DYNAMIC MODELING OF WATER DISTRIBUTION NETWORK BY COMPLEX NETWORK APPROACH

by

FANG ZENG

(Under the Direction of Ke Li)

ABSTRACT

The complex topology and adaptive behavior of infrastructure systems are driven by both self-organization of the demand and rigid engineering solutions. Therefore, the engineering complex system requires a method balancing holism and reductionism. To model the growth of water distribution networks (WDNs), a complex network model was developed following the combination of local optimization rules and engineering considerations. The demand node generation is dynamic and follows the scaling law of urban growth. The proposed model can generate WDN similar to reported real-world WDNs on some structural properties. Comparison with different modeling approaches indicates that a realistic demand node distribution and co-evolvement of demand node and network are important for the simulation of real complex networks. Combining with a pipe sizing algorithm, this dynamic model was used to investigate the effect of system size on cost efficiency of WDNs and find the optimum planning horizon for WDN expansion design.

INDEX WORDS: Water distribution network, Complex network, Dynamic model

DYNAMIC MODELING OF WATER DISTRIBUTION NETWORK BY COMPLEX NETWORK APPROACH

by

FANG ZENG

B.S., Suzhou University of Science and Technology, China, 2008

M.S., Nankai University, China, 2011

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2016

DYNAMIC MODELING OF WATER DISTRIBUTION NETWORK BY COMPLEX NETWORK APPROACH

by

FANG ZENG

Major Professor: Ke Li

Committee: Caner Kazanci

Jason Christian

Electronic Version Approved:

Suzanne Barbour

Dean of the Graduate School

The University of Georgia

December 2016

ACKNOWLEDGEMENTS

First of all, I would like thank my advisor, Dr. Ke Li, in his effort of guiding me through my research works during my PhD study. The enormous helps and instructions from him on how to perform researches from an engineering perspective are invaluable.

Moreover, I would like to thank Dr. Caner Kazanci, and Dr. Jason Christian for joining my dissertation committee. Their insightful suggestions and encouragements are most helpful for me to finish the dissertation work.

Finally, I want to thank my family in supporting me to pursuit my PhD studies. Their trust and help is invaluable through these 5 years of my life, in all the aspects.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1: INTRODUCTION

1.1 Dissertation statements

The trend of population growth and rapid urbanization worldwide demand a surge of the service providing infrastructures. An earlier projection shows that the urban infrastructure in the United States will double by 2030 from the status in 2000 [1]. The aging of existing infrastructure adds more financial and technical challenges. It is projected that approximately $3.6 trillion is needed by 2020 to maintain the infrastructure in the United States [2]. While this is a tremendous challenge, it provides an unprecedented opportunity for us to shape the sustainability and resilience of the future infrastructures. The key to seize the opportunity is the development of understanding of infrastructure systems as complex adaptive systems and transformation of the understanding into tangible principals and guidance for infrastructure planning and engineering practices. Landmarked by Ottino's essay [3] published in Nature, searching for a theory or organized framework to guide the engineering of complex systems has been ongoing for more than a decade.

Infrastructure systems are composed of many interacting components that yield outcomes as "the whole is greater than the sum of the parts", which makes them resistant to investigation using the traditional "divide and conquer" engineering approach. On the other hand, the complexity topology and adaptive behavior of infrastructure systems are driven by both self-organization of the demand and rigid engineering solutions. Therefore, the understanding of complex infrastructure system needs a balanced holism and reductionist methodology [4]. This requires predictability of the complex network topology, complex behavior and their evolution dynamics over time with organic integration of engineering rules.

Most infrastructure systems are spatial networks. As reviewed in [5], extensive research has been carried out to analytically characterize, describe and model spatial infrastructure networks. Aggregated indicators, such as node degree distribution, clustering coefficient, were used to describe the network topologies and derive useful information about the connectivity and efficiency of the systems. Study of many real infrastructure network however found deviation from small-world [6] and scale-free [7] networks. For planar networks where spatial limitation is significant, the node degree is usually peaked instead of following power law as observed in the scale-free network [8, 9], and shortest paths between long-range nodes are not small featured by small-world network [10]. Another important observation is that the pattern of these macro level indicators may change as the spatial and temporal scale of study changes. For example, the distribution of clustering coefficient of internet networks follow a scaling law of $C(k) \sim k^{-1}$ (k: node degree) when studied at the domain level, yet it turns into constant at router level [11]. These observations show the difficulties of connecting the system level studies that focus on macro scale emergence and the engineering solutions at the micro scale. Overall, the complex system studies to date are focused on observing emergent behavior and patterns at long term and large scale and looking for possible explanations of them. The short-term details of engineering are easily overlooked. As a result, deviations from real networks were observed. On the other hand, the engineering models focused on short term (e.g. in planning period of 20 to 30 years) therefore fail to reveal the longer-term emergence.

Using water infrastructure as an example, a dynamic model is proposed that bridges system studies with engineering practices by simulating the long term dynamic growth of infrastructure with staged optimization in each planning period, and incorporates engineering rules into the network growth. As one of the critical infrastructures for human health and economic wealth, water

distribution networks (WDNs) are sparse planar networks that are heavily limited by spatial constraints [9]. In this work, a complex network model was constructed to simulate the growth of water infrastructure guided by network efficiency optimization objectives and engineering criteria of loop formation. Then this model is used to study practical WDN design problems by combining a pipe optimization algorithm.

1.2 Scope of research project

The objective of this thesis is to provide a framework to model WDNs dynamics through bridging complex network studies with engineering practices. Through the thesis the following questions would be addressed, based on the proposed dynamic WDN modeling approach:

1. Is it possible to use simple rules as those proposed in the thesis to explain the formation of the global structure of WDNs?
2. What is the effect of urban growth pattern and engineering criteria on the efficiency and reliability of WDNs?
3. How to apply the proposed model to guide WDNs design?

1.3 Structure of this dissertation

This work contains three chapters. In chapter 2, I will introduce complex network theory, focusing on existed dynamic models on spatial network based on local optimization which provides me the tool to simulate the growth of WDNs. Then, a review of previous research on simulating WDNs is presented. In chapter 3, first, I will summarize the structural properties of real-world WDNs. Second, I will introduce the dynamic model to simulate the growth pattern of WDNs. At last, characteristics of simulated WDNs will be investigated. In chapter 4, I will present

two applications of the developed model, followed by the introduction of the pipe optimization algorithm used for model applications.

CHAPTER 2: LITUREATURE REVIEW

2.1 Complex network

2.1.1 Background of complex network

Networks are almost everywhere in our life. There are tangible networks, such as transportation networks and river networks; and there are also intangible networks, such as social networks. Networks have been long studied in many fields, including geography, physics, engineering, computer science, biology and sociology.

Historically, graph theory is used to study topological properties of network by modelling it as a collection of nodes and edges which connect them. It begins with publication by Leonhard Euler on solution to the Konigsberg bridge problem in 1736. After that, graph theory has been widely used to study special networks.

In real world, many networks are complex networks whose structure is complex and changing over time. Its study can date back in 1959, when random graph theory is introduced [12]. Random graph theory is used to study families of networks other than specific ones by combine graph theory with probability theory. With the advancement of computer and availability of large database for many real-world networks, it enables researchers to empirically study the structural properties of complex networks. However, the Erdős–Rényi model from random graph theory is unable to explain some topological properties of many real-world complex networks which are not completely regular and completely random. New models have been developed. The most well-known ones are the Watts–Strogatz model [6] that produces small-world networks and the Barabási–Albert model [7] that generates scale-free networks. The Barabási–Albert model is an

especially dynamic model that can reproduce the growing process of scale-free networks. A variety of models were developed based on the Barabási–Albert model to model dynamic complex networks with other structural properties.

While most studies on complex networks focus on their topological properties, the spatial aspect cannot be neglected, which significantly affects their structural properties. For modelling spatial complex networks, distance is incorporated into models, including the Erdős–Rényi model, the Watts–Strogatz model and the Barabási–Albert model. Also, spatial variants of the Barabási–Albert model were developed to model the growth of spatial complex networks. Another class of dynamic models of spatial complex networks is based on local optimization. For infrastructure networks with strong geographic constraints, their growth is more appropriate to be described by local optimization based models than Barabási–Albert models, since their behavior is significantly affected by engineering optimization and they usually are not scale-free.

2.1.2 Dynamic models for spatial networks based on local optimization

Based on the assumption that every new edge is generated to efficiently connect the new node to the existing network, a few simple dynamic models are developed to study the growth of some spatial complex networks. Even though those models are based on the mechanism of local optimization, some global structural properties can be reproduced.

Fabrikant et al. [13] introduced a heuristically optimized trade-off (HOT) model for internet growth by considering two objectives -- connection cost and transmission delay. Specifically, a new node *i* representing a router or autonomous system that is randomly added in the network is connected to a previous node j according to a local objective function:

$$min_{j<i}(\alpha d_{ij} + h_j) \tag{2.1}$$

where d_{ij} is the Euclidean distance from node *i* to node *j*, h_j is a measure of centrality of node *j* and α is the weight for the d_{ij} and a function of the final number of nodes (*n*). d_{ij} and h_j respectively represent the connection cost and transmission delay, and α gauges the relative importance of the connection cost (Euclidean distance). It was found the power law degree distribution of the internet can be produced when α is not too small or too large. If α is small, a star structure is produced. Or if α is as large as $\sqrt{n}$, this model will generate a dynamic version of the Euclidean minimum spanning tree (MST).

Gastner and Newman [8] proposed two similar models for spatial distribution network growth. In those two models, the positions of all nodes are given at the outset, rather than randomly added to the network one by one over time. The new edge is generated by minimizing the weighted sum of two objectives: connection costs and distribution efficiency. Here, the efficiency is measured based on route factor (*q*):

$$q = \frac{1}{n}\sum_{i=1}^{n}\frac{l_{i0}}{d_{i0}} \tag{2.2}$$

where l_{i0} is the Euclidean distance along the shortest path of the network from node *i* to root 0, and d_{i0} is the direct Euclidean distance. For node i, $\frac{l_{i0}}{d_{i0}}$ measures the straightness of its path to root. For the first model, the objective function is as following:

$$min_{j<i}(d_{ij} + \alpha\frac{d_{ij} + l_{j0}}{d_{i0}}) \tag{2.3}$$

The results show that as α increases, route factor initially decrease sharply, while the average edge length increases only slowly. It indicates that a little more construction costs can achieve much more distribution efficiency. For the second model, it considers the short path instead of the straightness. The resulted objective function is:

$$min_{j<i}(d_{ij} + \alpha l_{j0}) \tag{2.4}$$

The generated network self-organizes to a state with a small route factor. And the relationship between α, route factor and average edges length is qualitatively similar to the first model. The difference lies in their shapes. The network generated by the first model has a dendritic structure with some node unconnected to the network, due to the preference of straight path. However, the network generated by the second model fills out the space with some approximately constant from the root, resulting from the optimization of short path.

Those two models for distribution networks only generate networks with tree structures. But many distribution networks have loops to remain functional in case of perturbation or disturbance. Barthélemy and Flammini [14] developed a more realistic growth model for street network with loops, which is also based on local optimization. At each time step, several nodes (centers) are added randomly in the network according to a given distribution. Then edges (roads) will be generated to connect the new nodes to all previous nodes in their relative neighborhood, which results in the formation of loops. If two new nodes (A, B) have a same previous node (M) in their relative neighborhood, the roads will be built to maximize the reduction of the cumulative distance from M to A and B:

$$\Delta = [d(M,A) + d(M,B)] - [d(M',A) + d(M',B)] \tag{2.5}$$

$$\text{Constraint: } |MM'| = const. \ll 1 \tag{2.6}$$

where M' is interaction node between A and B, and A simple calculation for M' is:

$$\overrightarrow{MM'} \propto \vec{u}_A + \vec{u}_B \tag{2.7}$$

where $\vec{u}_A$ and $\vec{u}_B$ are the unit vectors from *M* in the direction of *A* and *B* respectively. The structural properties of generated street networks by this model match those found in empirical studies. It indicates that local optimization is a possible mechanism for the growth of street networks. Based

on this simple model, Barthélemy and Flammini [15] studied the co-evolution of population density and topology of street networks. Two economical mechanisms, accessibility and house price, are taken into account for the choice of new location.

Based on local cost benefit analysis, Louf et al. [16] used a generic model for the growth of spatial networks. At each time step, a new node *i* is connected to a previous node *j* to maximize the net benefit of the new edge:

$$R_{ij} = B_{ij} - C_{ij} \tag{2.8}$$

where B_{ij} is the expected benefit associated with construction of the edge between node *i* and node *j*, and C_{ij} is the expected cost associated with such a construction. The transportation networks were taken as a case example. The results show that, as the cost becomes relative more important, there is a crossover between a star graph and MST. When cost and benefit are of the same order of magnitude, a reminiscent of hub-and-spoke structure is emergent, which has a hierarchical structure. It is found that this structure is most efficient due to the high diversity of edge length.

2.2 Previous works

Currently, several static models of WDNs have been introduced in the purpose of generating synthetic WDNs for case study analysis. Only one dynamic model was developed based on a static model.

2.2.1 Static model

In water supply management, case study data is crucial to build theory and test new technology and strategy. However, the availability of WDNs data is limited due to the consideration of terrorism and security. Thus, currently, the study of WDNs is based on a few benchmark cases and the results can hardly be generalized. In order to complement the lack of

real-world study cases, research has been done to generate synthetic WDNs. Currently, there are four simulation models: “Modular Design System” (MDS) [17-20], “WaterNetGen” [21], “HydroGen”, spanning tree based algorithm [22].

MDS was first developed by Möderl et al. [17] based on MATLAB to create a variety of virtual WDNs with different boundary conditions and properties based on graph theory. Each WDN is formed by combining various small basic blocks consisting of some connected junctions, presenting different sections of the network with different properties (layout, junction density). The generated WDNs have grid structures because all blocks are built in grid. In order to generate more realistic WDNs, Sitzenfrei et al. [20] developed “Graph-based Catenation Approach” (GCA) to concatenate blocks according to the boundary conditions (housing and population densities) and follow a layout strategy. Further, Sitzenfrei et al. presented “Virtual Infrastructure Benchingmarking” (VIBe) software [23] to generate urban structure (digital elevation model, river system, land use, population density) based on multi-layer cellular automata, including infrastructure (sewer systems and WDNs). The virtual WDNs are created based on MDS and GCA according to the virtual environment [19]. As a result, VIBe exports the sewer system models and WDN models in the format used by SWMM and EPANET respectively for simulation. Through this software, WDNs can be stochastically analyzed and its interaction with urban structure can be also investigated.

WaterNetGen is developed by Muranho et al. [21] as an EPANET extension to automatically generate virtual WDN model. Different from WDS which generates virtue WDNs by combining blocks of predefined connectivity, WaterNetGen mimic real world WDNs which can be organized in district metered areas (DMAs): First, clusters of junctions are generated; then,

junctions in each cluster are connected by pipes according to distance and node degree; finally, clusters are interconnected based on distance. The generated WDNs have irregular geometry.

Similar to WaterNetGen, Hydrogen simulates WDNs based on DMAs: clusters of junctions are firstly generated in a circular area, then inter-cluster and intra-cluster nodes are connected by pipes according to different rules. One of the advantages of Hydrogen is that reservoirs, pumps and valves are generated automatically. The other advantage is allowance of predefining different redundancy for different clusters measured by meshedness coefficient.

In reality, the layout of WDNs is constrained by street networks. Mair et al. [22] developed an algorithm based on spanning tree to generate WDNs by using data of street networks. The initial WDNs are sub-graph of street networks generated by the spanning tree algorithm. The final networks are looped based on the length of the alternative path.

2.2.2 Dynamic model

Based on VIBe [23], Sitzenfrei et al. [24] developed dynamic virtual infrastructure benchmarking (DynaVIBe) to simulate the dynamic adaptation of water infrastructure to urban environment. For the study of WDNs, DynaVIBe has been used to investigate the effects of different design strategies for WDNs expansion on hydraulic performance, water quality and construction costs [25]. Specifically, the dynamic WDN model in DynaVIBe is based on the static model of MDS which generates WDNs by combining blocks which have grid structures. Pipe sizing is done assuming that the future water demand is known.

Based on graph theory, those models generate synthetic WDNs for case study by simply manipulating nodes and edges. Those synthetic ones resemble the real-world WDNs on some structural properties. But none of those models describes the mechanism of the evolution of the

real-world WDNs. Thus, they are incapable to explain the structural properties and predict the complex behavior of WDNs.

CHAPTER 3: MODELING OF WDN GROWTH PATTERNS

3.1 Structure of water distribution networks

As one of the most important man-made systems, water distribution systems are developed by engineers to deliver water from source to customers. A water distribution system is a complex system consisting of many interacted components. To analyze water distribution systems, they are transformed to networks with nodes (reservoirs, tanks, pipe junctions) and links (pipes, pumps, valves). The configuration of a pipeline network can be either branched or looped (Figure 3.1). A branched network has a tree structure, thus there is only one unique path to deliver water from the source to a demand node (i.e. consumers). The advantage of branched networks is that water is efficiently delivered to customers with low capital costs. However, such networks are very unreliable. If one pipe fails, downstream consumers would be disconnected to the source. On the contrary, a looped network is more reliable due to the existence of redundant pipes. For a fully looped network, water can be delivered to each customer through multiple paths, thus all customer can have their water supply regardless of any single-pipe break condition. The other advantages of looped networks include higher supply capacity and reduced water age [26]. Such network requires more initial investment. In practice, designers need to tradeoff between cost saving and network reliability. Thus, real-world WDNs are mostly partially looped [27]. Branched structures are usually used for low-density rural areas while looped structures are used for urban areas [28].

Figure 3.1. Configuration of a branched network (left) and a looped network (right).

The performance of WDNs heavily depends on their structure. On the traditional research field of WDN optimization, efforts have been made to optimize the layout of WDNs to achieve certain system reliability. Layout related reliability is called topological reliability, which is focused on network connectivity and pipe redundancy. A few metrics have been introduced two decades ago to measure topological reliability, including reachability/connectivity [29], node pair reliability [30] and system redundancy [31]. Among them, system redundancy is used by most researchers when a layout optimization problem is concerned [32-35]. It measures the number of distinct paths with no shared pipes between demand nodes and sources.

