A GIS-BASED DYNAMIC MODELLING OF WATER DISTRIBUTION NETWORK

by

TAUFIQ RASHID

(Under the Direction of Ke Li)

ABSTRACT

With the aging of water infrastructure systems and the challenge of building resilient and sustainable infrastructures for future, renovating water distribution networks (WDN) using holistic models is necessary. The understanding of infrastructure systems as complex adaptive system can be used as a framework to guide network master planning and engineering practices. A theoretical complex network model has been created following the combination of local optimization rules and engineering considerations. The demand nodes were generated dynamically following the scaling law of urban growth and considering geographic restrictions. To model the growth of real-world water distribution networks, Geographic Information System (GIS) was integrated into the complex network model. The model was used to develop WDN for Athens Clarke County and the result was compared with real-world WDN's reported by other literature. In general, the comparison showed that the structural properties of the generated WDN agree with the real-world WDNs.

INDEX WORDS: Water distribution network, Complex network, Geographic Information System, Dynamic Model

A GIS-BASED DYNAMIC MODELLING OF WATER DISTRIBUTION NETWORK

by

TAUFIQ RASHID

B.Sc., Khulna University of Engineering and Technology, Bangladesh, 2015

A Thesis Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment

of the Requirements for the Degree

MASTER OF SCIENCE

ATHENS, GEORGIA

2018

A GIS-BASED DYNAMIC MODELLING OF WATER DISTRIBUTION NETWORK

by

TAUFIQ RASHID

Major Professor: Ke Li
Committee: Lan Mu
Caner Kazanci

Electronic Version Approved:

Suzanne Barbour
Dean of the Graduate School
The University of Georgia
August 2018

DEDICATION

I dedicate this research to my sister and all other individuals who helped me tremendously to come to the United States for graduate school. Without their support, this journey would not be possible.

ACKNOWLEDGEMENTS

First of all, my gratitude goes to my parents and other family members for their lifelong support.

My sincere regards go to my major professor, Dr. Ke Li whose continuous guidance was imperative for this research. He was always very generous with his time and efforts to guide me on how to solve research problems. Whenever I faced problems with research, rather than just providing guidelines on how to solve it, he tried to solve the problem along with me. He was always very appreciative even with my smaller efforts. His thought towards religion and afterlife also inspired me a lot. The experience of working under his guidance for the past two years will be highly beneficial for my future career.

I also want to highly acknowledge Dr. Lan Mu for her continuous support in the last two years. She introduced me to the world of GIS and it was through her I came to learn about this exciting area of study and eventually fell in love with it.

Moreover, I would like to thank Dr. Caner Kazanci for joining my thesis committee and providing his insightful suggestions and encouragements. I extend my appreciation to Dr. K.C. Das for his advice and support to build my professional career. I also want to thank my friends, Emad Ahmed and Kadeghe Fue for their continuous support for my research and coursework.

Finally, I want to acknowledge my friends Kishan Mahmud and Nadia Noor for their support and motivation during my tough time with research. They provided me the courage to remain persistent and focused on solving the research questions.

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1

INTRODUCTION

1.1 Problem Overview

Water distribution system analysis is a vital part of water utility management program. Water distribution systems are susceptible to infrastructure growing older and worsening, water quality degradation, and capacity deficiencies. The requirement to restore, change, and repair drinking water distribution systems are becoming a huge concern in the United States. One of the key investment sector for municipalities include water distribution system and wastewater collection systems [1]. According to the American Water Works Association (AWWA), more than $ 1 trillion will be required to improve and repair the current water infrastructure system in the United States in the next 25 years [2]. Water distribution networks (WDN) are expensive system to construct, run and manage. Piping system accounts for most of the cost of building a water distribution system. The capital cost and expense required for maintaining WDN are large which make it inevitable to devise a suitable plan before constructing or repairing the network. The operating cost can be reduced, and system performance can be improved by modeling the network [3].

