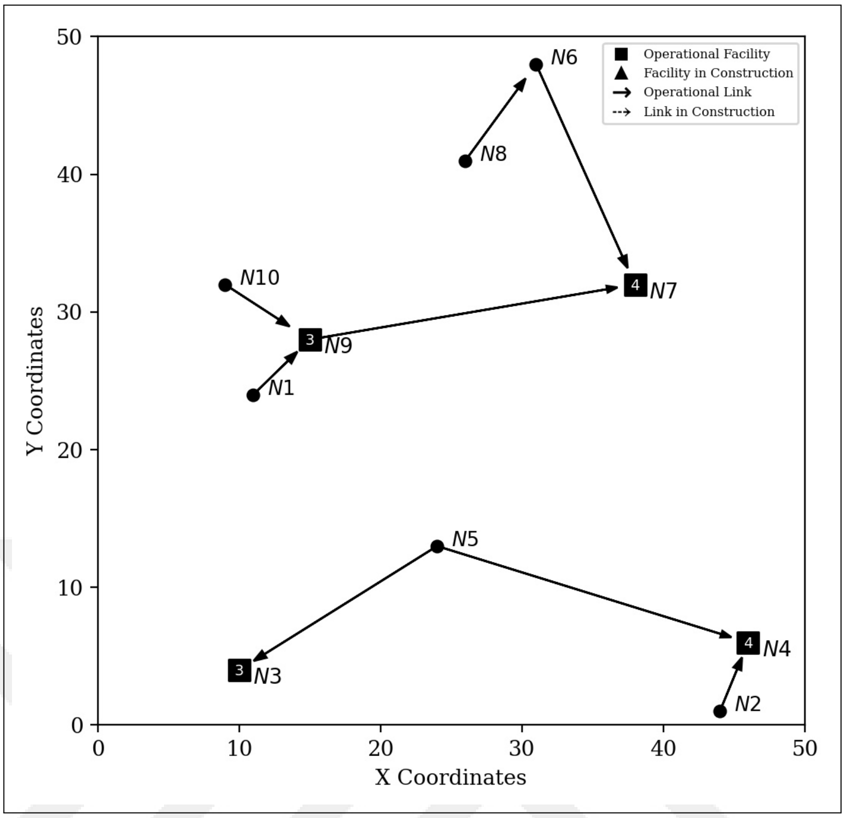


Figure A.5. Complete network in $E.A3$ ($t \in [2, 10]$)

APPENDIX B: DETAILS OF E.B8

Table B.1. Constructed facilities (Z_i^{lt}) for $E.B8$

Node (i)	**Level** (l)	**Time Period** (t)
6	5	-2
8	5	-2
7	1	-1
4	3	0
3	1	18

Table B.2. Established links (X_{ij}^{pt}) for $E.B8$

Link (i, j)	**Link Type** (p)	**Time Period** (t)
(1,8)	2	0
(2,8)	2	0
(3,6)	3	0
(4,6)	3	0
(5,6)	3	0
(7,9)	2	0
(9,8)	3	0
(10,8)	3	0
(2,4)	2	3

Table B.3. Facility expansions ($EX_i^{l\zeta t}$) for $E.B8$

Node (i)	**Initial Level** (l)	**Expanded Level** (ζ)	**Time Period** (t)
7	1	4	11

Table B.4. W_i^{kt} for $E.B8$ ($t \in [1, 10]$)

W_i^{kt}		Time Period (t)									
Source (i)	**Demand Point (k)**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
8	1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
4	2									0.50	
6	2				0.01	0.13	0.19	0.31	0.39		0.51
8	2	1.00	1.00	1.00	0.99	0.87	0.81	0.69	0.61	0.50	0.49
6	3	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
4	4		0.67	0.67	0.67	0.65	0.63	0.60	0.59	0.16	0.54
6	4	1.00	0.33	0.33	0.33	0.35	0.37	0.40	0.41	0.84	0.46
6	5	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
6	6	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
7	7	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.06
8	7	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.94
8	8	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.0
8	9	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
8	10	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

Table B.5. W_i^{kt} for $E.B8$ ($t \in [11, 20]$)