As complex systems, control of WDNs is beyond the scope of tradition engineering. For new engineering, it is needed to analyze network structure and complex behaviors, in order to understand their relationships. Due to the consideration of terrorism and security, the data of real-world WDNs is often not available for study. Currently, only a few papers have been published to explore the structural properties of real-world WDNs by the tools of complexity theory. Even though the number of WDNs sample studied in those research is very limited, it reveals valuable information of network structure properties shared by most of real-world WDNs. In this section, those results are summarized.

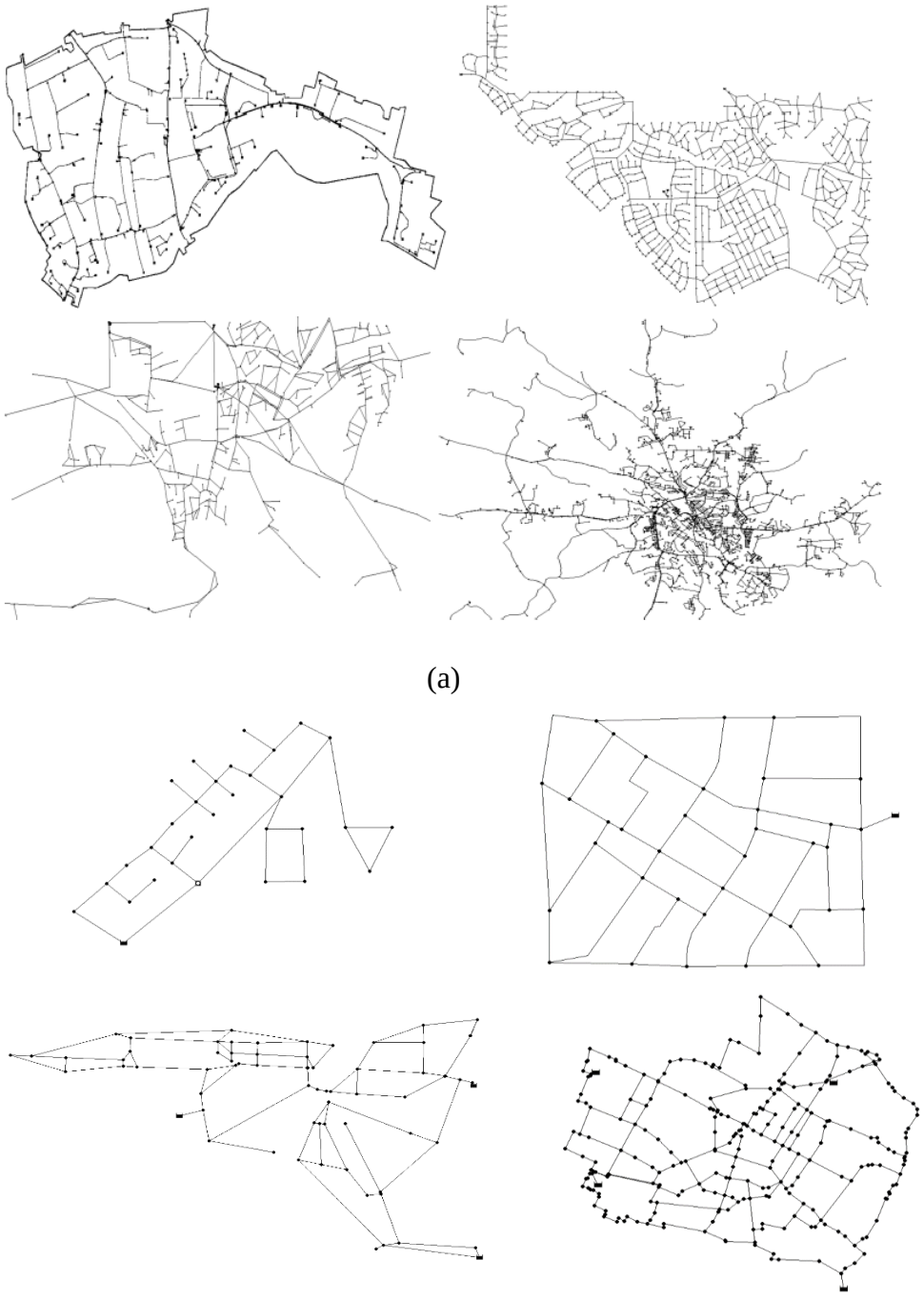
3.1.1 Complex network analysis

The first paper systematically studying the structure of WDNs was done by Yazdani and Jeffrey [9]. They quantitatively characterized the structure properties of four real-world WDNs by complex network metrics. Other researchers used the same metrics in their works to measure WDN structure [36, 37]. In complex network theory, a network is modeled as a graph G = G (N, E) where N is the set of nodes and E is the set of edges connecting the nodes. WDNs can be approximately modeled as planar graphs. A planar graph is a graph can be embedded in the plane with edges only intersecting at nodes. WDNs are usually planar because it is impractical to lay pipes in multiple layers underground [9]. Furthermore, WDNs are spatial networks which are heavily constrained by geographical and economic considerations. The formation of the structure of WDNs is significantly affected by those properties. Table 3.1 listed the structure properties of eight real-world WDNs reported by literatures, including East-Mersea (UK), Colorado Springs (USA), Richmond (UK) and Kumasi (Ghana) reported by Yazidani and Jeffery [9] (Figure 3.2(a)), Balcksburg (USA), Fossolo (Italy), Pescara (Italy) and Modena (Italy) reported by De Corte and Sorensen [37] (Figure 3.2(b)). Metrics used by those research are described below:

(1) Edge density (e): the ratio of the number of total existing edges to the maximum possible edges. This metric indicates the sparseness/denseness of the network connections.

(2) Average node degree ($<k>$): node degree is the number of edges connected to a node. In addition to average node degree, node degree distribution describing the heterogeneity of network connectivity is also analyzed in this work.

(3) Edge-per-node ratio (m/n): the ratio of number of edges to number of nodes. Similar to average node degree, it measures the level of network connectivity.



(a)

(b)

Figure 3.2. Real-world water distribution networks reported by Yazidani and Jeffery [9] (a): East-Mersea (top-left), Colorado Springs(top-right), Richmond (bottom-left) and Kumasi (bottom-right). Real-world water distribution networks reported by De Corte and Sorensen [37] (b): Blacksburg (top-left), Fossolo (top-right), Pescara (bottom-left) and Modena (bottom-right).

Table 3.1. Structure properties of real-world WDNs (n, nodes; m, edges; e, edge density; m/n, edge-per-node ratio; $<k>$, average node degree; k_{max}, maximum node degree; g_{max}, diameter; l, characteristic path length; c_b, central-point-dominance; q, route factor; r_m, meshedness; λ_2, algebraic connectivity; $\Delta\lambda$, spectral gap).

	n	m	e	m/n	$<k>$	k_{max}	g_{max}	l	c_b	q	r_m	λ_2	$\Delta\lambda$
East-Mersea (UK)[a]	755	769	2.70×10^{-3}	1.01	2.04	4	97	34.48	0.36	1.54	9.97×10^{-3}	1.97×10^{-4}	3.91×10^{-2}
Colorado Springs (USA)[a]	1786	1994	1.25×10^{-3}	1.11	2.23	4	69	25.94	0.42	1.45	5.86×10^{-2}	2.43×10^{-4}	2.83×10^{-2}
Kumasi (Ghana)[a]	2799	3065	7.83×10^{-4}	1.10	2.19	4	120	33.89	0.45	1.46	4.77×10^{-2}	9.40×10^{-5}	9.08×10^{-3}
Richmond (USA)[a]	872	957	2.52×10^{-3}	1.09	2.19	4	135	51.44	0.56	1.67	4.95×10^{-2}	6.09×10^{-5}	7.27×10^{-2}
Blacksburg (USA)[b]	31	35		1.13	2.26	4	9	4			0.09		
Fossolo (Italy)[b]	37	58		1.57	3.14	4	8	4			0.32		
Pescara (Italy)[b]	71	99		1.39	2.79	5	20	9			0.21		
Modena (Italy)[b]	272	317		1.17	2.33	5	38	14			0.09		

[a]Data from Yazdani and Jeffrey's works [9].
[b]Data from De Corte and Sörensen's work [37].

(4) Diameter (g_{max}): the maximum geodesic distance between all node pairs in the network. Geodesic distance is the number of edges in a shortest path.

(5) Characteristic path length (l): average geodesic distance between all node pairs.

(6) Central-point-dominance (c_b): average difference in relative betweenness centrality of the most central point and all other nodes. Betweenness centrality of a node is the sum of the ratio of the number of shortest paths between any two nodes that pass through the node to the total number of shortest paths between those two nodes. The relative betweenness centrality of a node is dividing betweenness centrality by the total number of node pairs. Betweenness centrality measures the importance of a node based as its importance to connect other nodes. Central-point-dominance indicates how the whole network is controlled by the most central point.

(7) Route factor (q): average ratio of Euclidean distance along the shortest path from a node to the root to its direct Euclidean distance to the root. This ratio measures the straightness of its path to root, which can be used to measure the efficiency of a distribution network.

(8) Meshedness coefficient (r_m): the ratio of the number of total existing independent loops to the maximum possible loops. Comparing the indicators such as clustering coefficient, meshedness coefficient is a better indicator for the topological redundancy of a WDN because that the loops in WDNs are mostly quadrilateral not triangular[9].

(9) Algebraic connectivity (λ_2): the second smallest eigenvalue of the normalized Laplacian matrix of the network. A larger algebraic connectivity means the network is more difficult to be bisected. Thus, this metric is usually used as the measurement for the network robustness.

(10) Spectral gap ($\Delta\lambda$): The difference between the first and second eigenvalues of the adjacent matrix of the network. Small spectral gap may indicate the presence of bottlenecks whose removal results in the split of the network into at least two large connected components. Similar to algebraic connectivity, this metrics can be used to measure network robustness.

As shown in Table 3.1, all WDNs have low values of edge density, which indicates that WDNs are sparse networks. The values of average node degree for those real networks are basically larger than 2 (value for tree) but much smaller than 4, and the values of edge-per-node ratio are larger than 1 but much smaller than 2. The value of average node degree and edge-per-node ratio for a tree-structured network have an average node degree of 2 and an edge-per-node ratio of 1, while a looped network with grid pattern network have an average node degree of 4 and an edge-per-node ratio of 2. The results indicate that real WDNs are usually less connected than a grid network. Those networks have a minimum node degree of 1 and a maximum node degree of 5, and most nodes have degrees less than 4. Thus, they are deviated from scale-free networks [7] where a power law or near power law exists for the node degree distribution. Furthermore, in Yazdani and Jeffrey's work [9], they found that the geodesic distances between many node pairs of the four large real WDNs are large, and the diameter and characteristic path length are also large. Thus, WDNs are not small-world networks [6] featuring short geodesic distances between nodes, meaning that it takes only a few stops to go from one node to another. That is due to the fact that water distribution pipes connect mainly according to spatial proximity. The value of central-point-dominance of those networks are smaller than 1, which is the value for a star-structured network, and larger than 0, which implies a network with equal important nodes in terms of shortest path. Thus, there is no huge difference on node importance for WDNs. The values of rout factors are

slightly larger than 1 which is the value for a network with nodes directly connected to each other. It implies that WDNs are efficient. Due to graphic limitation, WDNs is far less than maximum looped, indicated by low meshedness coefficient. In general, complex network metrics provide a way to compare the structure properties for different networks, including network efficiency, redundancy and robustness.

3.1.2 Fractal analysis

Although classical geometry (Euclidean geometry) can be used to study spatial objects with regular form (e.g., line, circle), many real-world objects have irregular forms which cannot be easily represented by classical geometry [38]. The irregular forms, although defying any geometrical order, may reveal the same degree of disorder and irregularity on many scales [39]. And fractal geometry could provide researchers a general framework to study such forms. Fractals are defined as tenuous spatial objects whose geometric properties include irregularity, scale dependence, and self-similarity [40]. The property of self-similarity means that the object can be "split into parts, each of which is (at least approximately) a reduced-size copy of the whole" [41], which has long been recognized as the characteristics of various natural phenomena [42]. From the past three-decade's study, it is acknowledged that, not only natural objects can be fractal, artificially planned and designed objects can also be fractal, such as urban form [39, 40, 42-48] and transportation networks [49-55].

The main tool of fractal analysis is the measure of fractal dimension, which links the fractal geometry to the traditional concepts of Euclidean geometry [53]. In Euclidean geometry, a straight line and a square have a dimension of 1 and 2, respectively. However, fractal dimension can be used to describe the dimension of curves. In fractal geometry, a curve should have a fractal dimension between 1 and 2, which is the feature of many real-world planar complex systems. It

characterizes the scaling properties of spatial objects [56], which gives an indication of how completely a fractal fills the space as scale changes [57]. Among various calculation methods, box-counting method is one of easiest to operate. To calculate fractal dimension, boxes with side length *l* are used to cover the spatial object, and then count the minimum number of boxes needed, *M*. The number of boxes will change with the side length of the box (Figure 3.3). Thus, the fractal dimension (*D*) of the node set is:

$$D = \lim_{l \to 0} \frac{\log M}{\log(1/l)} \tag{3.1}$$

Being part of the whole urban system, water distribution networks (WDNs) evolve to adapt the water demand of the population in a fractal-sprawled urban form, which makes the WDNs inheriting the fractal structure. The fractal growth of water distribution systems was firstly discussed by Chai and Li [58]. Recently, Kowalski et al. [59] analyzed a real WDN, and found it has the structure property of self-similarity and the fractal dimension calculated by box-counting method is 1.439.

Figure 3.3. Estimate the fractal dimension of the coastline of Great Britain by box counting. (image from https://en.wikipedia.org/wiki/Minkowski%E2%80%93Bouligand_dimension)

3.2 Model development

Mathematically, a spatial network can be modeled as a graph G = G (N, E) where N is the set of nodes and E is the set of edges connecting the nodes. In a WDN, the nodes could be reservoirs, tanks, pipe junctions. For the sake of simplicity, one root node (reservoir) is arbitrarily chosen in the beginning of the modeling. In real systems, more root nodes are possible. The complexity of WDNs comes from multiple sources. As a service providing infrastructure, WDN continuously expand over time to meet the growing demand of population in a self-evolving manner. In addition, the growth and maintenance of the networks depend crucially on the social decision making at a variety of temporal and spatial scales where the water distribution components interact with the social-economic environment. The current engineering optimization is in general multi-objective including efficiency, cost, and reliability [60]. Each object requires a unique set of optimization rules. The overall optimization outcome is then a compromise of all rules. Moreover, although designed and locally optimized with rigid engineering methods, a solution to a local problem is often implemented without the consideration of the whole system. The wide variety of options adds more complexity into the system. Therefore, modeling the complexity behavior of WDN needs to include the dynamic growth of water demand (nodes) and the evolving of the pipeline networks (edges) with the consideration of engineering optimization rules as depicted below. This will allow the exploration of emergent properties of short term and local scale engineering decisions, which provides important insights into bridging complexity and engineering.

There are two sub-modules in the proposed complex network model for simulating WDN dynamics: the first module simulates the generation of new water demand according to urban development; the second module simulates the expansion of WDN due to the new water demand. Specifically, at each time step, multiple new nodes (i.e. new water demand) would be generated

for the given area, after that each new node would be connected to the existing network through one pipe forming a tree structure and, in order to improve reliability of the network, extra edges will be generated to form loops. Schematic flow chart is shown in Figure 3.4.

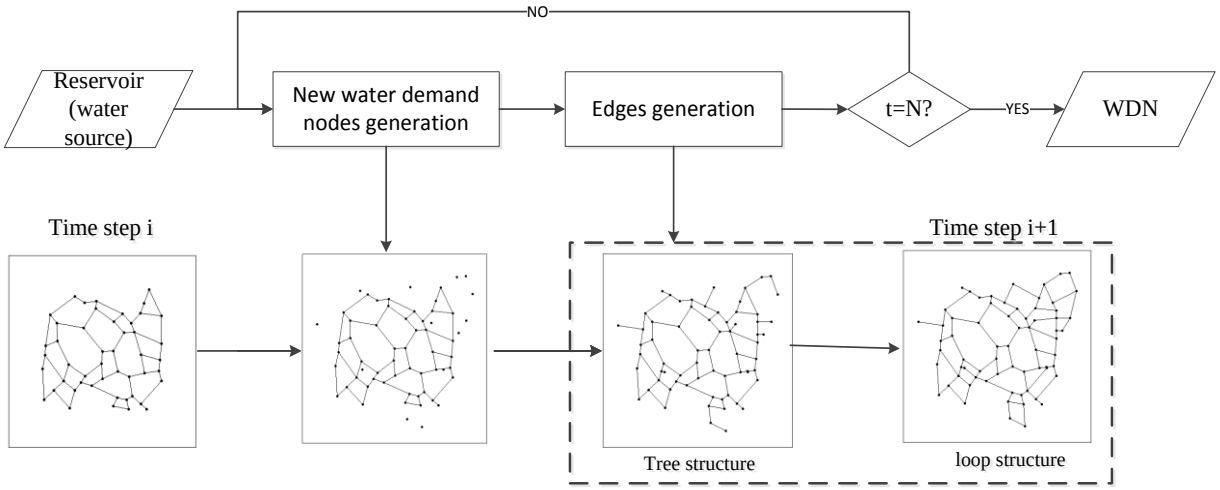


Figure 3.4. Schematic flow chart.

3.2.1 Water demand nodes generation

One challenge to model real engineered complex systems is the generation of demand nodes, the pattern of which is complex by nature. The topology of the water demand nodes follows the growth and distribution of population in the city. Urban growth has been a long time endeavor that inspired a wide range of top-down and bottom-up models. Although planning practices are in favor of detail oriented models such as UrbanSim [61] or SLEUTH [62] that are capable of locating demands in fine grids of a specific area, these models are burdened with great data and calculation needs that are not necessary for the exploration of system behavior under changing guiding rules. Empirical studies have shown that the growth of city can be described as power laws. A simple scaling function of the population (P) over the urban area (A) can be used to describe the distribution of any urban forms [63, 64]:

$$A \propto P^{\alpha} \tag{3.2}$$

where α is the scaling factor characterizing how urban area grows with population. In this study, the water demand nodes generation is governed using this scaling rule. The coefficient α varies as the location and time changes, reflecting different urban forms that are limited by the spatial and economic factors [65-68]. Marshall [68] studied the development of US urban areas during 1950-2000, and found the central tendency value of α is towards 2 and interquartile values are in the range of 1.3~3.7.