Water distribution systems are complex network of many interacting components such as pipes, pumps, reservoirs, valves, and other accessories. Due to the presence of many interacting components, this kind of infrastructure system produce result as “the whole is greater than the sum of the parts”, which makes them resistant to study using the traditional “divide and conquer” engineering approach. On the other hand, the complex topology and adaptive behavior of such infrastructure systems are driven by both self-organization of the demand and rigid engineering solutions. Hence, a balanced holism and reductionist methodology are required to understand complex water infrastructure system [4]. This requires predictability of the complex network topology, complex behavior, and their evolution dynamics over time with organic integration of engineering rules. Complex system studies put more emphasis on examining evolving behavior and configuration at long-term and large scale to find probable description of them. The short-term specifics of engineering are not considered which results in difference from real networks. The engineering models, however, put more emphasis on short-term (e.g. in planning period of 20 to 30 years) and unable to show the long-term emergence. To simulate the growth of WDNs, a complex network model has been created by our research group following the network efficiency optimization objectives and engineering criteria of loop formation. The demand nodes were projected dynamically using the scaling law of urban growth. The evolvement of the network was directed by local optimization rule [5]. The proposed model can generate WDN similar to reported real-world WDNs on some structural properties. However, the proposed model does not follow the constraint of geographic policy and restriction. Obstacles such as water bodies, buildings etc. can significantly affect the modeling of WDN. For example, the presence of a waterbody on the path of a proposed network means that the pipes cannot pass through the waterbody. So, the pipes need to be modeled to pass in an alternative way. To model the growth of real-world WDN using complex network theory, the network has to expand following the boundary of geographic limitation.

There has been significant improvement in modeling water distribution system in the past 30 years. The modeling practice has transformed from time-consuming hand calculation exercise to thorough demonstrations of networks containing numerous pipes. Water distribution simulation models have been effectively integrated with Geographic Information System (GIS) to aid data input and output and to save and recover information more efficiently and easily [3]. GIS is turning out to be a very important tool for water distribution modeler in terms of a decision support tool and as a source for modeling data [6]. In this study, the complex network model was integrated into GIS so that all the geographic policy and restriction can be considered while generating the WDN. The proposed model can generate WDN for any cities in the United States. The integration of the model into GIS will also provide additional benefits such as finding suitable locations for water system facility sites. Since GIS can store a large amount of data, so ideal values for model parameters such as pipe length, pipe diameter etc. can be easily calculated. To add the geographic restriction, the shapefile of the target study area was imported into ArcGIS. Then a new layer was created from the existing shapefile to include only those features through which the water network can pass. Then the modified shapefile was imported into MATLAB so that the demand node generation can be governed by geographic consideration. The model can generate WDN only within the boundary provided by the shapefile. In this way, the complex network model can generate realistic WDN following the constraint of geographic policy and restriction.

1.2 Objective

The overall objective of this study is to create a model which will be able to generate real-world water distribution network following the complex network approach. This will be achieved through the following steps:

- Select a study area for which the WDN will be generated.
- Modify the study area to exclude areas which fall under geographic constraints for water distribution network.
- Import the modified study area to the complex network model.

- Modify the model so that it can use the study area as a reference while generating water distribution network
- Run the model to generate a water distribution network for the selected study area.
- Compare the generated water distribution network with real-world network.

CHAPTER 2

LITERATURE REVIEW

2.1 Complex System

We are living in a progressively interconnected world of techno-social systems where the organizations are made up of various technological layers which are interacting within the social component that guide their function and growth [7]. So, it is necessary to increase our attention towards the dynamic character and multi-level organization of events. A complex system is made of many elements and the behavior of the system cannot be simply deduced from the behavior of its elements. The degree of complexity depends on the amount of information required to explain the behavior of such a system [8]. Sometimes it is hard to make sense of complex system because it requires us to think hypothetically and often contradicts with the traditional norms and principle [9]. The association across different levels in complex system are not easily understandable. For instance, to study about ecological systems, one needs to think about how genes, individuals, populations, and species are interconnected. An ecosystem can be observed from the point of the specific organism to the point of the environment as a whole [10]. The skill to comprehend and tackle complex systems require consideration of many events which is hard to understand for traditional engineering science communities [4]. According to the National Academy of Engineering, engineers in the U.S. require a broader education system which would help them to understand complex system [11]. Due to the interaction of many phenomena, infrastructure systems generate results as “the whole is greater than the sum of the parts”. So, it is not possible to study them using the traditional “divide and conquer” engineering approach. However, the complexity topology and adaptive behavior of infrastructure systems are determined by both self-organization of the demand and stiff engineering solutions. So, a balance of holism and reductionism is required to understand complex system [4].