W_i^{kt}		Time Period (t)									
Source (i)	**Demand Point (k)**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**
8	1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
4	2										0.07
6	2	0.62	0.74								
8	2	0.38	0.26	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.93
3	3										0.05
6	3	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.95
4	4	0.53	0.53	0.52	0.52	0.50	0.49	0.47	0.46	0.45	0.39
6	4	0.47	0.47	0.48	0.48	0.50	0.51	0.53	0.54	0.55	0.61
6	5	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
6	6	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
7	7	0.06	0.06	1.00	1.00	1.00	1.00	1.00	1.00	0.99	0.97
8	7	0.94	0.94							0.01	0.03
8	8	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
8	9	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
8	10	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

Table B.6. Y_{ij}^{kpt} for $E.B8$ ($t \in [1, 10]$)

Y_{ij}^{kpt}			**Time Period (t)**									
Link (i, j)	**Demand Point (k)**	**Link Type (p)**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
(1,8)	1	2	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
(2,4)	2	2				0.01	0.13	0.19	0.31	0.39	0.50	0.51
(2,8)	2	2	1.00	1.00	1.00	0.99	0.87	0.81	0.69	0.61	0.50	0.49
(3,6)	3	3	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
(4,6)	2	3				0.01	0.13	0.19	0.31	0.39		0.51
(4,6)	4	3	1.00	0.33	0.33	0.33	0.35	0.37	0.40	0.41	0.84	0.46
(5,6)	5	3	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
(7,9)	7	2	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.94
(9,8)	7	3	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.94
(9,8)	9	3	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
(10,8)	10	3	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

Table B.7. Y_{ij}^{kpt} for $E.B8$ ($t \in [11, 20]$)

Y_{ij}^{kpt}			**Time Period (t)**									
Link (i, j)	**Demand Point (k)**	**Link Type (p)**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**
(1,8)	1	2	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
(2,4)	2	2	0.62	0.74								0.07
(2,8)	2	2	0.38	0.26	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.93
(3,6)	3	3	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.95
(4,6)	2	3	0.62	0.74								
(4,6)	4	3	0.47	0.47	0.48	0.48	0.50	0.51	0.53	0.54	0.55	0.61
(5,6)	5	3	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
(7,9)	7	2	0.94	0.94							0.01	0.03
(9,8)	7	3	0.94	0.94							0.01	0.03
(9,8)	9	3	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
(10,8)	10	3	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

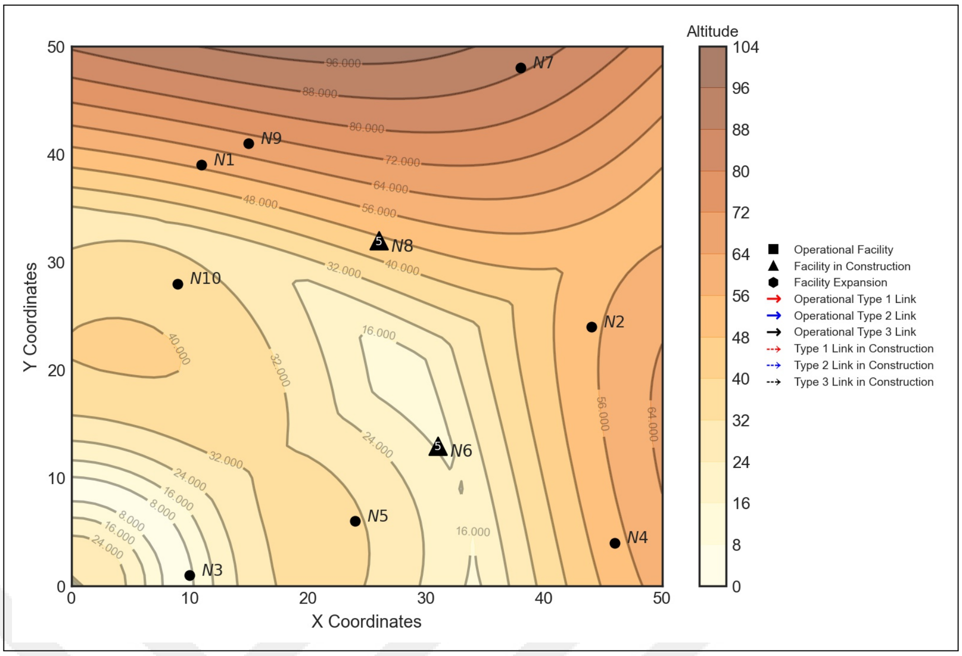


Figure B.1. Complete network in $E.B8$ ($t = -2$)