To simulate the realistic growth, the demand nodes are generated dynamically in each planning period. Starting with only one root node representing the reservoir located at the center of the given area, at each time step, the nodes generation module would add multiple water demand nodes at random which follows the scaling rule of equation (3.2). Assuming water demand of each node is identical, the number of nodes is linearly proportional to the population. Therefore, with a uniform distribution, new nodes are added within a square region with side length $2R$, where R is calculated by:

$$R \propto N^{\alpha/2} \tag{3.3}$$

where N is the total number of water demand nodes. The side length increases with N, simulating the sprawling of urban area. The number of initial water demand nodes is N_0 and they are located in the square region with side length $2R_0$. The number of new water demand nodes generated at each time step is determined according to population growth rate. According to U.S. census bureau, the average yearly population growth rate in United States was 1.3% in the 20th century. As the design of WDN expansion is done with a typical planning horizon of about 20 years [69], the growth rate of new water demand nodes is set to be 30% for each time frame in this study. An engineering constraint is imposed as the distance threshold r_0 to prevent generating unrealistic demand node due to in-fill development. If a new node is within the distance r_0 to an existing node in the network, the new node would be removed and existing node is taken as a connected new

water demand node. If two new nodes are generated within distance r_0, they will be combined into one.

In general, the most important parameter for water demand nodes generation is the scaling factor α, which reflects urban growth pattern. As shown in Figure 3.5, the larger the value of α, the more sprawled the urban form. For a sprawled city, nodes are generally much denser at the central urban areas than suburban areas.

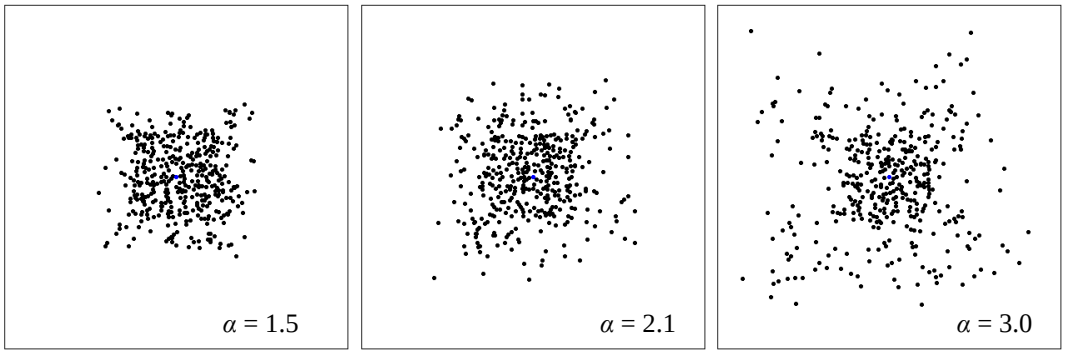


Figure 3.5. Water demand nodes generated based on different value of α.

3.2.2 Edges generation

WDNs are networks embedded in two-dimensional space with one or more root nodes. They grow outward from the root as their service area expands. As the demand grows, pipes were added into the system to connect the new demand nodes to existing networks. To optimize the topology of a WDN, both the total length of pipes and shortest path distances from the root to demand nodes need to be considered when connecting a new node to the system. The total length of pipes represents the total capital cost. The shortest path distance provides the shortest travel distance of water from the reservoir to a consumer, which represents operational energy demand and water age. Without consideration of reliability, minimization of the total length of pipes can be achieved by a minimum spanning tree (MST) [70]. A star structure that connects each node directly to the root would minimize the total travel distance from all demand nodes to the root. In

practice, a tradeoff between the total length of pipes and the distance from water source to consumers' needs to be considered. In their spatial distribution network model with static demand nodes, Gastner and Newman [8] proposed a growth model with local minimization process considering both total length of edges and shortest path distances from root node to other nodes. In this study, the same local minimization mechanism was adopted and modified to reveal the real situation of water network development. An unconnected new water demand node *i* is connected to an edge in the existing network according to the objective function:

$$min(d_{ij} + \beta d_{j0}) \tag{3.4}$$

where $d_{i,j}$ is the Euclidean distance from a water demand node *i* to node *j* which is the closest point on the potential edge to node *i*, d_{j0} is the Euclidean distance along the shortest path of the network from root node (reservoir) to node *j*, and β regularizes the balance between $d_{i,j}$ and d_{j0}. A large value of β gives higher weight to shortest path distance to the root when calculating the objective function. When β is set to 0, the generated network will be a MST, while a β value of 1 will result in a star structure (Figure 3.6).

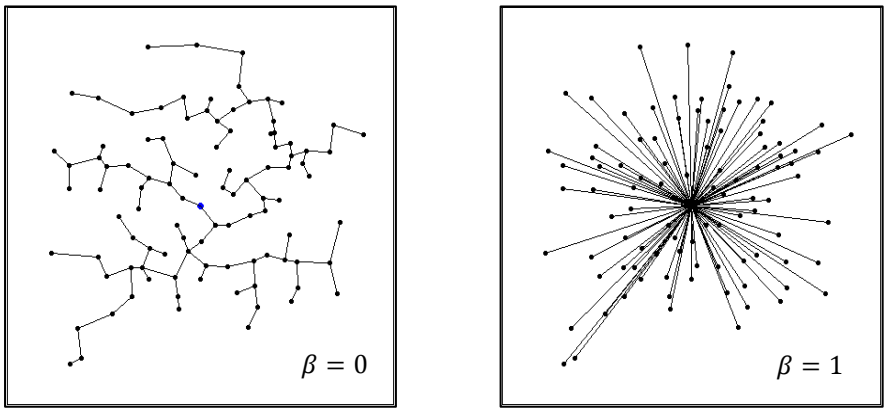


Figure 3.6. A minimum spanning tree generated at β = 0 and a star structure generated at β = 1.

The modification on the Gastner and Newman model is that node *j* is not necessarily an end node of an edge. Rather, the node *j* for water network model can be any point on an existing edge that is close to the unconnected node (Figure 3.7). In case a new node is close to an edge but

far (larger than distance r_0) from both end nodes of the edge, as illustrated in the Figure 3.7(b), a connection node is introduced on the edge at the perpendicular crossing point. The connection node has no water demand, while that edge would be split into two edges connected by the connection node. Through an exhaustive search, the model algorithm would identify the optimized edge in the existing network that a new water demand node should be connected to according to the objective function.

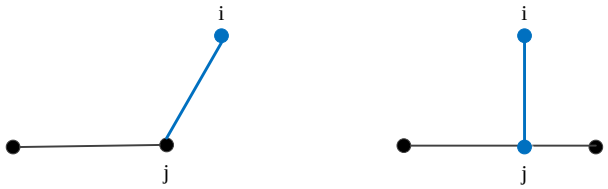


Figure 3.7. A new water demand node *i* is connected to an existing edge: (a) through an end node *j*, (b) through a connection node *j*.

The aforementioned edge generation will lead to a tree structure as most existing complex network studies did. In practice, the water network contains loops that ensure the reliability of the services in case of perturbation or disturbance. This is handled through another optimization mechanism. Specifically, for the new water demand node *i*, an extra edge can be added to connect it to another edge through node *k* (end node or connection node) by minimizing the distance:

$$min(d_{ik}) \tag{3.5}$$

The addition of edge to form loops will increase the cost of network therefore need to be constrained by the following network graph and engineering rules:

(1) $d_{ik}<d_{max}$: node *k* has to be within the neighborhood of new node *i*, which is limited by a maximum distance d_{max}. Actual water networks are mostly partially looped due to economic or geographical limitations. The parameter d_{max} can be adjusted to simulate real

water networks with different level of looping. The larger the value of d_{max}, the more possibilities of forming loops (Figure 3.8).

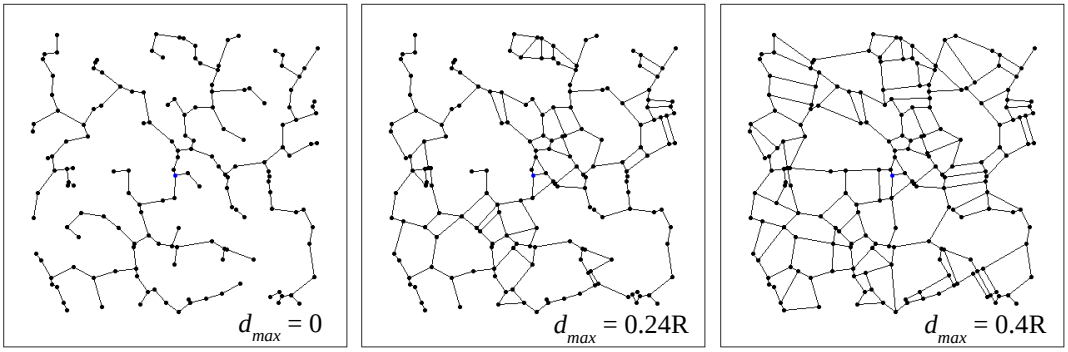


Figure 3.8. WDNs generated based on different values of d_{max}.

(2) k_i<3: degree of new node i shall be lower than 3. Most nodes in a tree-structured network have a degree of 2 while for a looped network with grid pattern, node degree is mostly 4. A real WDN is usually less connected than a grid network [37]. Thus, if the degree is 3 or higher, it commonly indicates that new node i has already been well connected and no extra edge would be needed.

(3) g_{ik}>2: g_{ik} is the geodesic distance between node k and new node i. This constrain assures that new node i is not directly connected to one edge repeatedly which results in a triangular cycle, as shown in Figure 3.9. Cycles in WDNs are mostly quadrilateral not triangular [9].

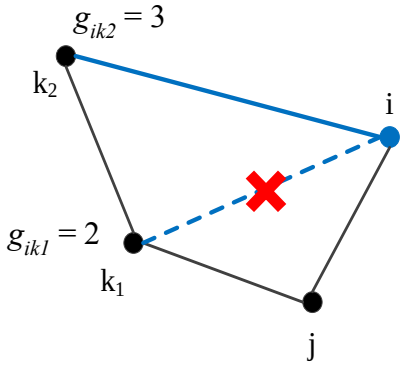


Figure 3.9. Illustration of looping constraint: g_{ik}>2. Node i can not be connected to k_1 because g_{ik1} is not larger than 2.

(4) $\theta > \theta_{min}$: θ measures the angle between the extra edge with any other incident edges and is constrained by a minimum angle θ_{min}. The rationale for this constraint is that the design of WDNs layout in practice is restricted by the street right of ways. Thus, the layout of WDNs follows certain street pattern [71]. Very small angle between two roads is not common in real world. In this study, θ_{min} is set to 45° which is also used by [72] to obtain a more realistic street pattern.

(5) The extra edge shall not be intersecting with other edges as WDNs are usually planar [9].

3.3 Results and discussion

3.3.1 An example of generated WDN

The proposed model has three adjustable parameters: 1) the scaling factor α allows the simulation of different urban forms; 2) the weighting coefficient β allows the simulation of relative importance of shortest path distance, which is implied in engineering optimization process; and 3) the maximum distance d_{max} allows the testing of the impact of engineering looping criteria on the emergent properties of WDNs. To simulate the growth of the real water network, the model runs in a time step of 20 years, which is the typical planning period for the water network master plan. At the beginning of each time step, new nodes are generated following the scaling rule of urban growth. Edges are then generated according to the local optimization and loop formation rules. Figure 3.10(a-d) shows the growth of a generated WDN at different simulation steps with the parameter setting at α = 2.1, β = 0.1, $d_{max} = 1.4R_0$. The generated WDN expands as time step progresses, consisting of increasing number of water demand nodes and more complex pipe networks connecting the customers. A zoomed-in view of the center of the final WDN layout is provided in Figure 3.10(d), showing a near-fully looped structure around the center due to high

population density. To obtain networks with different structural properties, α, β and d_{max} can be adjusted which are discussed in section VI.

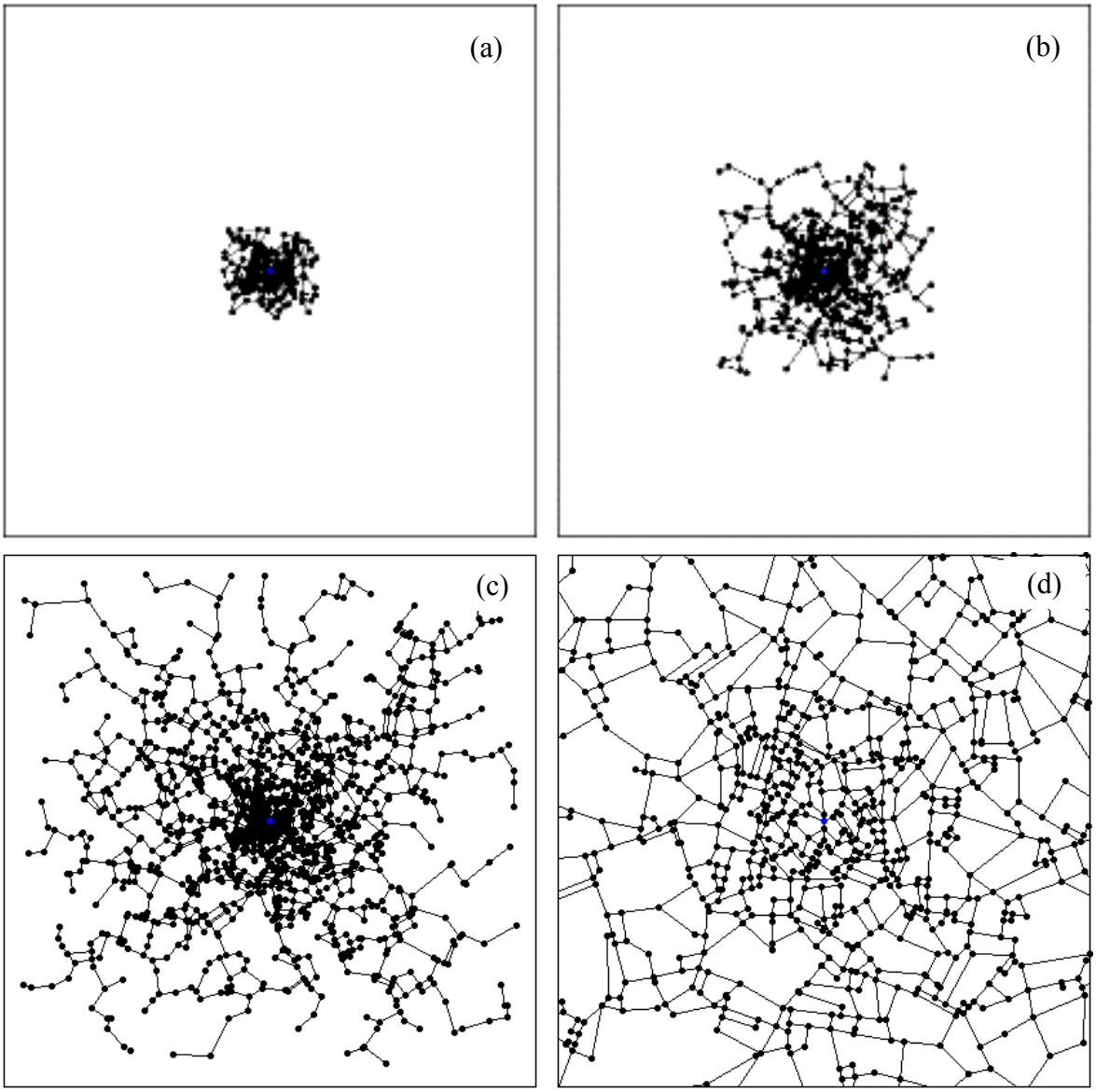


Figure 3.10. Snapshot of the generated WDN at different time steps (α = 2.1, β = 0.1, d_{max} = 1.4R_0, N_0 = 100): (a) t = 4, (b) t = 7, (c) t = 10, (d) zoomed-in center of the network of t = 10.

3.3.1.1 Fractal pattern of nodes

The node distribution of Figure 3.10(c) is analyzed by fractal theory. Its box counting dimension is calculated by Fraclac, a free software developed by Karperien. As shown in Figure

3.11, five different box sizes are used to cover network nodes. A fractal dimension of 1.5156 is obtained by linear regression (Figure 3.12). This is in agreement with the fractal nature of urban form [43, 44, 73]. Most studies on modeling infrastructure networks generate nodes randomly with a uniform distribution [8, 13, 14]. Yook et. al [74] used fractal distribution of nodes to build a network model for the internet and indicated that it was more realistic than the ones using randomly distributed nodes. They found that the fractal dimension for the internet and population of the world is 1.5±0.1. Since water demand nodes were generated based on the scaling law between the population and urban area with an average α, it is expected that the fractal dimension of generated WDN is in the range of internet and population.

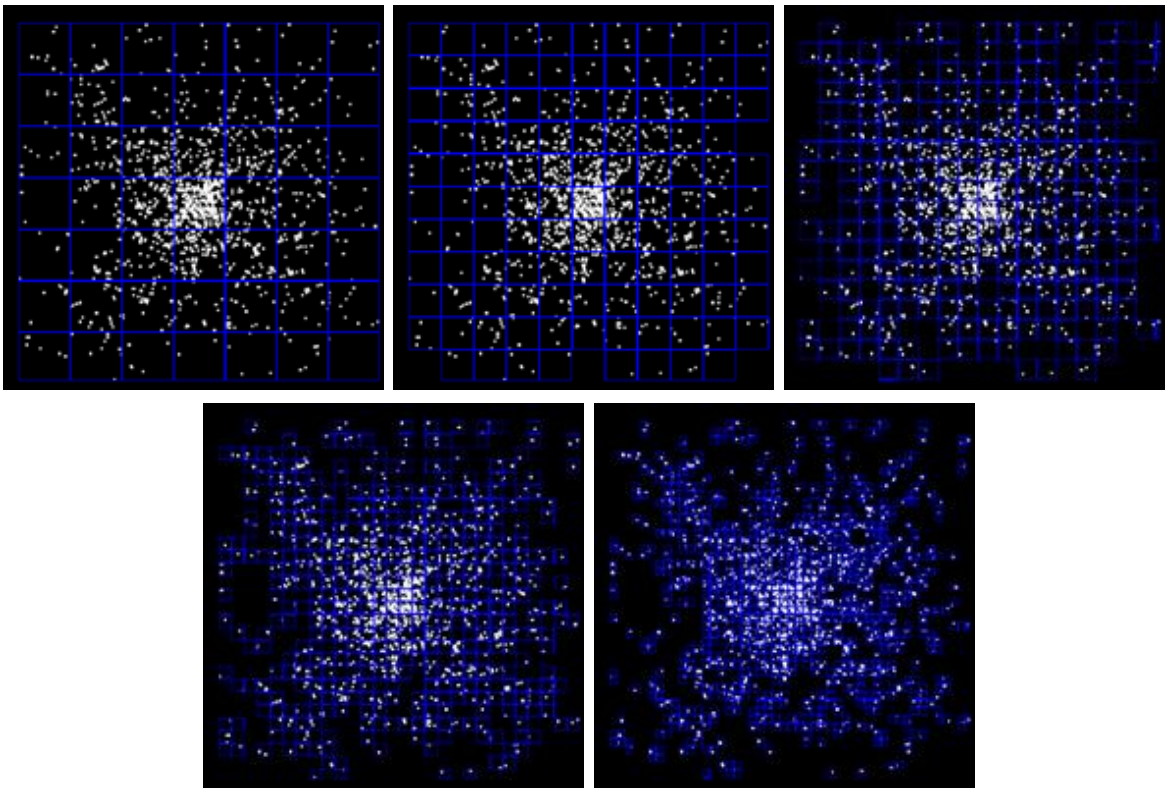


Figure 3.11. Nodes covered by boxes with 5 different box sizes.