Many networks in real world are complex networks whose structure is dynamic and complex. It was studied in 1959, when random graph theory was proposed [12]. Random graph theory allows researchers to empirically study the structural properties of complex networks. It is used to analyze families of networks by combining graph theory with probability theory. There were some limitations of this graph theory. Some topological properties of many real-world complex networks which are not completely regular and completely random, could not be explained by using this random graph theory. Barabási–Albert model [13] was a notable complex model produced after the random graph theory. It was a dynamic model which can generate scale-free networks. Some dynamic models are built to study the growth of spatial complex networks. These models are based on the assumption that every new edge is created to efficiently connect the new node to the existing network. Some global structural properties of these models can be reproduced. Fabrikant et al. [14] proposed a simple model of Internet growth which under very general assumptions and parameter values results in power-law-distributed degrees. In their model, 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). Their results suggested that power law degree distribution of internet is perhaps the manifestation of trade-offs, the degrees obey power law when α is not too small or too large. If α is small, a star structure is produced. If α grows at least as fast as$\sqrt{n}$, this model will generate a dynamic version of the Euclidean minimum spanning tree (MST). Gastner and Newman [5] analyzed spatial distribution or collection networks such as pipelines and sewers, emphasizing on their cost in terms of total edge length and their efficiency in terms of the network distance between vertices, as measured by the route factor. They introduced two models for spatial distribution network. In those two models, the positions of all nodes are given at the outset, rather than randomly add to the network one by one over time. They emphasized on the cost of a network, which is represented by the total length of all its edges, and its efficiency in terms of the directness of routes from point to point. 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 relatively straight trunk lines and short branches. This model is mostly unrealistic because many vertices are never joined to the network, even if they are quite close to the root which happens due to the preference of straight path. The network generated by the second model produces a symmetric network that fills out the space with some approximately constant from the root because it gives preference to shortest paths irrespective of shapes. A more realistic growth model has been developed for street network with loops by Barthélemy and Flammini [15] where the set of nodes and the network of links that connects them grow simultaneously. The model algorithm generates new sections of roads based on local optimization and tries to connect the nodes using shortest possible path. 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:

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

The maximization of Δ is done under the 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 statistical properties of this model match those found in empirical studies which shows that local optimization is a potential method for the growth of street networks. Barthélemy and Flammini [16] developed a study to model the connection between the evolution of the topology of street networks and the population density. It is based on the fact that the population density decreases exponentially from a core district. Accessibility and house price were the two economical mechanisms used to govern the individual choice of new location.

2.2 Water Distribution System

A water distribution network (WDN) is an interrelated collection of sources, pipes, and hydraulic control elements (e.g., pumps, valves, regulators, tanks) supplying recommended water quantities at preferred pressures and water qualities to consumers [17]. Even though the size and complexity of water distribution systems differ significantly, the main objective for all types of WDNs are to supply water from the source to the customer. Transporting water from source to customer requires a network of pipes, pumps, valves and other accessories. There are two types of piping system: transmission mains and distribution mains. Transmission mains are used to deliver large amounts of water over long distances. These are usually used to carry water from a treatment facility to a storage tanks throughout several cities. Distribution mains follow the general topology and alignment of road networks in cities. They are smaller in diameter than the transmission mains and are used to supply water to individual customers [6]. Storage features such as tanks and reservoirs are needed for storing water to allow for variation in water demand. Piping network, storage, and the supplementary framework are together denoted as the water distribution system. To analyze water distribution systems, pipes and other connections are represented by edges, with the fixed junctions (reservoirs, tanks, and pipe intersections) shown by nodes. The structure of pipeline network can be either branched or looped (Figure 2.1).

Figure 2.1. Structure of a branched network (left) and a looped network (right).

Having a tree-like structure, a branched network has only one distinctive route to supply water from the source to a demand node. The benefit of this type of network is that water is efficiently supplied to customers with low capital costs. However, if one pipe stops working, downstream consumers would be disconnected from the source which makes branched network very unreliable. In contrast, the presence of redundant pipes makes a looped network more reliable. Water can be supplied to each customer via multiple paths in a completely looped network. So, the failure of one pipe won't disconnect all the consumers from the source. Higher supply capacity and lowered water age also make looped network more beneficial than branched network [18]. The downside of such network is that they require more initial investment. The ideal design of a real-world WDN is a complex problem where the designers have to tradeoff between cost minimization and network reliability [19]. So, most of the real-world WDNs are partially looped [6]. In low-density rural areas, branched networks are usually used while looped structures are used for the urban area [20].