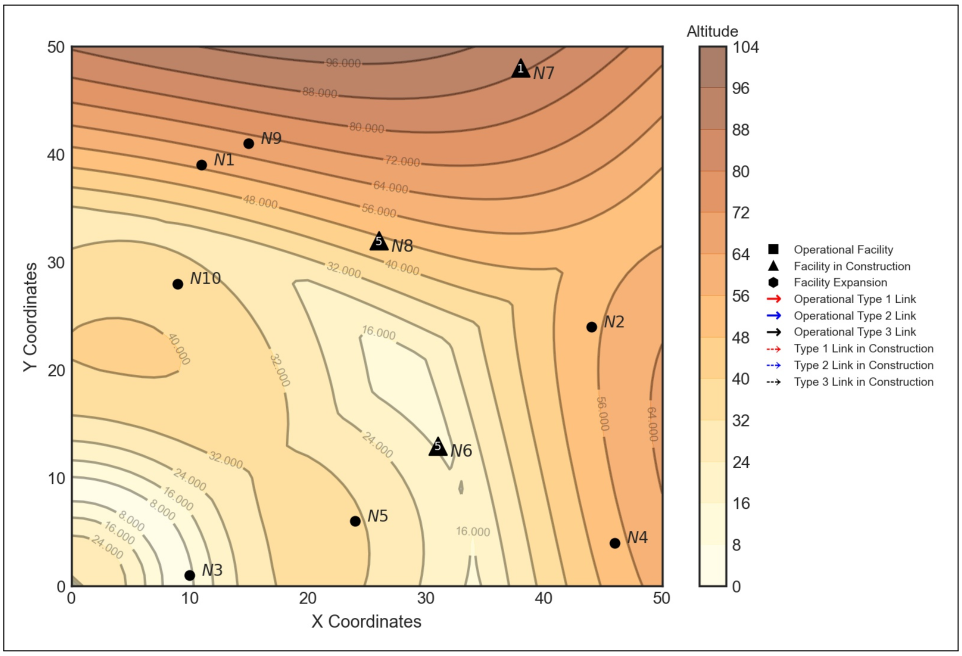


Figure B.2. Complete network in $E.B8$ ($t = -1$)

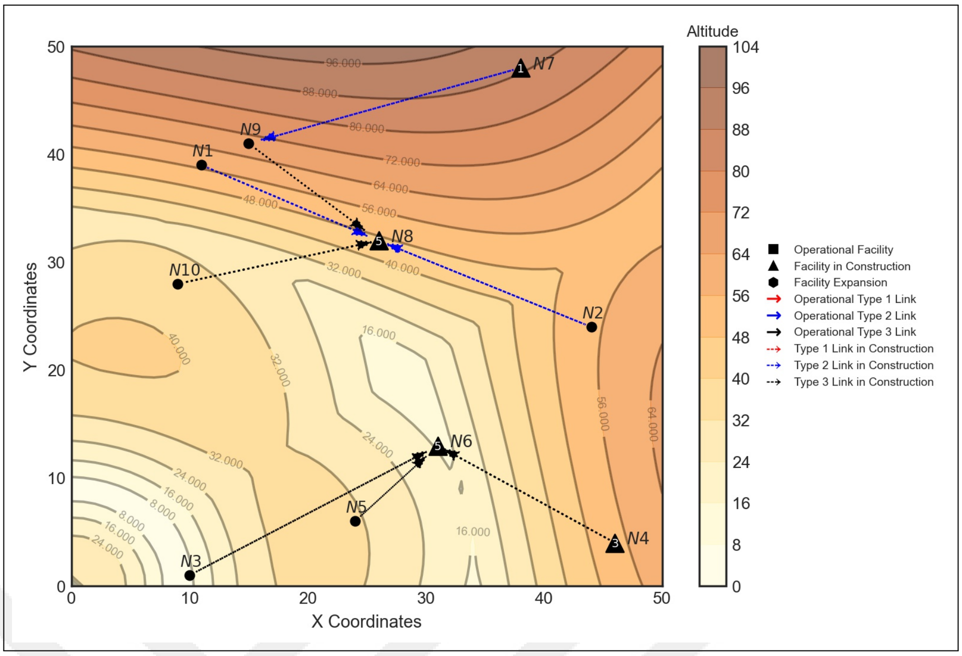


Figure B.3. Complete network in $E.B8$ ($t = 0$)