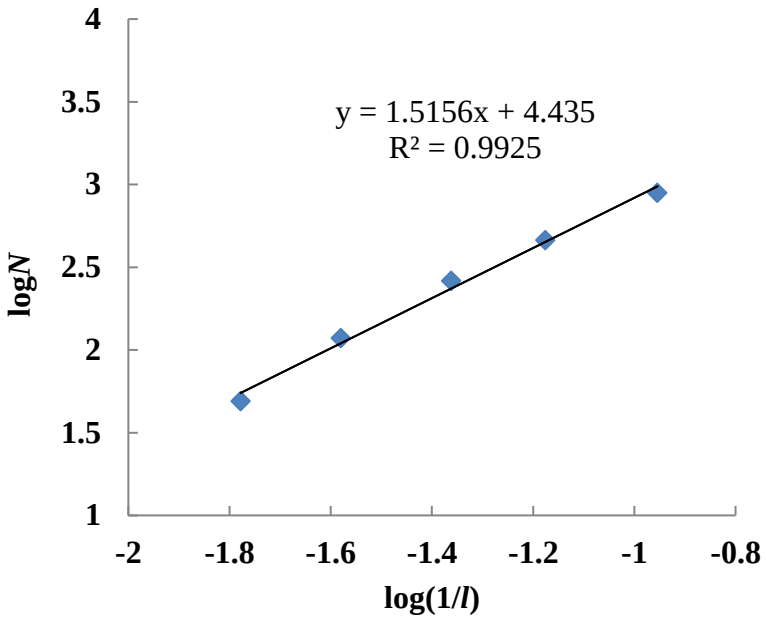


Figure 3.12. Fractal analysis of the generated WDN.

3.3.1.2 Comparison with real-world WDNs.

The structural properties of the generated WDN (Figure 3.10(c)) as well as eight real-world networks (Figure 3.2) are listed in Table 3.2, which are measured by some important complex network metrics described in section 3.1.1. Besides of those metrics, two new metrics has been proposed to characterize network structure based on scale information: median pipe length (L), which is the median value of pipe length; and median loop size (A), the median value of area enclosed by an independent loop. In general, the emergent structural properties of the generated WDN agree with the values reported for real-world WDNs.

Table 3.2. Structural properties of eight real-world WDNs and the generated WDN (n, nodes; m, edges; e, edge density; $<k>$, average node degree; k_{max}, maximum node degree; q, route factor; L, median pipe length (m); A, median loop size (km^2); r_m, meshedness; λ_2, Algebraic connectivity).

	n	m	e	$<k>$	k_{max}	q	L	A	r_m	λ_2
East-Mersea[a]	755	769	2.70×10^{-3}	2.04	4	1.54			0.01	1.97×10^{-4}

Colorado Springs[a]	1786	1994	1.25×10^{-3}	2.23	4	1.45	34	0.019	0.06	2.43×10^{-4}
Kumasi[a]	2799	3065	7.83×10^{-4}	2.19	4	1.46			0.05	9.40×10^{-5}
Richmond[a]	872	957	2.52×10^{-3}	2.19	4	1.67	40		0.05	6.09×10^{-5}
Blacksburg[b]	31	35	7.53×10^{-2}	2.26	4		129		0.09	4.19×10^{-2}
Fossolo[b]	37	58	8.71×10^{-2}	3.14	4		134	0.018	0.32	7.37×10^{-2}
Pescara[b]	71	99	3.94×10^{-2}	2.79	5		333	0.099	0.21	3.26×10^{-3}
Modena[b]	272	317	8.60×10^{-3}	2.33	5		188	0.299	0.09	4.06×10^{-3}
Generated WDN	1437	1970	1.91×10^{-3}	2.74	5	1.21	231	0.098	0.19	1.08×10^{-3}

[a]Data from Yazdani and Jeffrey's works [9].
[b]Data from De Corte and Sörensen's work [37].

As shown in Table 3.2, the generated WDN at the final stage (t = 10) has 1437 nodes and 1970 edges, which is less than Colorado Springs and Kumasi and more than other networks. The small value of edge density (e) indicates that the generated WDN is as sparse as the real-world cases. Similar to the results for real-world WDNs, average node degree ($<k>$) for the generated network is larger than 2 but much smaller than 4, which indicates that the generated WDN is much less connected than a grid network. The degree distribution (Figure 3.13) indicates this WDN is not scale-free network whose degree distribution follows a power law. The generated network has a minimum node degree of 1 and a maximum node degree of 5, and most nodes have degrees of 3 or 2. The geodesic distance distribution is also shown in Figure 3.13. It indicates that the generated network is not a small-world network characterized by short geodesic distance between nodes. Due to the fact that water distribution pipes connect mainly according to spatial proximity, a small-world model is not practical for WDNs. This is in agreement of the observation of real-world WDNs. The generated network has a smaller route factor (q) than the four real-world WDNs, which means that it is more efficient to distribute water. This is expected because the proposed simulation model has ignored certain practical constrains in WDN development such as geography and hydraulics factors. In order to compare the level of looping for networks on different scales, a

scale parameter is assigned to the generated network which leads to a median pipe length (L) of 231 m. The number of loops (meshedness coefficient, r_m) and loop size (median loop size, A) for the generated network are all within the reasonable ranges. According to algebraic connectivity (λ_2), the generated WDN is more robust than the real cases reported by Yazdani and Jeffrey and less robust than those reported by De Corte and Sörensen.

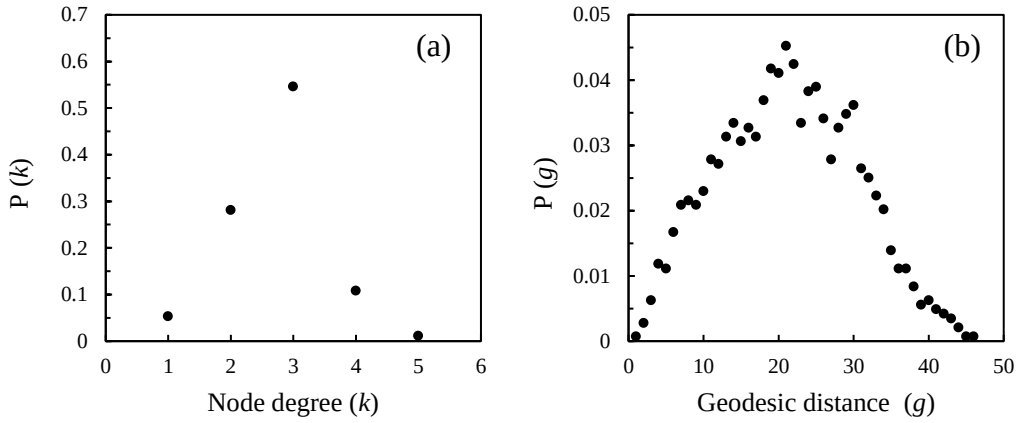


Figure 3.13. Degree distribution (a) and Geodesic distance distribution (b) of the generated WDN.

3.3.2 Comparison with different model methods

Using the optimization rule similar to equation (3.4), Gastner and NTewman [8] developed a spatial distribution network model. It generates tree structure with uniform distributed nodes whose positions are given at the outset. In their model, q drops very quickly from around 3.9 to 1.5 when β increases from 0 to 0.05, then decreases very slowly to around 1.1 as β further increases to 0.5. In our model, as β increases from 0 to 0.5, q gradually decreases from 1.27 to 1.11. The values of q for studied span of β are all within the range (1.1 -1.6) observed for real-world spatial networks [8]. The small change of q as β changes for the proposed model indicates that the design factor β has limited effect on q in real systems.

To find out the reason for this difference, three variations of the proposed model were constructed. The first variation, identified as a co-evolving tree model, removes the loop forming engineering rules to eliminate the loops but keeps the power law of demand nodes generation which coevolve as the network grows. The second model, identified as a static nodes tree model, removes the loop formation and gives out all demand nodes at the outset. The third model, a simple tree model, is similar to the second variation but the demand nodes follow a uniform distribution instead of fractal. The changing of q as β changes was plotted in Figure 3.14, which are average values based on 100 simulations.

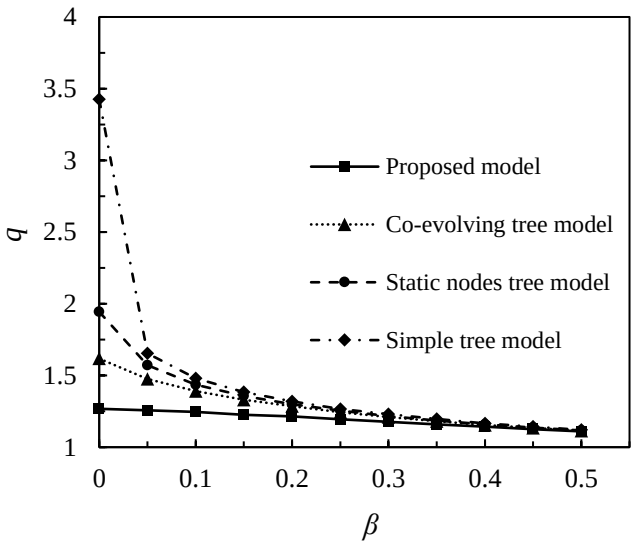


Figure 3.14. Plot of route factor (q) as a function of β (parameter settings for proposed model: α = 2.1, $d_{max} = 1.8R_0$, $N_0 = 50$, $t = 12$).

Significant differences of q for the four models appear at $\beta = 0$. At $\beta = 0$, the router factors are 1.27±0.02, 1.62±0.11, 1.94±0.16 and 3.43±0.89 respectively. This indicates that the largest discrepancy from the real network is due to the demand nodes distribution. Co-evolvement of demand nodes with the network is also important. Due to redundant pipes, formation of loops results in shorter paths to reservoir, which therefore reduces the route factor, but it is not as

important as the modeling methods for the demand nodes. Therefore, to model a real complex network, it is important to use realistic node distribution and dynamically model the node growth. A simplified model could lead to a trend that deviated from real networks. For example, with a value of 0.05 for β, the simple tree model will project a route factor of 1.65, which is approximately 30% higher than the value of 1.26 obtained from the proposed model. When the factor β is large enough, the modeling results converge. However, since the cost optimization of WDNs is in favor of MST, large β is rare in practical networks.

3.3.3 Parameters study

The proposed model has three adjustable parameters: 1) the scaling factor α allows the simulation of different urban forms; 2) the weighting coefficient β allows the simulation of relative importance of the shortest path distance, which is implied in engineering optimization process; and 3) the maximum distance d_{max} allows the testing of the impact of engineering looping criteria on the emergent properties of WDNs. The significance of the impact of each parameter on the efficiency, redundancy, and robustness were investigated by changing one parameter while fixing the other two. 100 simulations were run for each parameter setting. The efficiency is measured by the average Euclidean distance along the shortest path from all water demand nodes to the reservoir, denoted as d. This is due to the fact that the work required for delivering a certain amount of water to service point is determined by the distance of delivery. The redundancy is measured by a modified version of meshedness coefficient, which is the ratio of the number of total existing independent loops of the network to the maximum possible loops among reservoir and water demand nodes (r'_m). The robustness is measured by algebraic connectivity (λ_2). The simulation results are summarized in Figure 3.15.

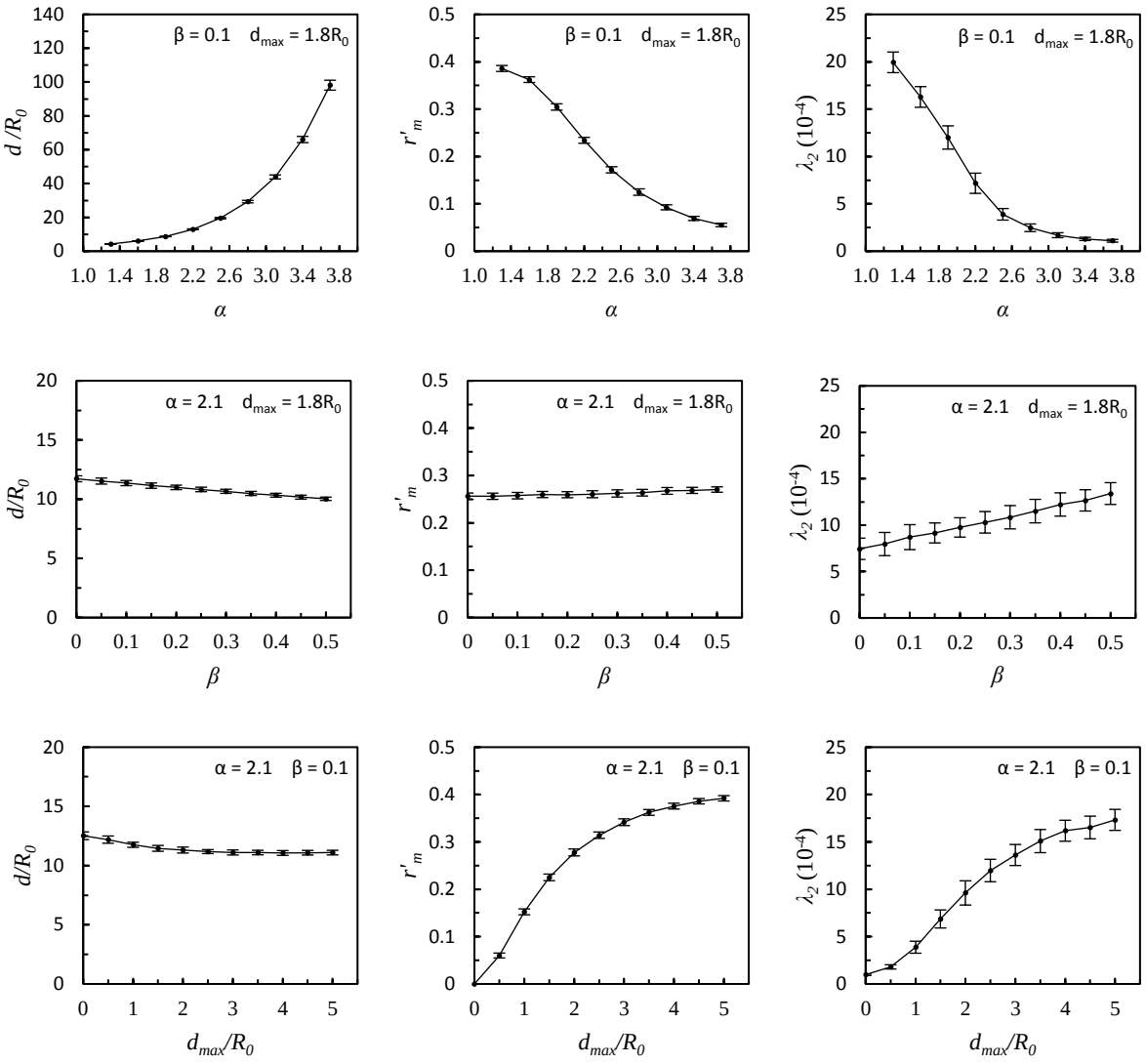


Figure 3.15. The effect of parameters (α, β, d_{max}) on the structural properties of generated WDNs (N_0=50, t=12).

The first column in Figure 3.15 shows the impacts of the parameters on the efficiency of WDNs. The scaling factor α has the most significant impact on the network efficiency. An exponential increase is observed as the scaling factor increases. Increasing the α value from 2.2 to 3.4 causes the d to increase over 5 times which means a surge in the cost of delivering service. Increasing β and d_{max} would reduce the value of d, but only slightly. The increases of β and d_{max} from 0 to their highest values studied only reduce d by 14% and 11%, respectively. This implies

that the efficiency of engineered system is largely decided by the urban growth pattern. For a sprawled city, there is limited space left for the engineers to increase the efficiency of WDNs. Staged engineering optimization, are important in a short period, yet, in the long run, the improvement of efficiency by engineering optimization is limited. An optimization considering the long-term emergency might bear more benefit. This result agrees with Frank's reanalysis[33] of studies conducted during 1950s-1980s that investigated the effect of land use on the infrastructure costs, which concluded that low density and highly sprawled development result in highest costs.

As shown in the second column of Figure 3.15, the redundancy indicator, r'_m, is most significantly affected by d_{max}. As d_{max} increases, more loops are formed in the networks, which results in higher redundancy. When it is large enough, d_{max} is no longer a constraint for loop formation. r'_m also decreases exponentially as the urban scaling factor α increases, showing the importance of scaling factor on the redundancy of WDNs. As city grows outwardly, different engineering criteria are necessary to maintain the same degree of redundancy. The local optimization parameter β is mainly controlling the growth of tree branches of the network, therefore it has limited impact on network redundancy.

The robustness of the network, as represented by the algebraic connectivity λ_2, is also controlled by the loop forming criteria d_{max} and the scaling factor α. Increasing β could increase robustness slightly by more direct connections to the reservoir. As α increases, λ_2 decreases exponentially. Yet, the increase of d_{max} can overturn the impact of sprawling.

The only factor that is important to all three properties of WDNs is the scaling factor. As the city sprawls, the costs of service delivery increase drastically and engineering solutions can only help a little. This finding may be not as optimistic as the belief that technology advancement could solve the sustainability challenge. On the other hand, the redundancy and robustness

reduction due to city expansion can be resolved through engineering methods. It is expected that increasing robustness could be achieved without sacrifice of the efficiency. Consideration of construction costs is another dimension which is not included in this analysis.

CHAPTER 4: MODEL APPLICATION

Complex network models are usually used to explain the emergence of structural properties of networks. For example, the well-known Barabási–Albert model explains the emergence of scale-free structure using a preferential attachment mechanism [7]. Similarly, the proposed complex network model of WDNs can help us understanding how the structural properties of WDNs emerge as a result of some simple rules. However, for a real-world WDN, both the structural and the hydraulic properties affect its performance (efficiency and reliability). The hydraulics of a network are obtained by solving a system of mass and energy balance equations. Except the layout, extra inputs are required for hydraulic modeling including water demand and elevation for nodes, as well as the diameter and roughness for the pipes. On the other hand, the design of WDN is typically to minimize costs while satisfying population demand and hydraulic requirement. Cost is a function of both layout and pipe sizes. Thus, in order to guide the design of WDNs, WDNs generated by complex network model is sized. As a result, the sized networks can be evaluated regarding hydraulic performance and costs can also be estimated. In this chapter, pipe sizing strategy is firstly introduced to optimize the synthetic WDNs generated by the proposed model. Then, two applications of the proposed model combined with pipe sizing algorithm to study practical WDN problems are presented.

4.1 Pipe sizing

In engineering practice, capital cost of a WDN is majorly determined by pipe diameters. Therefore, pipe sizing is a classical problem for engineers and researchers. A lot of optimization algorithms have been developed to size pipes, including deterministic methods and evolutionary

algorithms. Recently, the most popular ones are evolutionary algorithms. In this thesis, characterization of hydraulic properties of a given WDN would be performed based on the pipe sizing results obtained by the parallel hybrid estimation of distribution algorithm and particle swarm optimization (PEDPSO) [75], which can find good results with less computational time and higher reliability. The goal is to minimize the capital cost of pipes while meeting the minimum pressure requirement for each demand node (25m).

4.1.1 PEDPSO

PEDPSO is a hybrid optimization algorithm based on particle swarm optimization (PSO) and estimation of distribution algorithm (EDA). Similar to other search-based optimization algorithms such as genetic algorithms, PEDPSO will search the optimized solution from the vast solution space (which is to the power of total number of pipes). Specifically, for a given network to be optimized, each specific pipe sizing configurations (candidate solution) will be encoded as a vector of integers, each element in the vector represents the diameter of a corresponding pipe. Each encoded vector (and the candidate solution it represents) is called an "individual". This "individual" could be evaluated based on its cost (as the size for each pipe is known) as well as the head deficit (HD) measuring how much the given network is violating the head constraints. The final energy function *E* to be minimized is then the combination of these two terms:

$$E = C\left[1 + \sum_{k=1}^{K}\left(H_{K,min} - H_{K,actual}\right)\right] \tag{4.1}$$

where *C is* the total cost of the "individual", $H_{k,min}$ denotes the required minimal head of the node *k*, $H_{k,acutal}$ denotes the actual head on it, and *K* is the total number of nodes with an actual head less than the required one. The involved hydraulic calculation for water heads is done by an external hydraulic solver.

As the name suggests, the particle swarm optimization (PSO) routine of the algorithm constructs a population (i.e. “swarm”) consisting of multiple individuals (in this work set as 200), which will be updated simultaneously at each iteration of optimization. Specifically, each individual updates its pipe sizing configuration according to: 1) towards the best (i.e. minimizing energy function *E*) configuration it has achieved; 2) towards the best population-wise configuration, which is the average of individual solutions; and 3) adding randomness to the above two directions, to enlarge the searching space while ensuring that the optimization will not fall into the local optimum.

The randomness added during the updating of individual network configuration would lead to varied results for each individual. In order to pursuit population-wise optimum results, the Estimation of Distribution Algorithms (EDAs) are applied. For each iteration of PEDPSO, the whole population is ranked and split into two sub-populations with better half (i.e. lowest energy functions) updated by EDA and other half updated by PSO. Then, the two updated sub-populations are combined for the next iteration. This is the key difference between PEDPSO and traditional PSO, in which the whole population is updated only by PSO. Using the distribution estimation-sampling scheme has been shown to be more effective as it helps population diversity control and avoids premature convergence. The details of PEDPSO can be found in Qi al. et [75].

4.1.2 Combination of PEDPSO and graph theory

To further save computational time, a near-optimum solution is pursued by combing PEDPSO and graph theory, which reduces the number of pipe variables in optimization. Specifically, by assuming water is efficiently delivered to node from source through the shortest distance path, the whole WDN is decomposed into a shortest-distance tree and remaining edges. An example of a Hanoi WDN is shown in Figure 4.1. For a completely new WDN design, only

new pipes in the shortest-distance tree are subjected to pipe sizing while other remaining new pipes are all set to minimum allowable diameter. For a WDN expansion design, the existing network may need upgrading in order to meet the hydraulic requirements of growing water demands. In such case, parallel pipes will be built that are parallel to existing pipes. That means that they are connecting two already directly connected nodes. In this work, only existing pipes in the shortest-distance tree are considered to place parallel pipes. Thus, only new pipes in the shortest-distance tree as well as possible parallel pipes are subjected sizing by PEDPSO, while other new pipes are all set to minimum allowable diameter. Hydraulic simulation is still done to the whole looped network rather than the tree network, which ensures the minimum required pressure to the whole network. A similar method was used by Zheng et al. [76] to obtain a near-optimum solution for WDN design.

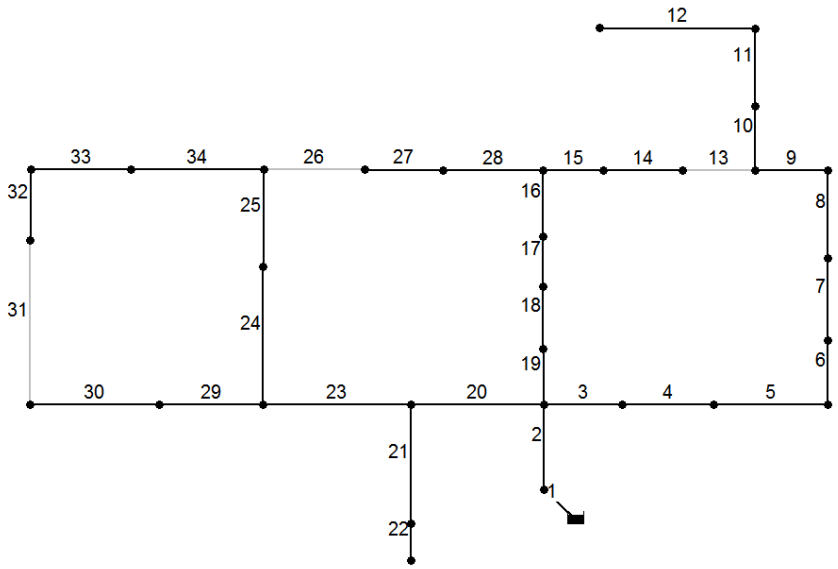


Figure 4.1. Hanoi water distribution network: shortest-distance tree is composed of black edges and remaining edges are pipe 13, 26 and 31 which are shown in gray.

4.1.3 Hydraulic analysis

Hydraulic simulations are done by an external solver EPANET 2 [77] according to the maximum hourly water demand. Head loss (h_L) calculation is based on Hazen-Williams (H-W) equation:

$$h_L = 1.852L(Q/c)^{1.852}D^{-4.871} \quad (4.2)$$

where Q is pipe flow rate (m^3/s); D is the pipe diameter (m); c is pipe roughness coefficient, set to 130 for all pipes in this work.

4.1.4 Pipe options

The capital costs of new pipe include excavation, bedding, material cost, installation cost, and backfill, which are estimated according to the methodology of Clark et al. [78]. Parallel pipes are placed parallel to existing pipes, which implies disruption and reconstruction of pavement roads, thus an additionional 20% cost is accounted for parallel pipes. Pipe diameter options and costs are given in Table 4.1.

Table 4.1. Pipe diameter options and capital costs.

Diameter (inch)	Diameter (mm)	New pipe ($/m)	Parallel pipe ($/m)
6	152.4	84.12	100.94
8	203.2	94.78	113.74
10	254.0	121.36	145.63
12	304.8	147.83	177.40
14	355.6	181.76	218.11
16	406.4	201.71	242.05
18	457.2	247.05	296.46
20	508.0	286.58	343.90
24	609.6	355.87	427.05
30	762.0	406.33	487.60
36	914.4	467.16	560.59
42	1066.8	550.95	661.14
48	1219.2	635.96	763.15

54	1371.6	812.34	974.80
60	1524.0	1069.95	1283.94
66	1676.4	1273.72	1528.46
72	1828.8	1330.77	1596.93

4.2 Application I: Effect of system size on cost efficiency of the system

4.2.1 Introduction

According to 2006 Community Water Supply Surveys (CWSSs), there are 49,133 community water systems (CWSs) in the US, which vary greatly in size (served population from less 100 to over 500,000). About 82 percent of CWSs serves 3,300 or fewer people, which is considered as small. Those systems just serve only 11 percent of the total population. On the other hand, about 1 percent of CWSs serve more than 100,000 people, which provide water to 46 percent of the total population.

Conventionally, small water supply system (WSS) is thought to have high cost for water supply and lack technical, managerial and financial capacity to meet modern water treatment requirements [79]. The economies of scale in water industry have been investigated based on local cost structures in many countries [80-85]. Although different results were obtained for different studies, evidence of economies of scale has been found in many studies [85]. On the other hand, water quality [86] and distribution pumping cost [87] may be limiting factors to the effectiveness of large centralized water supply systems. Along with economies of scale, evidence of diseconomies of scale was also found in many countries [85].

Therefore, consolidation of small water systems and decentralized water distribution system are suggested as ways to improve the efficiency of water supply systems. Some research has been done to try to find the optimum size of a water supply system based on empirical studies

of local cost structures. Mizutani and Urakami [88] investigated the Japanese water supply organization and concluded that its optimal size would be one serving a population of approximately 766,000 persons. Saucer [89] studied water firms in rural areas of former East and West Germany, and found the optimal water firm is to be on average about three times larger than the existing one.

In this section, simulated water distribution systems with different sizes are constructed to investigate the effect of system size on the cost efficiency of water supply system. The layout of those synthetic test cases is simulated by a complex network model introduced in section 3.2. The optimization of test cases is done based on method described in the first section of this chapter.

4.2.2 Methodology

4.2.2.1 Water supply system design

Population and density

In this study, water supply systems with seven different sizes are investigated, which are serving a population from 62624 to 1001984 (Table 4.2). The population density is set to be 978.5 capital/km^2 for all test cases, which is the average urban population density according to 2010 US census data.

Table 4.2. System size.

System size	Population	Water demand (MGD)
1	62624	8.545
2	140904	19.282
3	250496	34.291
4	391400	53.574
5	563616	77.238
6	767144	105.066
7	1001984	137.166

Water demand

Shammas and Wang [90] showed typical demand values for total water demand in the US ranging from 227.1~1324.9 l/capita/day. In this study, average daily water demand is set to 518.4 l/capita/day, and the maximum hourly water demand is assumed to be 3 times of the average daily demand. The total water demand for each system size is shown in Table 4.2.

Water distribution network

Layout

For each system size, 100 WDNs are generated. The layout of water distribution networks is simulated by a dynamic complex network model (section 3.2). In order to make the generated WDNs comparable between different sizes, urban growth pattern and pipe generation strategy are set to be the same for all test cases with α = 2.1, β = 0.1 and d_{max} = 1500 (m), where α is scaling factor of urban growth, β is weighting coefficient of shortest path distance for tree structure generation and d_{max} is maximum allowable distance for loop formation. The difference lies in the initial size of networks including initial number of nodes (N_0) and initial side length of the city ($2R_0$, m), as listed in Table 4.3, while population density, population growth rate (13.8 %) and time steps (t = 10) for the simulation are all the same. Each node represents a population of 800.

Table 4.3. Parameters for layout generation.

System size	N_0	R_0
1	25	1.25
2	55	1.87
3	98	2.50
4	153	3.12
5	220	3.75
6	300	4.37
7	391	5.00

Pipe sizing

The generated WDNs at t = 10 are sized by combining PEDPSO and graph theory, which is introduced in section 4.1. Parameters in the optimization algorithm for different system sizes are listed in Table 4.4, including reservoir head, minimum diameter (D_{min}), maximum diameter (D_{max}) and number of iterations.

Table 4.4. Parameters for pipe sizing.

System size	Reservoir head (m)	D_{min} (inch)	D_{max} (inch)	Iteration
1	49	8	24	400
2	59	8	36	700
3	70	8	48	1500
4	76	8	60	2500
5	84	8	72	4000
6	91	8	84	6000
7	99	8	96	8000

4.2.2.2 Cost estimation models

Costs of water supply systems considered in this study include distribution pipelines capital, pumping energy cost, DWTP construction cost and operation & maintenance (O&M) cost. Distribution system maintenance/repair is not considered currently. The design life for water supply systems is 50 years.

Pipeline cost

The total capital cost for a distribution network is calculated as:

$$Y = \sum_{i=1}^{N} C_i L_i \tag{4.3}$$

where Y is the total capital cost of pipelines ($); C_i is the cost per unit length of pipe i ($/m); L_i is the total length of pipe i (m); N is the total number of pipes.

Energy Cost

In this study, energy is only consumed to provide enough head for the central reservoir. The total annual energy cost is calculated according to the following equation:

$$y_2 = e * c \tag{4.4}$$

where y_2 is the total annual energy cost ($/yr); e is annual energy consumption (kWh/yr); c is the unit energy cost ($/kWh), which is set to 0.1044$/kWh according to average US electricity retail price in 2014 (http://www.eia.gov/electricity/state/). Annual energy consumption is estimated according to normal demand condition with a pumping efficiency 75%. Considering interest rate, the total energy cost during the system design life is calculated as:

$$Y_2 = y_2 + \frac{y_2}{(1+r)} + \frac{y_2}{(1+r)^2} + \cdots\cdots + \frac{y_2}{(1+r)^{m-1}} \tag{4.5}$$

where Y_2 is total energy cost during the design life of water supply system ($); r is the annual interest rate; m is system design life (yr).

DWTP Construction Cost and O&M cost

For all test cases, a centralized drinking water treat plant (DWTP) is assumed to treat all raw water, which is set to be in the same location as water source. It is modeled as a conventional DWTP including chemical addition (typical chemicals include Potassium Permanganate, Alum or Ferric Salts, Polymers, and Chlorine), flocculation, sedimentation, filtration, and disinfection according to the specifications of McGivney and Kawamura [91]. The capital cost and O&M cost of a conventional DWTP are estimated as following [92]:

$$Y_3 = 8.7684 * x^{0.5957} * 10^6 \tag{4.6}$$

$$y_4 = 0.4384 * x^{0.2946} * 10^6 \tag{4.7}$$

where Y_3 and y_4 are the total capital cost ($) and annual O&M cost ($/yr) of DWTP respectively; x is the design flow (MGD). The total DWTP O&M cost during its design life is:

$$Y_4 = y_4 + \frac{y_4}{(1+r)} + \frac{y_4}{(1+r)^2} + \cdots\cdots + \frac{y_4}{(1+r)^{m-1}} \quad (4.8)$$

where Y_4 is total O&M cost during the design life of DWTP ($).

Total and unit costs of water supply system

The total and unit cost of a water supply system for the its design life are calculated as:

$$Y_T = Y_1 + Y_2 + Y_3 + Y_4 \quad (4.9)$$

$$Y_u = \frac{Y_T}{Q * m} \quad (4.10)$$

where Y_T and Y_u are the total cost ($) and unit cost ($/m^3) of a water supply system during its design life; Q is annual water consumption (m^3/yr).

4.2.3 Results and discussion

4.2.3.1 Simulated WDN layout

Figure 4.2 shows an example of generated WDN for each system size. Larger networks consist of more number of nodes and edges as listed in Table 4.5. Table 4.5 also listed some graphic metrics that were used by researchers to measure structural properties of WDNs [9, 37, 93], including average node degree (<k>), meshedness coefficient (r_m) and route factor (q). Node degree is the number of edges connected to a node. Average node degree is the average number of edges per node, which measures network connectivity. Tree-structure networks have an average node degree of about 2 while complete grid networks have an average node degree of about 4. As shown in Table 4.5, the generated networks have an average node degree lager than 2 and smaller than 3, which agree with real WDN patterns that are usually partially looped and less connected than complete grid networks [37]. Meshedness coefficient is the ratio of the number of total existing independent loops to the maximum possible loops, which is an indicator for the network redundancy. De Corte and Sörensen [37] studied 10 real networks and found that their meshedness

coefficient is mostly between 0.01 and 0.1, while the highest one is 0.32. The meshedness coefficient of generated networks indicates that those networks are well redundant. Route factor measures the straightness of paths from other nodes to root, which is the average ratio of shortest distance from a node to the root through edges to its direct Euclidean distance to the root. It can be used as an indicator of network efficiency. The route factor for generated networks (1.22-1.25) is smaller than the four real WDNs (1.45-1.67) reported by Yazdani and Jeffrey [9]. This is expected because the layout simulation model has ignored certain practical constrains in WDN development such as geography and hydraulics factors. But it is still within the range (1.1 -1.6) observed by Gastner and Newman for other real spatial networks [8]. Comparing the structural properties between different system sizes, the difference is minor.