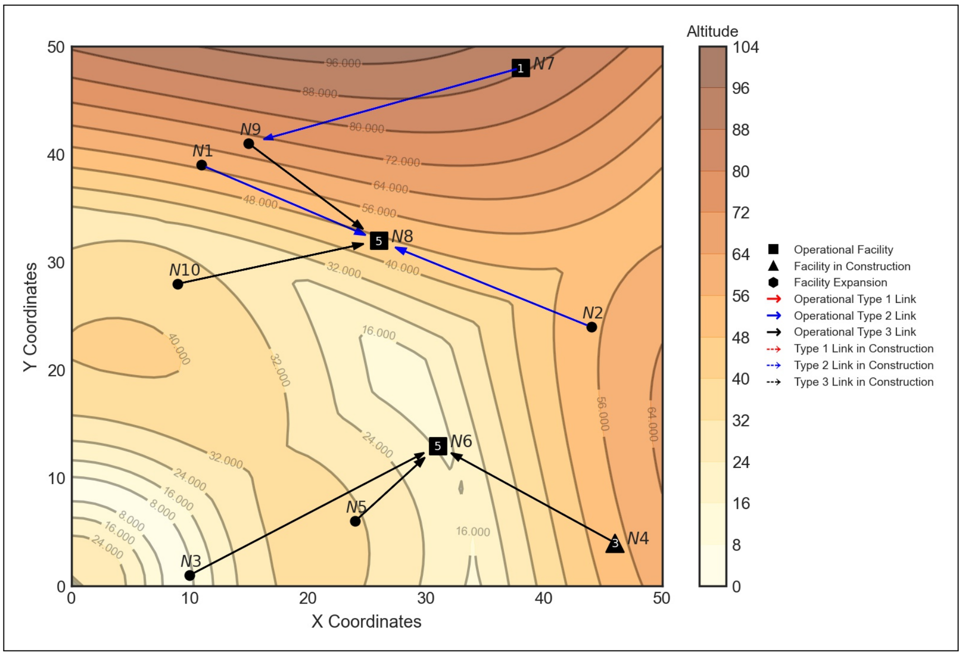


Figure B.4. Complete network in $E.B8$ ($t = 1$)

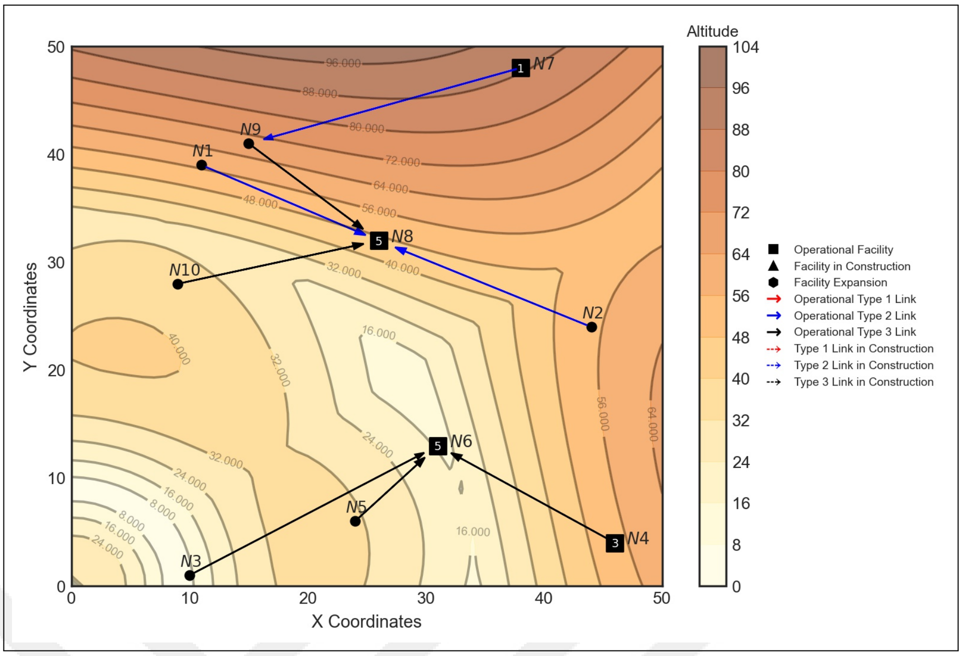


Figure B.5. Complete network in $E.B8$ ($t = 2$)