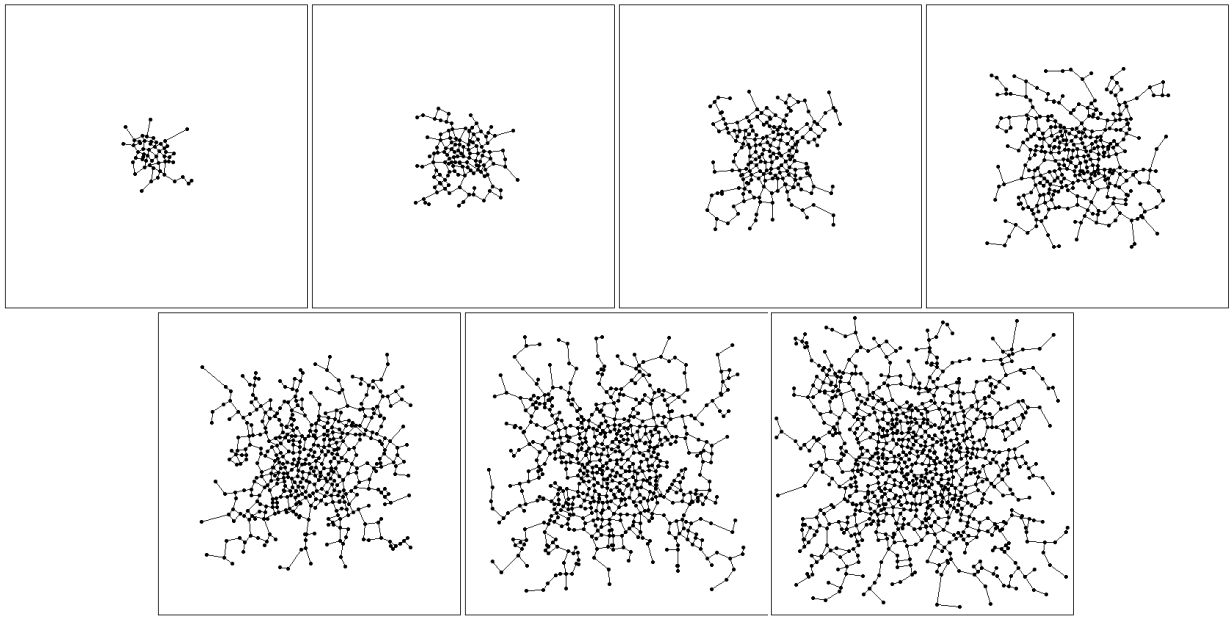


Figure 4.2. An example of generated WDN for each system size.

Table 4.5. Structural properties of simulated WDNs for all system sizes.

System size	Node number		Edge number		Average Degree		Route Factor		Meshedness coefficient	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
1	51	4	69	6	2.70	0.08	1.24	0.05	0.193	0.020

2	118	6	163	10	2.76	0.06	1.25	0.05	0.199	0.015
3	211	9	293	14	2.77	0.04	1.24	0.04	0.197	0.011
4	334	11	466	17	2.79	0.03	1.23	0.03	0.200	0.008
5	477	13	666	19	2.79	0.03	1.23	0.02	0.200	0.007
6	653	13	913	20	2.80	0.02	1.22	0.02	0.200	0.005
7	855	18	1196	29	2.80	0.02	1.22	0.02	0.201	0.006

4.2.3.2 Size effects on cost efficiency of water supply system

The average total costs for all system sizes with a 50-year design life are shown in Figure 4.3. Those costs are estimated based on an annual interest rate of 3%. As shown in Figure 4.3, the total costs of WDSs become significantly larger with system size. Although all four items cost more as system size increases, their growth rate are different with the order: pipeline capital > pumping energy > DWTP capital > DWTP O&M. Even though the total costs of DWTP (capital cost and O&M cost) is more than the total costs of distribution system (pipeline capital cost and pumping energy cost) when system is small, as system grows larger, pipeline capital cost is approaching DWTP capital cost and pumping energy has exceeded DWTP O&M cost for the largest study system size.

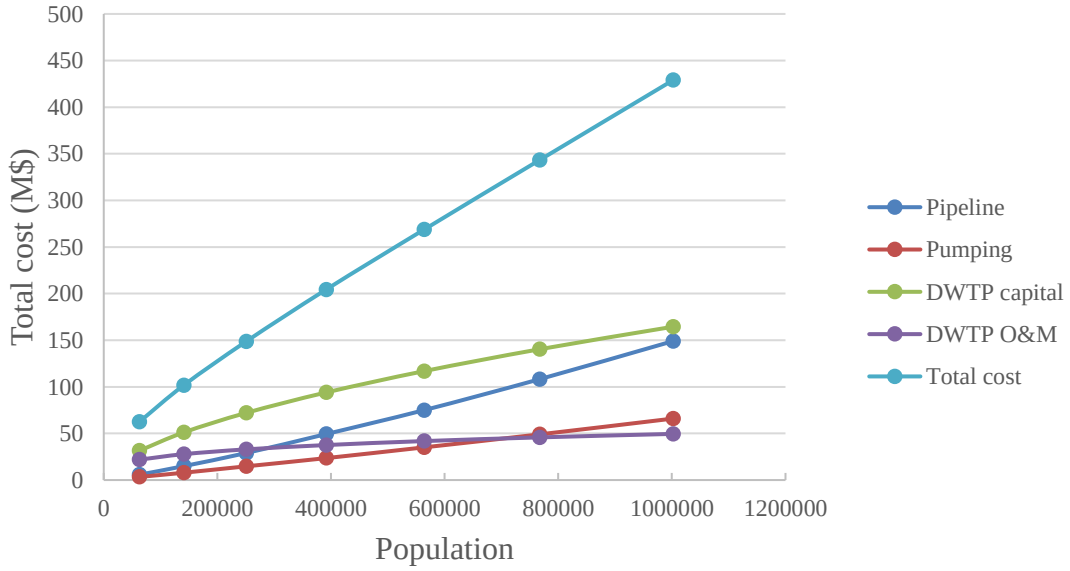


Figure 4.3. Total costs for all system sizes.

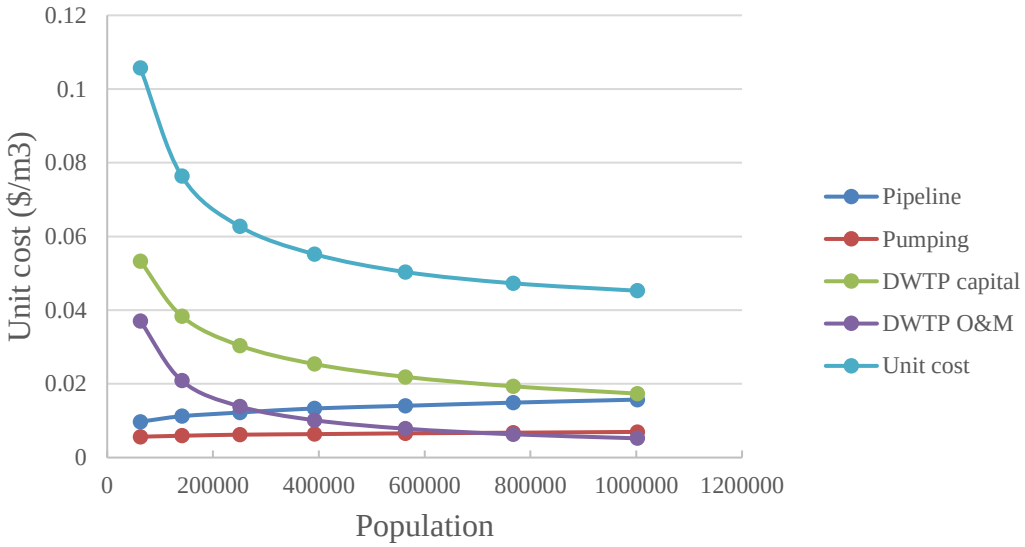


Figure 4.4. Unit costs for all system sizes.

The unit costs for all system sizes are show in Figure 4.4. The units cost of the whole system are decrease as system size increases. Agreed with other DWTP studies, economies of scale are existed in DWTP costs including construction costs and O&M costs. On the other hand, the unit costs of distribution system including pipeline construction cost and pumping energy cost are increasing as system size increases. Even though their increase is compromised by the decrease of DWTP costs for the studies cases, which lead to the decreased unit cost for whole systems, they should be able to exceed DWTP costs at a system size larger than what have presented in this study.

4.3 Application II: Impact of planning horizon on network design

4.3.1 Introduction

As a critical part of urban systems, water distribution networks (WDNs) continuously expand themselves over time to meet the growing demand of population growth [25]. The master planning of WDNs is generally done over short-term planning horizons (typically about 20 years) with the objective of minimizing system cost to meet water demand by the end of the planning

horizon. Such practice prohibits excess capacity, thus the system might fail to meet the new water demand after the planning period and capacity expansion would be required [94].

Long-term planning is favored considering economies of scale in pipes. Walski [69] derived an equation describing the relationship between pipe cost and its capacity: cost = $2.67 \times \text{capacity}^{0.567}$. This equation indicates that installing a large pipe costs less than several smaller pipes at the same capacity. On the other hand, it may avoid overdesign and save interest cost to build smaller parallel pipes over time instead of installing a larger pipe. There is a trade-off between economies of scale and excess capacity cost due to interest rate.

The problem of expansion size and timing is generally studied in the field of capacity expansion [95-97]. On WDN design, research has been done on expansion scheduling over a planning period [98-101]. Only a few papers were devoted to study the optimum planning horizon for WDN expansion. Applying the model introduced by Manne [97] to determine the optimal expansion size for a new facility, Scurato [102] obtained the optimum expansion cycle for a water pipeline considering an infinite time period and assuming a linear water demand growth. The optimum planning cycle was found to depend on interest rate. Higher interest rates lead to shorter optimum expansion cycles. Since water pipelines usually have a life span over 100 years, Walski [69] used a single pipe section to address the problem of optimum planning horizon for a total of an 100-year period. He found that optimum planning horizon drops from roughly 60 years to 40 years as annual interest rate increases from 1% to 5%. Considering the complexity feature of WDNs, this paper investigated the optimum planning horizon by simulating the dynamic growth of a whole network using spatial network models.

WDNs are complex networks which consist of large number of interconnected components, as suggested by Yazdani and Jeffrey [9]. Currently, complex network researchers have introduced

new concepts and tools to describe and model the complex behaviors of complex networks. This research provides foundation to simulate the complex topology of WDNs and their dynamics. Since most water networks are planar networks limited by geography and engineering optimization, a test case of WDN expansion was generated in this study by incorporating engineering rules into a spatial distribution network model proposed by Gastner and Newman [8]. The planning of networks serving both linear and exponential population growth patterns was investigated. The impact of planning horizon on the design of test case is investigated under different annual interest rates. Optimization of the test case over a planning horizon is done based on an evolutionary algorithm, the parallel hybrid estimation of distribution algorithm and particle swarm optimization (PEDPSO) [75].

4.3.2 Methods

4.3.2.1 Test case

The test case in base scenario is a synthetic WDN serving a squared city with a population of 106029 that is uniformly distributed in the urban area. The average urban population density from 2005 and 2010 US Census data is around 1000 capita/km^2. In this study, the population density of the study city is set to 1060.29 capita/ km^2. The water network consists of a reservoir situated in the center of the city, 145 junctions and 193 pipes (Figure 4.5). Shammas and Wang [90] showed typical demand values for total water demand in the US ranging from 227.1~1324.9 l/capita/day. In this study, the average daily water demand is set to 518.4 l/capita/day, and the maximum hourly water demand is assumed to be 3 times of the average daily demand.

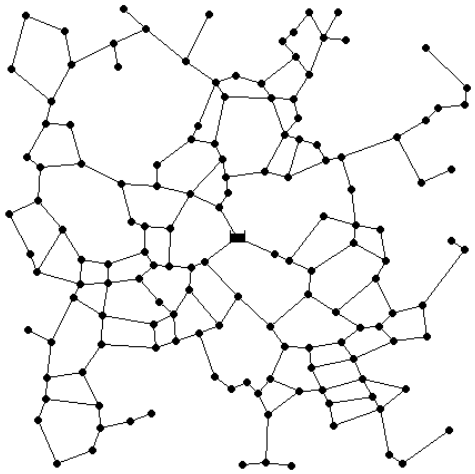


Figure 4.5. Test case WDN in base scenario.

4.3.2.2 Population growth and water demand

The total study span for the test case is 100 years. Two population growth scenarios are investigated. One is exponential growth with a growth rate of 9.7% per 10 years which is the US population growth rate in the first decade of 21st century reported by 2010 US census. The other one is linear growth pattern. System-wide water demand at different time points is shown in Table 4.6 for both growth patterns. At the end of the study span, both scenarios had the same size of population, as well as the same water demand needs.

Table 4.6. System sizes.

	Linear growth			Exponential growth		
	Population	Area (km^2)	Water demand (MGD)	Population	Area (km^2)	Water demand (MGD)
Base	106029	100	14.517	106029	100	14.517
10 years	121968	132	16.7	116424	120	15.941
20 years	137907	169	18.882	127512	145	17.459
30 years	153846	211	21.065	139986	174	19.167
40 years	169785	256	23.247	153846	210	21.065
50 years	185724	307	25.429	168399	252	23.057
60 years	201663	362	27.612	185031	304	25.334
70 years	217602	421	29.794	203049	366	27.801
80 years	234234	488	32.071	222453	440	30.458

90 years	250866	560	34.349	243936	529	33.4
100 years	267498	637	36.626	267498	637	36.626

4.3.2.3 Modeling of WDN layout expansion

In this work, the growth of WDN is simulated by a dynamic complex network model (section 4.2) under linear and exponential population growth patterns respectively. The layout expansion is simulated in 10 time steps for a total study span of 100 years with each time step set as 10 years. Urban growth pattern and pipe generation strategy are set to constant with α = 2.0, β = 0.1 and d_{max}= 1200 (m), where α is scaling factor of urban growth, β is weighting coefficient of shortest path distance for tree structure generation and d_{max} is maximum allowable distance for loop formation.

4.3.2.4 Modeling of WDN Pipe sizing

In this study, four different planning horizons are investigated: 10 years, 20 years, 50 years and 100 years. With the 10-year planning horizon, 10 sequenced designs of WDN expansion are required for the whole study span (100 years), while only a one-time design is needed with 100-year planning horizon. The strategy for each design is the same, which is described as below.

Problem description

For each design, the problem is to minimize the total capital cost of all pipes while meeting water demand requirements. The decision variables are diameters of new pipes connecting new demand nodes and parallel pipes to the existing pipes. The design of WDN expansion is performed in phases to meet gradual water demand growth. Each phase is assumed to be 10 years. For 10-year planning horizon, the designed WDN expansion is constructed through one phase. The design is done based on the demand by the end of planning period, while the cost is calculated at the

beginning of that planning period. For a longer planning horizon that consists of several phases, new pipes connecting new water demand nodes are constructed by different phases which follow the projected time of new nodes, while all parallel pipes are assumed to be constructed in the first phase of the planning period. The total cost (C) for the design is calculated as the present value at the beginning of the planning period. Mathematically, the objective function is described as:

$$C = \sum_{i=1}^{Z_1} c_i^0 L_i + \sum_{j=1}^{Z_2} \frac{c_j L_j}{(1+r)^{(k_j-1)y}} \tag{4.11}$$

where c_i^0 and c_i are the cost per unit length of the pipe used in parallel pipe i and new pipe j respectively; L_i and L_j are the length of the parallel pipe i and new pipe j respectively; Z_1 and Z_2 are the total number of the parallel pipes and new pipes built in the network respectively; r is annual interest rate; k_j is the phase index in which new pipe j is constructed; y is the number of years in each phase (10 years). One hydraulic constraint is included in the optimization model: minimum allowable pressure (25m) requirements for each demand node. In the optimization process, costs involving pumping are not considered. Reservoir head is predefined before optimization. In this study, reservoir head is set to increase as WDN expands (Table 4.7) which is independent of planning horizons.

Pipe options

Parameters in optimization algorithm for different system sizes are listed in Table 4.7, including reservoir head, minimum diameter (D_{min}), maximum diameter (D_{max}).

Table 4.7. Design parameters for pipe sizing.

	Reservoir head (m)		D_{min} (inch)	D_{max} (inch)
	Linear growth	Exponential growth		
Base	50	50	-	-
10 years	55	53	6	36

20 years	60	56	6	36
30 years	65	60	6	36
40 years	70	64	6	36
50 years	75	69	6	42
60 years	80	74	6	42
70 years	85	80	6	42
80 years	90	86	6	48
90 years	95	93	6	48
100 years	100	100	6	48

4.3.2.5 Total present cost

For each planning horizon scenario, the total present cost (*PC*) of WDN expansion over 100 years is calculated as following:

$$PC = \sum_{p=1}^{P} \frac{C_p}{(1+r)^{(p-1)Y}} \tag{4.12}$$

where C_p is total cost for design p; P is total number of designs; and Y is the total number of years in the planning horizon.

4.3.3 Results and Discussion

4.3.3.1 Simulated WDN layout expansion

Using the dynamic WDN layout modeling introduced in Figure 2.1, WDN expansion at 10 different time steps is simulated based on two growth patterns (linear/exponential), as shown in Figure 4.6 and Figure 4.7.

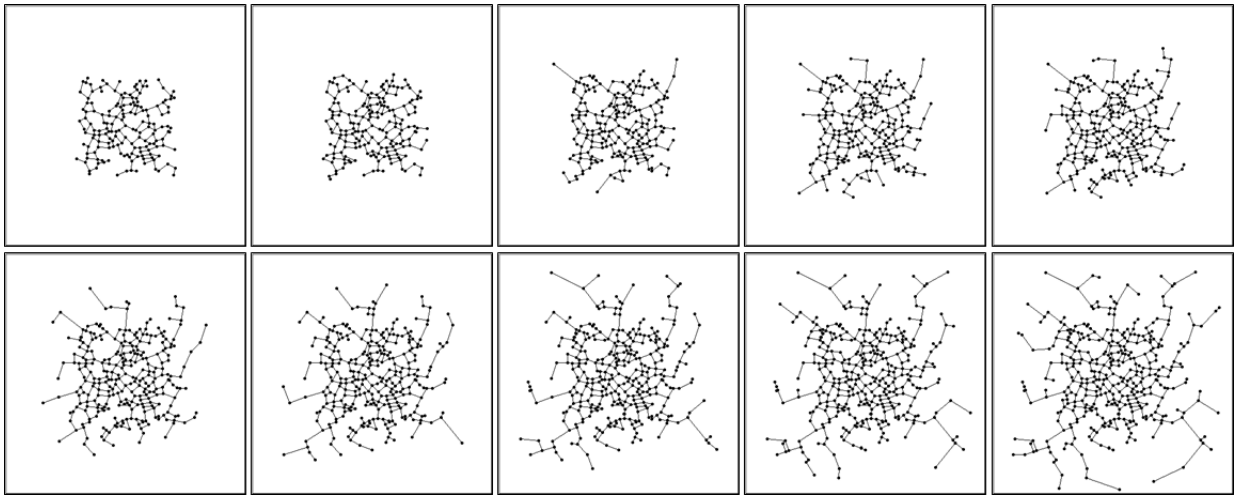


Figure 4.6. Layouts of WDN expansions at 10 different time steps based on linear growth: (1) 10 years; (2) 20 years; (3) 30 years; (4) 40 years; (5) 50 years; (6) 60 years; (7) 70 years; (8) 80 years; (9) 90 years; (10) 100 years.

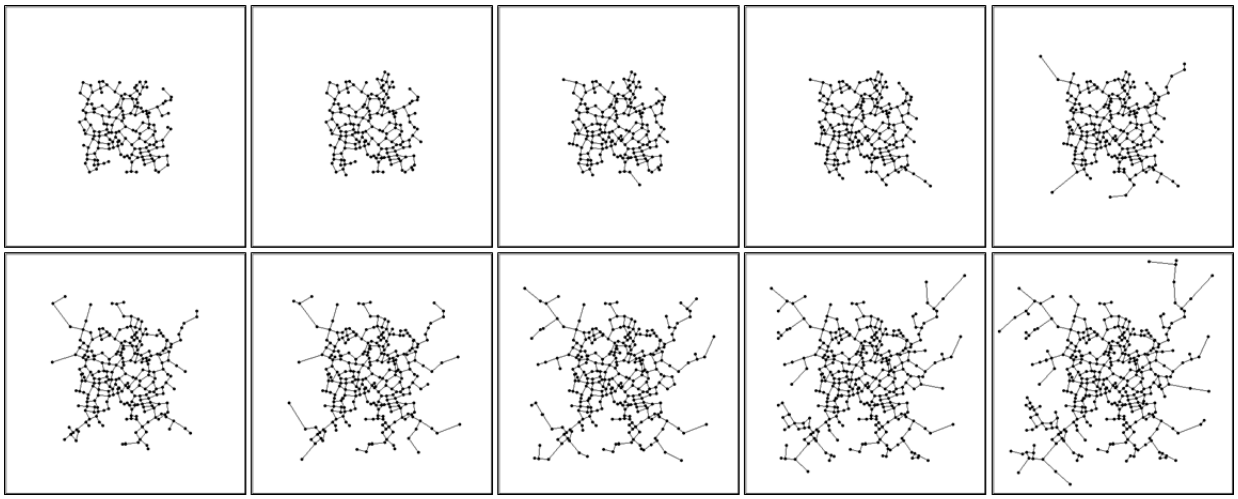


Figure 4.7. Layouts of WDN expansions at 10 different time steps based on exponential growth: (1) 10 years; (2) 20 years; (3) 30 years; (4) 40 years; (5) 50 years; (6) 60 years; (7) 70 years; (8) 80 years; (9) 90 years; (10) 100 years.

The simulated WDNs expand as time step progresses, incorporating an increasing number of nodes (n), edges (m) and independent loops (l), as listed in Table 4.8. Table 4.8 also listed some graphic metrics that were used by researchers to measure structural properties of WDNs [9, 37,

93], including average node degree (<k>), maximum node degree (k_{max}), meshedness coefficient (r_m) and route factor (q). Node degree is the number of edges connected to a node. Average node degree is the average number of edges per node, which measures network connectivity. Tree-structure networks have an average node degree of about 2 while complete grid networks have an average node degree of about 4. As shown in Table 4.8, the generated networks have an average node degree lager than 2 and smaller than 3, which agree with real WDN patterns that are usually partially looped and less connected than complete grid networks [37]. The maximum node degree for all generated networks is small (4 or 5), which indicates that WDNs are not scale-free networks [7] featured with the existence of highly connected nodes. Meshedness coefficient is the ratio of the number of total existing independent loops to the maximum possible loops, which is an indicator for the network redundancy. As listed in Table 4.8, even though the number of independent loops is increasing as WDN expands, the redundancy of the whole network do not necessary increase. Route factor measures the straightness of paths from other nodes to root, which is the average ratio of shortest distance from a node to the root through edges to its direct Euclidean distance to the root. It can be used as an indicator of network efficiency. The route factor for generated networks (1.22-1.25) is smaller than the four real WDNs (1.45-1.67) reported by Yazdani and Jeffrey [9]. This is expected because the layout simulation model has ignored certain practical constrains in WDN development such as geography and hydraulics factors. But it is still within the range (1.1 -1.6) observed by Gastner and Newman for other real spatial networks [8].

Table 4.8. Structural properties of the simulated WDN at different time steps (n, number of nodes; m, number of edges; l, number of independent loops; L_{avg} (m), average edge length; $<k>$, average node degree; k_{max}, maximum node degree; r_m, meshedness coefficient; q, route factor).

	n	m	l	L_{avg}	$<k>$	k_{max}	r_m	q
Linear growth								
10 years	165	216	52	674	2.62	4	0.16	1.24
20 years	184	247	64	670	2.68	5	0.18	1.23
30 years	210	282	73	678	2.69	5	0.18	1.23
40 years	233	309	77	700	2.65	5	0.17	1.22
50 years	259	346	88	689	2.67	5	0.17	1.23
60 years	282	376	95	704	2.67	5	0.17	1.23
70 years	308	406	99	713	2.64	5	0.16	1.23
80 years	338	440	103	716	2.60	5	0.15	1.23
90 years	359	469	111	729	2.61	5	0.16	1.23
100 years	378	490	113	751	2.59	5	0.15	1.23
Exponential growth								
10 years	157	209	53	667	2.66	4	0.17	1.23
20 years	171	229	59	658	2.68	4	0.18	1.24
30 years	185	250	66	656	2.70	4	0.18	1.24
40 years	206	281	76	642	2.73	4	0.19	1.23
50 years	223	302	80	668	2.71	4	0.18	1.23
60 years	254	339	86	667	2.67	4	0.17	1.24
70 years	282	373	92	685	2.65	5	0.16	1.24
80 years	313	409	97	700	2.61	5	0.16	1.25
90 years	346	451	106	710	2.61	5	0.15	1.25
100 years	379	489	111	715	2.58	5	0.15	1.25

4.3.3.2 Pipe sizes

In this study, the simulated test case is sized under different planning horizons and annual interest rates. The scenarios for planning horizon include 10 years, 20 years, 50 years and 100

years. The scenarios for interest rate include 1.0%, 1.5%, 2.0%, 2.5%, 3.0%, 4.0%, 5.0% and 6.0%, where the rate is assumed to be constant over time steps. Thus, the total number of experiments is 32 for each growth pattern.

Since the pattern of pipe diameters under different annual interest rates for both growth patterns are similar, only the optimization result for one typical scenario is analyzed in this paper, which is 3.0% annual interest rate for linear growth pattern. In this scenario, the total number of new pipes built during 100-year study span is 532 with 10-year planning horizon, 493 with 20-year planning horizon, 387 with 50-year planning horizon and 328 with 100-year horizon. Since the layout of WDN is the same for all planning horizons, the difference in the pipe number indicates that WND design strategy under shorter-term planning horizon results in more parallel pipes to be built in order to meet minimum pressure requirement as the demand increases. Distribution of pipe diameters under different planning horizons is shown in Figure 4.8, indicates that designs with longer-term planning horizon results in increased number of large pipes which provide excess capacity at the beginning of a planning period.

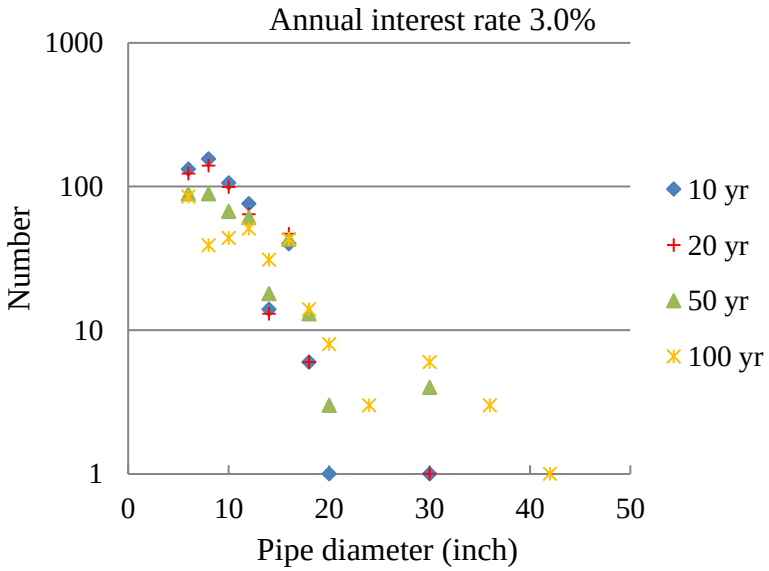


Figure 4.8. Distribution of pipe diameters by different planning horizons for linear growth pattern.

4.3.3.3 Cost analysis

The total present costs for different scenarios are shown in Figure 4.9. Figure 4.9 shows that the optimum planning horizon (i.e. achieving minimum total present cost) for the study case is dependent on annual interest rate, which was as expected. Generally, the results suggested that the shorter planning horizon is favored with a higher annual interest rate while the longer planning horizon is favored with a lower interest rate.

For the linear growth pattern, the optimum planning horizons are 100 years with annual interest rate lower than 1.5%, 50 years with interest rate of 2.0~2.5%, and 20 years with interest rate of 2.5%~6.0%. Walski [69] studied the impact of planning horizon on WDN design under different interest rates for a linear population growth pattern by using a single pipe section and concluded that optimum planning horizon was from about 60 years to 40 years as annual interest rate increases from 1% to 5%. The differences are due to the fact that this research modeled the whole network instead of a single section.

To further elaborate the mechanics of how planning horizon affects WDN designs under different interest rate, two scenarios with lower bound interest rate (1.0%) and upper bound interest rate (6.0%) were selected for further investigation. As analyzed in section 3.2, design under longer-term planning horizon results in larger pipes than shorter-term planning horizon, which are built at the beginning of a planning period to provide excess capacity. In the scenario of low interest rate as shown in Figure 4.10, even though excess capacity requires extra cost at beginning, a long planning horizon can actually save money in the long run because economies of scale play the major role, while the impact of interest rate can be ignored. However, when the interest rate becomes high, the total present costs for WDN expansion are mainly determined by the costs at

the beginning period. The long planning horizon design would lose its advantage from economies of scale due to too much extra cost at the beginning period (Figure 4.10).

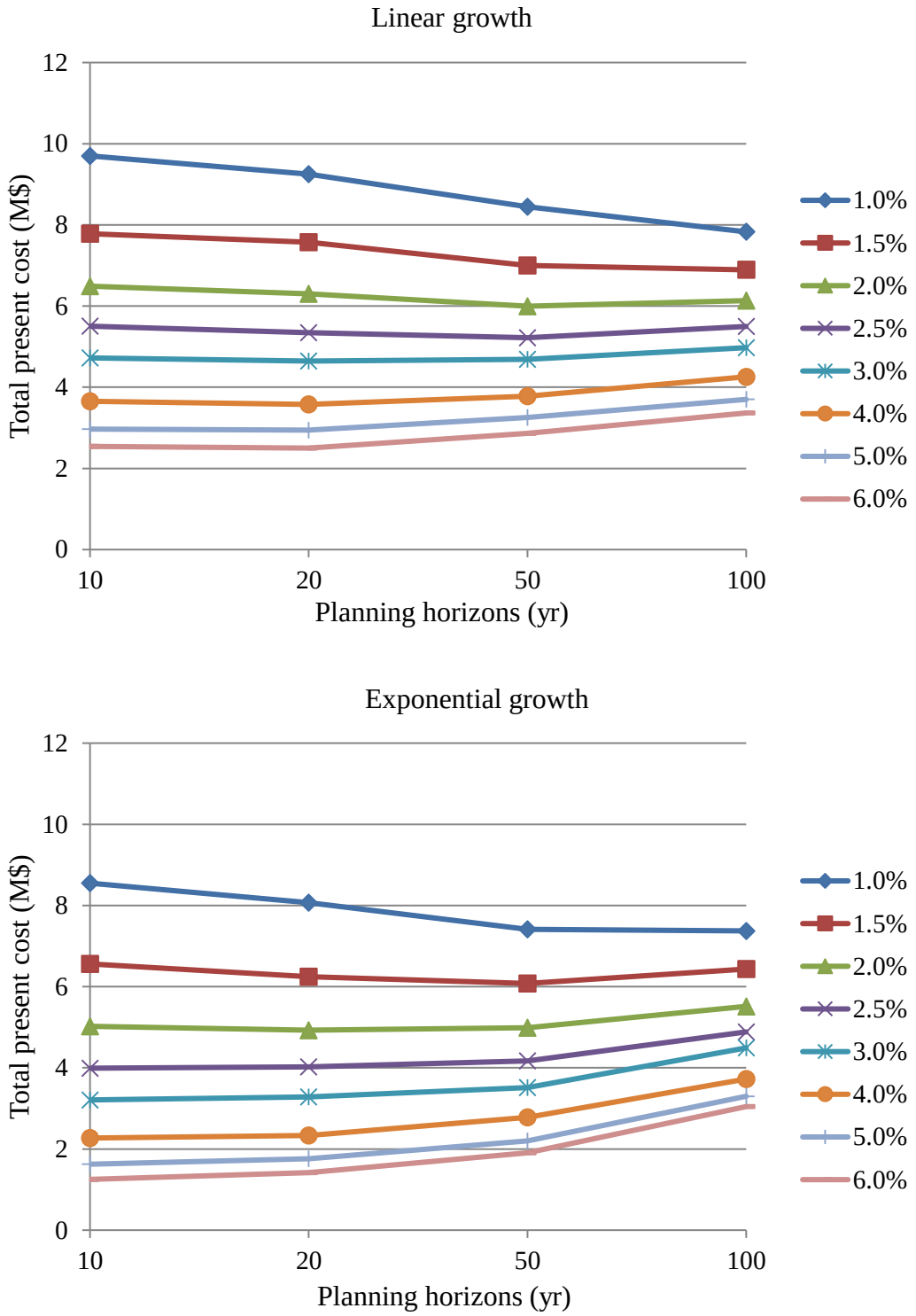


Figure 4.9. Total present cost as a function of planning horizon and interest rate.

For test case with exponential growth pattern, the optimum planning horizons are 100 years, 50 years and 20 years for annual interest rate of 1.0%, 1.5% and 2.0%, respectively. When the annual interest rate goes up to 3.0% or more, the optimum planning horizon is 10 years. In general, the optimum planning horizons are shorter than those under linear growth pattern. Comparing to linear growth, exponential growth imposes less new water demands at the beginning of the study period. Thus for a longer-term plan horizon, even though the total demand at the end of the study span is the same for both linear and exponential growth patterns, exponential growth will result in more excessive capacity at the beginning of study period, which also results in higher extra cost induced by interest. As shown in Figure 4.11, the difference in present cost at the beginning of study span between different planning horizons become much higher in exponential growth case. When interest rate is as high as 6.0%, the extra cost for excessive capacity can completely exceed the saving from economies of scale, resulting in the preference towards a shortest planning horizon.

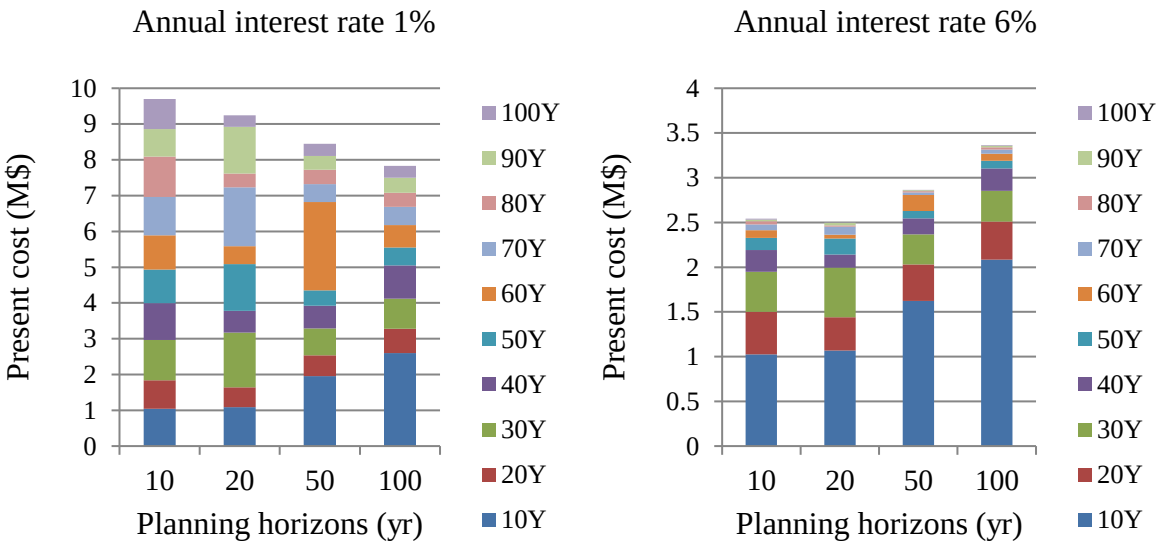


Figure 4.10. Present cost of pipes built at each time step for linear growth pattern.

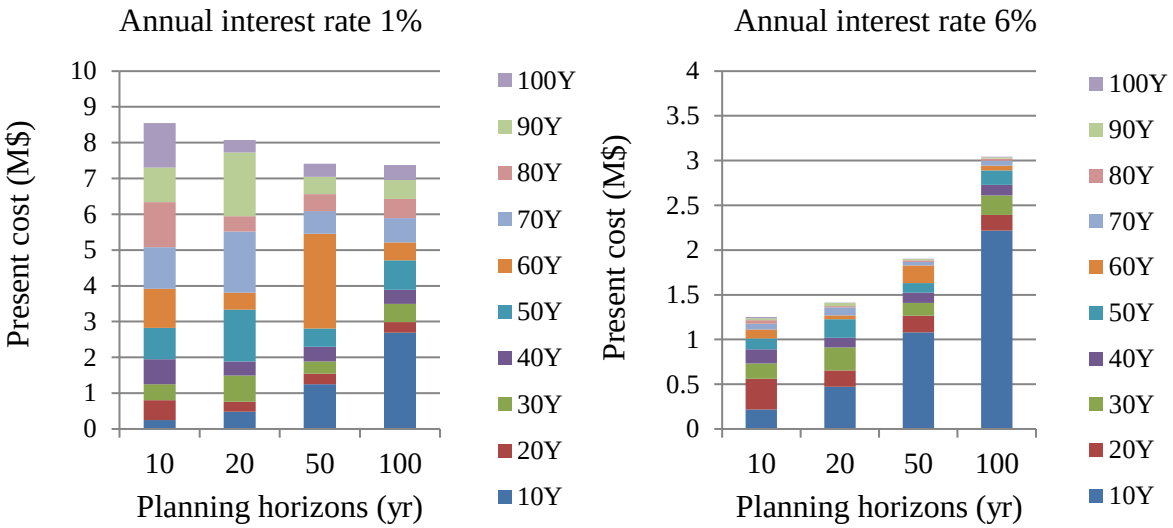


Figure 4.11. Present cost of pipes built at each time step for exponential growth pattern.