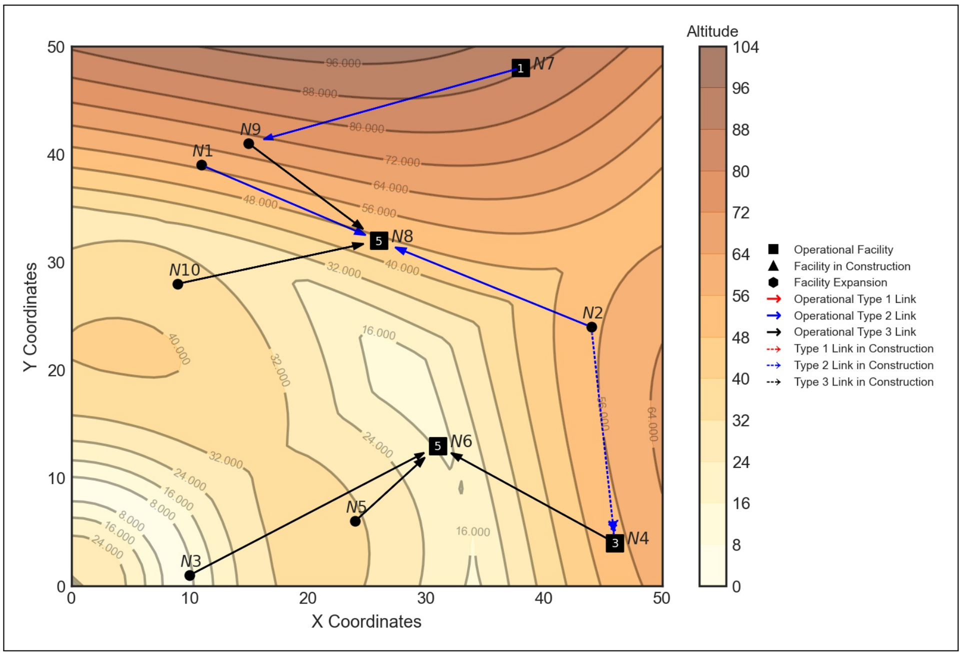


Figure B.6. Complete network in $E.B8$ ($t = 3$)

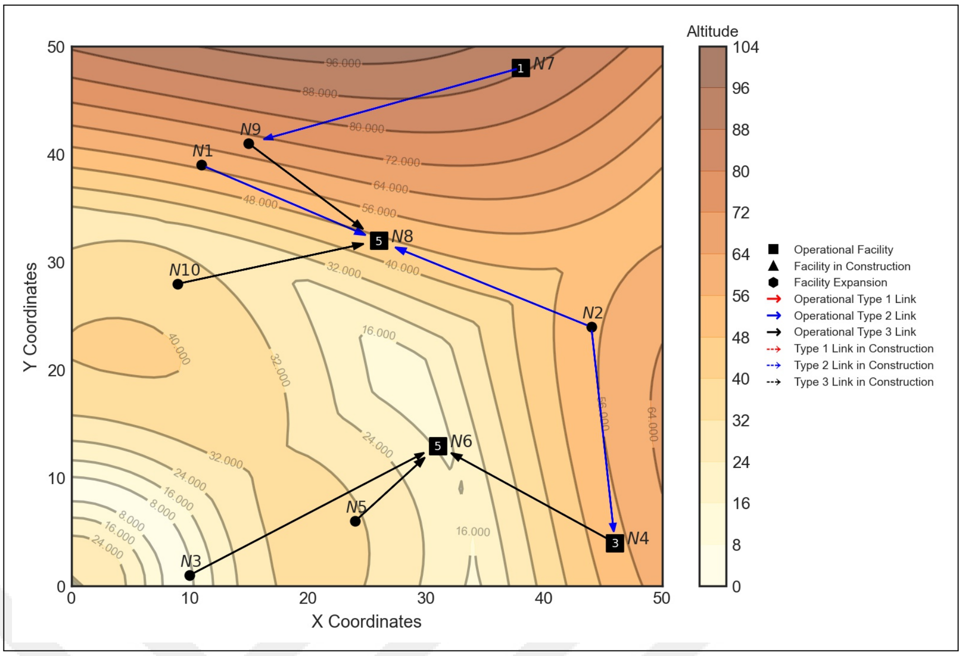


Figure B.7. Complete network in $E.B8$ ($t \in [4, 10]$)

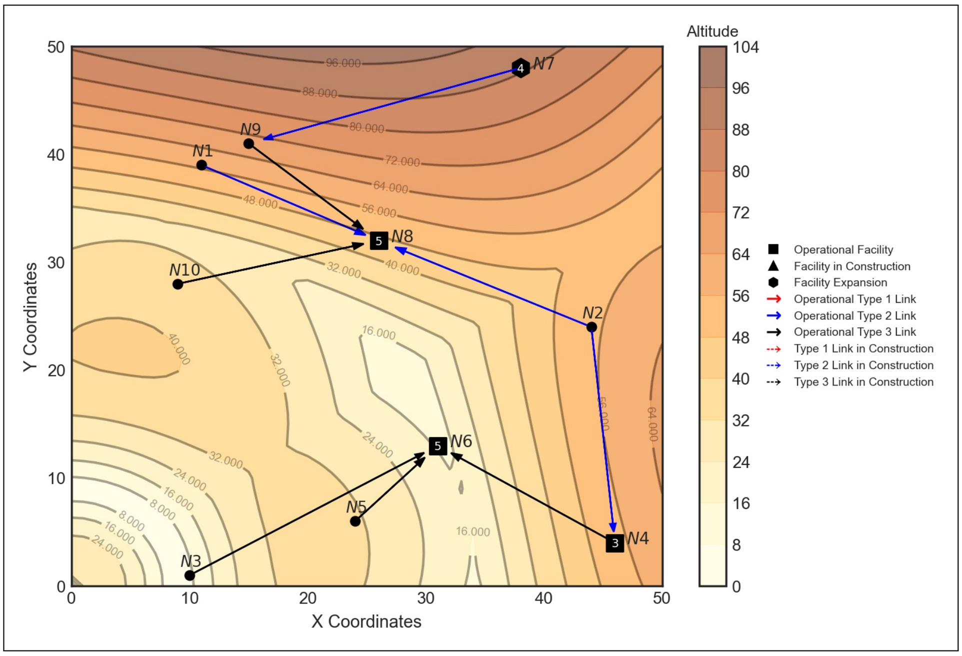


Figure B.8. Complete network in $E.B8$ ($t \in [11, 12]$)

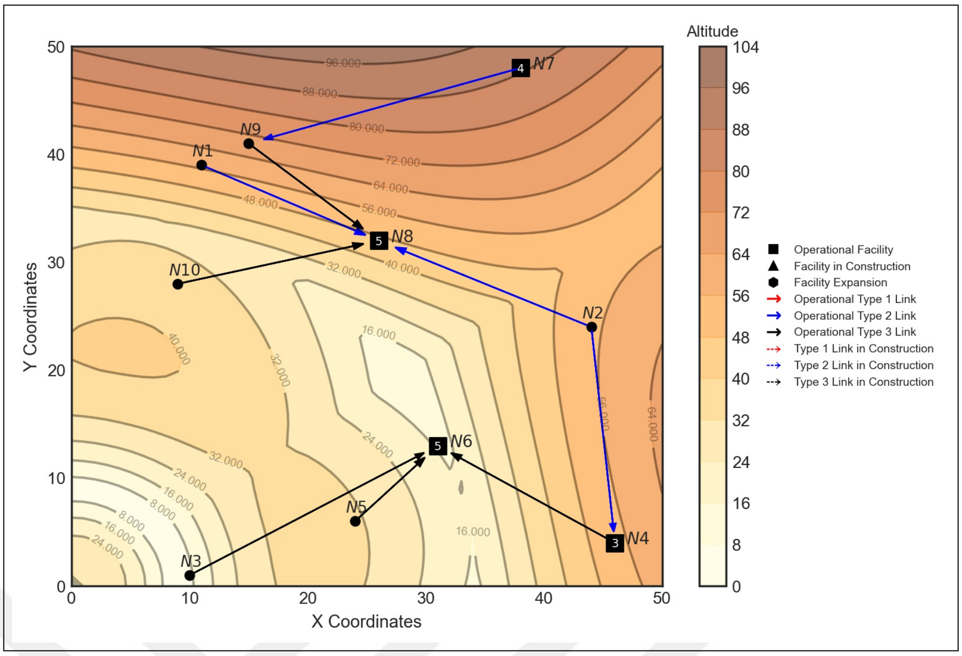


Figure B.9. Complete network in $E.B8$ ($t \in [13, 17]$)