4.3.4 Conclusions

In this study, the impact of planning horizon on WDN design is analyzed through a synthetic city with single water reservoir under various annual interest rate scenarios with two population growth patterns (exponential and linear). The results indicate that the choice of optimal planning horizon is sensitive to the interest rate. For both growth patterns, shorter planning horizons are favored with higher annual interest rates while longer planning horizons are favored with lower rates as the balance between economies of scale and the excessive capacity cost induced by interest. Exponential growth pattern generally favors a shorter planning horizon than linear one due to more excess capacity provided at the beginning of study period. For networks with multiple sources, the model has to be constructed case by case to determine the optimum planning period.

REFERENCE

1. Nelson, A.C., *Toward a new metropolis: The opportunity to rebuild America*. Vol. 14. 2004: Brookings Institution Washington, DC.
2. ASCE, *2013 Report Card for America's Infrastructure*, 2013.
3. Ottino, J.M., *Engineering complex systems.* Nature, 2004. **427**(6973): p. 399-399.
4. Rouse, W.B., *Complex engineered, organizational and natural systems.* Systems Engineering, 2007. **10**(3): p. 260-271.
5. Barthélemy, M., *Spatial networks.* Physics reports, 2011. **499**(1): p. 1-101.
6. Watts, D.J. and S.H. Strogatz, *Collective dynamics of `small-world' networks.* Nature, 1998. **393**(6684): p. 440-442.
7. Barabási, A.L. and R. Albert, *Emergence of scaling in random networks.* Science, 1999. **286**(5439): p. 509-512.
8. Gastner, M.T. and M.E. Newman, *Shape and efficiency in spatial distribution networks.* Journal of Statistical Mechanics: Theory and Experiment, 2006. **2006**(01): p. P01015.
9. Yazdani, A. and P. Jeffrey, *Complex network analysis of water distribution systems.* Chaos: An Interdisciplinary Journal of Nonlinear Science, 2011. **21**(1): p. 016111-016111-10.
10. Csányi, G. and B. Szendrői, *Fractal–small-world dichotomy in real-world networks.* Physical Review E, 2004. **70**(1): p. 016122.
11. Ravasz, E. and A.L. Barabási, *Hierarchical organization in complex networks.* Physical Review E, 2003. **67**(2): p. 026112.

12. Erdős, P. and A. Rényi, *On Random Graphs I.* Publicationes Mathematicae, 1959. **6**: p. 290-297.

13. Fabrikant, A., E. Koutsoupias, and C.H. Papadimitriou, *Heuristically optimized trade-offs: A new paradigm for power laws in the Internet*, in *Automata, languages and programming*2002, Springer. p. 110-122.

14. Barthélemy, M. and A. Flammini, *Modeling urban street patterns.* Physical Review Letters, 2008. **100**(13): p. 138702.

15. Barthélemy, M. and A. Flammini, *Co-evolution of density and topology in a simple model of city formation.* Networks and Spatial Economics, 2009. **9**(3): p. 401-425.

16. Louf, R., P. Jensen, and M. Barthelemy, *Emergence of hierarchy in cost-driven growth of spatial networks.* Proceedings of the National Academy of Sciences, 2013. **110**(22): p. 8824-8829.

17. Moderl, M., T. Fetz, and W. Rauch, *Stochastic approach for performance evaluation regarding water distribution systems.* Water Science and Technology, 2007. **56**(9): p. 29-36.

18. Moderl, M., et al., *Systematic generation of virtual networks for water supply.* Water Resources Research, 2011. **47**.

19. Sitzenfrei, R., M. Möderl, and W. Rauch, *Graph-based approach for generating virtual water distribution systems in the software VIBe.* Water Science and Technology: Water Supply, 2010. **10**(6): p. 923-932.

20. Sitzenfrei, R., M. Möderl, and W. Rauch. *WDS Designer—A Tool Algorithmic Generation of Water Distribution Systems based on GIS Data*. in *Proceedings of the World Environmental & Water Resources Congress EWRI*. 2010.

21. Muranho, J., et al., *WaterNetGen: an EPANET extension for automatic water distribution network models generation and pipe sizing.* Water Science & Technology: Water Supply, 2012. **12**(1).

22. Mair, M., W. Rauch, and R. Sitzenfrei. *Spanning tree based algorithm for generating water distribution network sets by using street network data sets*. in *World Environmental & Water Resources Congress EWRI*. 2014.

23. Sitzenfrei, R., et al., *A multi-layer cellular automata approach for algorithmic generation of virtual case studies: VIBe.* Water Science and Technology, 2010. **61**(1): p. 37-45.

24. Sitzenfrei, R., et al., *Dynamic virtuai infrastructure benchmarking: DynaVIBe.* Water Science and Technology: Water Supply, 2010. **10**(4): p. 600.

25. Sitzenfrei, R., et al., *Modeling Dynamic Expansion of Water Distribution Systems for New Urban Developments.* Bridges, 2014. **10**: p. 9780784412312.320.

26. Swamee, P.K. and A.K. Sharma, *Design of water supply pipe networks*2008: John Wiley & Sons.

27. Methods, H., et al., *Advanced water distribution modeling and management*2003: Haestad press.

28. Bhave, P.R. and R. Gupta, *Analysis of water distribution networks*2006: Alpha Science Int'l Ltd.

29. Wagner, J., U. Shamir, and D. Marks, *Water Distribution Reliability: Analytical Methods.* Journal of Water Resources Planning and Management, 1988. **114**(3): p. 253-275.

30. Quimpo, R. and U. Shamsi, *Reliability-Based Distribution System Maintenance.* Journal of Water Resources Planning and Management, 1991. **117**(3): p. 321-339.

31. Ormsbee, L. and A. Kessler, *Optimal Upgrading of Hydraulic-Network Reliability.* Journal of Water Resources Planning and Management, 1990. **116**(6): p. 784-802.

32. Afshar, M., *Evaluation of selection algorithms for Simultaneous layout and pipe size optimization of water distribution networks.* Scientia Iranica, 2007. **14**(1): p. 23-32.

33. Afshar, M., M. Akbari, and M. Mariño, *Simultaneous layout and size optimization of water distribution networks: engineering approach.* Journal of Infrastructure Systems, 2005. **11**(4): p. 221-230.

34. Loganathan, G., H.D. SHERALI, and M.P. SHAH, *A two-phase network design heuristic for minimum cost water distribution systems under a reliability constraint.* Engineering Optimization+ A35, 1990. **15**(4): p. 311-336.

35. Saleh, S.H. and T.T. Tanyimboh, *Coupled topology and pipe size optimization of water distribution systems.* Water Resources Management, 2013. **27**(14): p. 4795-4814.

36. Soldi, D., A. Candelieri, and F. Archetti, *Resilience and Vulnerability in Urban Water Distribution Networks through Network Theory and Hydraulic Simulation.* Procedia Engineering, 2015. **119**: p. 1259-1268.

37. De Corte, A. and K. Sörensen, *HydroGen: an artificial water distribution network generator*. Water Resources Management, 2014. **28**(2): p. 333-350.

38. Falconer, K., *Fractal Geometry: Mathematical Foundations and Applications*. 2nd ed. Book2003: Wiley.

39. Longley, P.A., M. Batty, and J. Shepherd, *The size, shape and dimension of urban settlements.* Transactions of the Institute of British Geographers, 1991. **16**(1): p. 75-94.

40. De Keersmaecker, M.L., P. Frankhauser, and I. Thomas, *Using fractal dimensions for characterizing intra-urban diversity: The example of Brussels.* Geographical Analysis, 2003. **35**(4): p. 310-328.

41. Mandelbrot, B.B., *The Fractal Geometry of Nature*. 1st ed. Book1982: W. H. Freeman.

42. Longley, P.A. and M. Batty, *On the Fractal Measurement of Geographical Boundaries.* Geographical Analysis, 1989. **21**(1): p. 47-67.

43. Shen, G., *Fractal dimension and fractal growth of urbanized areas.* International Journal of Geographical Information Science, 2002. **16**(5): p. 419-437.

44. Frankhauser, P., *Comparing the morphology of urban patterns in Europe a fractal approach*. European Cities Insights on outskirts, Report COST Action 10 Urban Civil Engineering. Vol. 2. 2004. 79-105.

45. Tannier, C., et al., *A Fractal Approach to Identifying Urban Boundaries.* Geographical Analysis, 2011. **43**(2): p. 211-227.

46. Batty, M., *The size, scale, and shape of cities.* Science, 2008. **319**(5864): p. 769-771.

47. Batty, M. and Y.C. Xie, *Self-organized criticality and urban development.* Discrete Dynamics in Nature and Society, 1999. **3**(2-3): p. 109-124.

48. Frankhauser, P., *The fractal approach. A new tool for the spatial analysis of urban agglomerations.* Population, 1997. **52**(4): p. 1005-1040.

49. Benguigui, L., *A fractal analysis of the public transportation system of Paris.* Environment and Planning A, 1995. **27**(7): p. 1147-1161.

50. Benguigui, L. and M. Daoud, *Is the Suburban Railway System a Fractal?* Geographical Analysis, 1991. **23**(4): p. 362-368.

51. Domenech, A., *A topological phase transition between small-worlds and fractal scaling in urban railway transportation networks?* Physica a-Statistical Mechanics and Its Applications, 2009. **388**(21): p. 4658-4668.

52. Kim, K.S., L. Benguigui, and M. Marinov, *The fractal structure of Seoul's public transportation system.* Cities, 2003. **20**(1): p. 31-39.

53. Lu, Y.M. and J.M. Tang, *Fractal dimension of a transportation network and its relationship with urban growth: A study of the Dallas-Fort Worth area.* Environment and Planning B-Planning & Design, 2004. **31**(6): p. 895-911.

54. Rodin, V. and E. Rodina, *The Fractal Dimension of Tokyo's Streets.* Fractals, 2000. **8**(4): p. 413.

55. Thomson, J., *Great Cities and Their Traffic*, ed. Gollancz1977, London.

56. Labarbera, P. and R. Rosso, *On the fractal dimension of stream networks.* Water Resources Research, 1989. **25**(4): p. 735-741.

57. McAdams, M.A., *The application of fractal analysis and spatial technologies for urban analysis.* Journal of Applied Functional Analysis, 2009. **4**(4): p. 569-579.

58. Chai, L.H. and H.B. Li, *A new theoretical analysis on organizing principles of water supply networks.* Journal of Water Supply Research and Technology-Aqua, 2007. **56**(4): p. 233-244.

59. Kowalski, D., B. Kowalska, and M. Kwietniewski, *Monitoring of water distribution system effectiveness using fractal geometry.* Bulletin of the Polish Academy of Sciences Technical Sciences, 2015. **63**(1): p. 155-161.

60. Amit, R. and P. Ramachandran, *Optimal Design of Water Distribution Networks A Review.* 2009.

61. Waddell, P., *UrbanSim: Modeling urban development for land use, transportation, and environmental planning.* Journal of the American Planning Association, 2002. **68**(3): p. 297-314.

62. Silva, E.A. and K.C. Clarke, *Calibration of the SLEUTH urban growth model for Lisbon and Porto, Portugal.* Computers, Environment and Urban Systems, 2002. **26**(6): p. 525-552.

63. Bettencourt, L.M.A., *The origins of scaling in cities.* Science, 2013. **340**(6139): p. 1438-1441.

64. Bettencourt, L.M.A., et al., *Growth, innovation, scaling, and the pace of life in cities.* Proceedings of the National Academy of Sciences of the United States of America, 2007. **104**(17): p. 7301-7306.

65. Chen, Y., *Characterizing growth and form of fractal cities with allometric scaling exponents.* Discrete Dynamics in Nature and Society, 2010. **2010**.

66. Craig, J. and J. Haskey, *The relationships between the population, area, and density of urban areas.* Urban Studies, 1978. **15**(1): p. 101-107.

67. Pumain, D., *Scaling laws and urban systems.* Santa Fe Institute, Working Paper n 04-02, 2004. **2**: p. 26.

68. Marshall, J.D., *Urban land area and population growth: a new scaling relationship for metropolitan expansion.* Urban Studies, 2007. **44**(10): p. 1889-1904.

69. Walski, T.M. *Long-term water distribution design*. in *World Environmental and Water Resources Congress 2013@ Showcasing the Future*. 2013. ASCE.

70. Graham, R.L. and P. Hell, *On the history of the minimum spanning tree problem.* Annals of the History of Computing, 1985. **7**(1): p. 43-57.

71. Mair, M., et al. *Identifying multi utility network similarities*. in *Proceedings of the World Environmental and Water Resources Congress, Albuquerque, New Mexico*. 2012.

72. Masucci, A.P., K. Stanilov, and M. Batty, *Exploring the evolution of London's street network in the information space: A dual approach.* Physical Review E, 2014. **89**(1): p. 012805.

73. Batty, M. and P. Longley, *Fractal cities: A geometry of form and function*. 1st ed. Book1994, San Diego: Academic Press.

74. Yook, S.H., H. Jeong, and A.L. Barabási, *Modeling the Internet's large-scale topology.* Proceedings of the National Academy of Sciences, 2002. **99**(21): p. 13382-13386.

75. Qi, X., K. Li, and W.D. Potter, *Estimation of distribution algorithm enhanced particle swarm optimization for water distribution network optimization.* Frontiers of Environmental Science & Engineering, 2016. **10**(2): p. 341-351.

76. Zheng, F.F., A.R. Simpson, and A.C. Zecchin, *A combined NLP-differential evolution algorithm approach for the optimization of looped water distribution systems.* Water Resources Research, 2011. **47**.

77. Rossman, L.A., *EPANET users manual*, 1994, U.S. Environmental Protection Agency: Cincinnati, Ohio.

78. Clark, R., et al., *Cost Models for Water Supply Distribution Systems.* Journal of Water Resources Planning and Management, 2002. **128**(5): p. 312-321.

79. Shih, J.-S., *Economies of scale and technical efficiency in community water systems*2004, Washington, DC: Resources for the Future.

80. Dharmaratna, D. and J. Parasnis, *An analysis of the cost structure of water supply in Sri Lanka.* Journal of the Asia Pacific Economy, 2012. **17**(2): p. 298-314.

81. Filippini, M., N. Hrovatin, and J. Zoric, *Cost efficiency of Slovenian water distribution utilities: an application of stochastic frontier methods.* Journal of Productivity Analysis, 2008. **29**(2): p. 169-182.

82. Garcia, S. and A. Thomas, *The structure of municipal water supply costs: Application to a panel of French local communities.* Journal of Productivity Analysis, 2001. **16**(1): p. 5-29.

83. Kim, H.Y. and R.M. Clark, *Economies of scale and scope in water supply.* Regional Science and Urban Economics, 1988. **18**(4): p. 479-502.

84. Nauges, C. and C. van den Berg, *Economies of density, scale and scope in the water supply and sewerage sector: a study of four developing and transition economies.* Journal of Regulatory Economics, 2008. **34**(2): p. 144-163.

85. Stone and Webster, *Investigation into evidence for economies of scale in the water and sewerage industry in England and Wales*, 2004: London.

86. Weber, W.J. and J.W. Norton, *Financial and Technological Analysis of Water Treatment Technology Implementation Using Distributed Optimal Technology Network (DOT Net) Concepts.* Ann Arbor, 2008. **1001**: p. 48109-2099.

87. Chung, G., et al., *A general water supply planning model: Evaluation of decentralized treatment.* Environmental Modelling & Software, 2008. **23**(7): p. 893-905.

88. Mizutani, F. and T. Urakami, *Identifying network density and scale economies for Japanese water supply organizations.* Papers in Regional Science, 2001. **80**(2): p. 211-230.

89. Sauer, J., *Economies of scale and firm size optimum in rural water supply.* Water Resources Research, 2005. **41**(11): p. n/a-n/a.

90. Shammas, N.K. and L.K. Wang, *Water supply and wastewater removal*, 2011, John Wiley & Sons: Chichester; UK.

91. McGivney, W. and S. Kawamura, *Cost Estimating Manual for Water Treatment Facilities*2008: John R. Wiley & Sons.

92. Kawamura, S. and W. McGivney, *Cost Estimating manual for water treatment facilities*, 2008, Wiley: Hoboken, NJ, USA.

93. Yazdani, A. and P. Jeffrey, *Applying network theory to quantify the redundancy and structural robustness of water distribution systems.* Journal of Water Resources Planning and Management, 2011. **138**(2): p. 153-161.

94. Hsu, N.S., et al., *Optimization and capacity expansion of a water distribution system.* Advances in Water Resources, 2008. **31**(5): p. 776-786.

95. Freidenfelds, J., *Capacity expansion: analysis of simple models with applications*1981, New York: North-Holland.

96. Luss, H., *Operations research and capacity expansion problems: A survey.* Operations research, 1982. **30**(5): p. 907-947.

97. Manne, A.S., *Capacity expansion and probabilistic growth.* Econometrica: Journal of the Econometric Society, 1961: p. 632-649.

98. Creaco, E., M. Franchini, and T.M. Walski, *Accounting for phasing of construction within the design of water distribution networks.* Journal of Water Resources Planning and Management, 2013. **140**(5): p. 598-606.

99. Creaco, E., M. Franchini, and T.M. Walski, *Taking account of uncertainty in demand growth when phasing the construction of a water distribution network.* Journal of Water Resources Planning and Management, 2014. **141**(2): p. 04014049.

100. Martin, Q.W., *Linear water-supply pipeline capacity expansion model.* Journal of Hydraulic Engineering, 1990. **116**(5): p. 675-690.

101. Mortazavi, M., G. Kuczera, and L. Cui, *Application of multi objective optimization for managing urban drought security in the presence of population growth| NOVA. The University of Newcastle's Digital Repository*. 2012.

102. Scarato, R.F., *Time-capacity expansion of urban water systems.* Water Resources Research, 1969. **5**(5): p. 929-936.