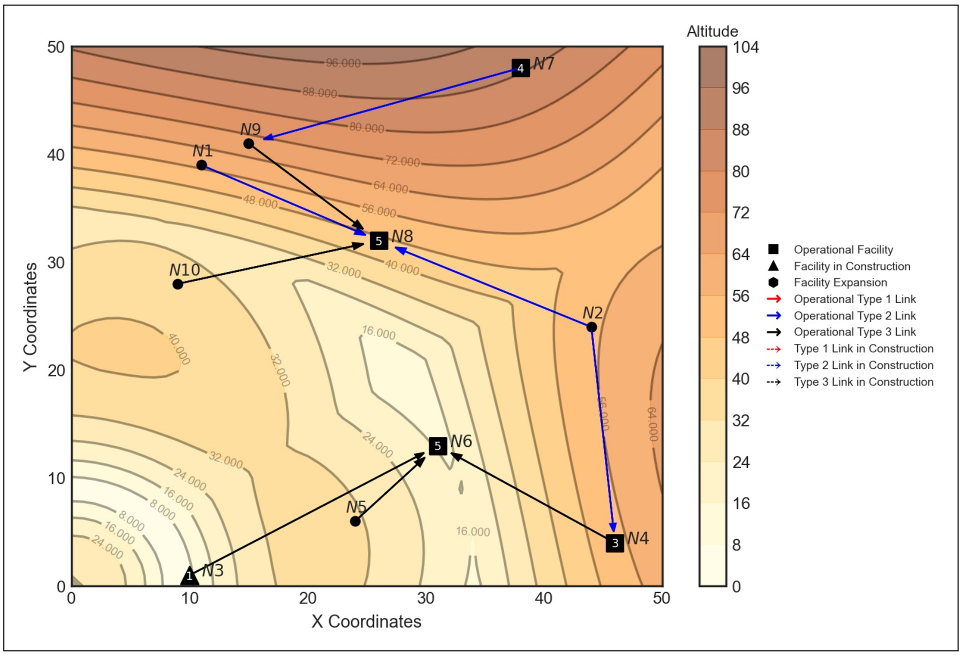


Figure B.10. Complete network in $E.B8$ ($t \in [18, 19]$)

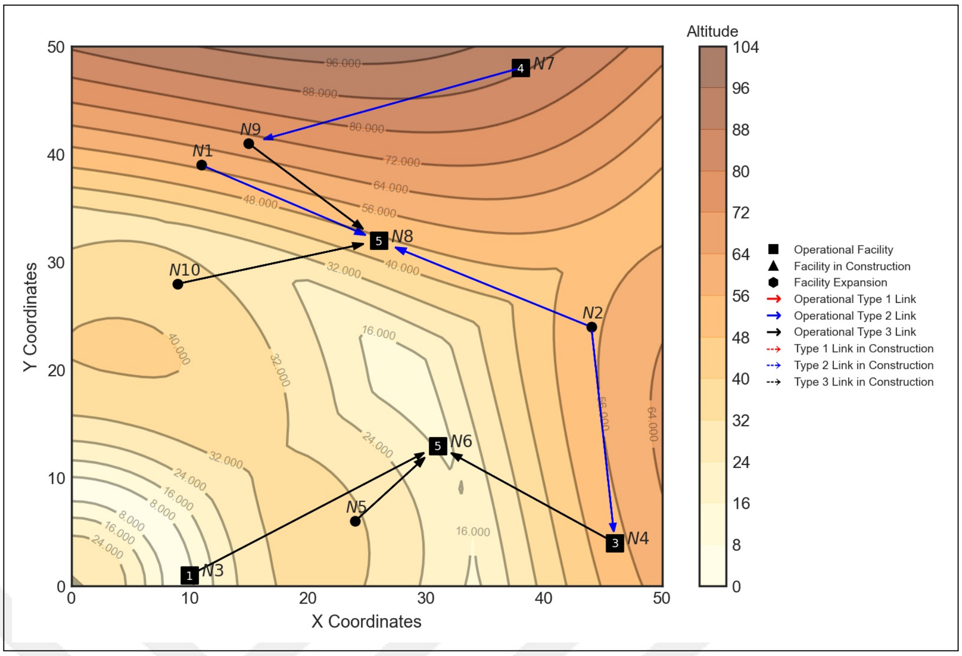


Figure B.11. Complete network in $E.B8$ ($t = 20$)