2.3 Complex System in Water Distribution Network

A water distribution system is a complex system consisting of many interacted components. Various interconnected sections in WDN are organized in nontrivial arrangement and interact in complex manner. The potential of using a range of probable combinations of pipe sizes, materials and connectivity layouts, and factors such as the position of the valves and pumping stations, pump scheduling, and indefinite demand for water are the primary reasons for high complexity and uncertainty in model and management of WDN [19]. To date, only small world phenomenon in WDNs is studied by using complex network approaches in water supply systems [21]. There was no systematic analysis of WDN structure and function by using complex network approach until done by Yazdani and Jeffrey [19]. They analyzed the structural distribution of water distribution systems and compared them to similar types of physical networks. The empirical data for four standard water distribution networks was studied by calculating the structure of the network paths cycles, connectivity, and efficiency. Similar method was employed by other researchers to measure the structure of WDN in their study [22, 23]. In complex network approach, a water distribution network can be represented as a graph G = G (N, E) where N is the set of nodes and E is the set of edges connecting the nodes [19]. WDNs are modeled as undirected graphs because the direction of water flow in the pipes change occasionally [23, 42]. WDNs can be approximately modeled as planar graphs because it is impractical to lay pipes in multiple layers on top of each other [19]. 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 Yazdani and Jeffery [19] (Figure 2.2), Balcksburg (USA), Fossolo (Italy), Pescara (Italy) and Modena (Italy) reported by De Corte and Sorensen [23,42] (Figure 2.3) are listed in table 2.1.

Figure 2.2. Real-world water distribution networks reported by Yazdani and Jeffery [9]: East-Mersea (top-left), Colorado Springs (top-right), Richmond (bottom-left) and Kumasi (bottom-right).

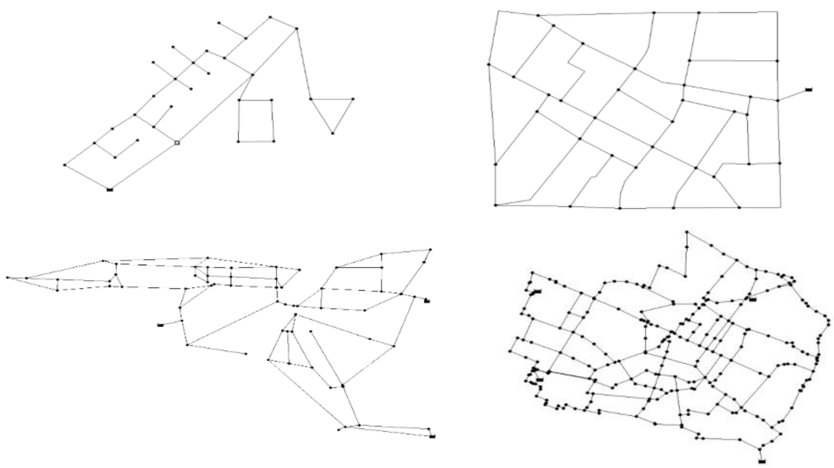


Figure 2.3. Real-world water distribution networks reported by De Corte and Sorensen [23, 42]: Blacksburg (top-left), Fossolo (top-right), Pescara (bottom-left) and Modena (bottom-right).

Table 2.1. 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 [19].
[b]Data from De Corte and Sörensen's work [23,42].

Several metrics are used by researchers to explain the structural formation of WDN. The similarity to or deviation of the structure of WDN from treelike or mesh graphs can be defined by calculating edge-per-node ratio (m/n). This is the ratio of number of edges to number of nodes. The values of edge-per-node ratio are larger than 1 but much smaller than 2 for real networks as shown in table 2.1. The sparseness/denseness of the network connections can be determined by evaluating edge density (e). This is the ratio of the number of total existing edges to the maximum possible edges. WDNs are sparse networks which is evident from the low edge density values of all WDNs. Average node degree ($<k>$) represents the number of edges connected to a node. Real

WDNs are usually less connected than a grid network. The values of average node degree for those real networks are basically larger than 2 but much smaller than 4. 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 average geodesic distance between all node pairs in WDN can be determined by using characteristic path length (l) and the maximum geodesic distance can be calculated by using diameter (g_{max}). Geodesic distance is the number of edges in a shortest path. The geodesic distances between many node pairs of the four large real WDNs reported in Yazdani and Jeffrey's research [19] are large, and the diameter and characteristic path length are also large. Large geodesic distances between nodes mean that WDNs are not small-world networks [58]. Central-point-dominance (c_b) shows how the network flow is controlled by centrally located point(s). It is basically the average difference in relative betweenness centrality of the most central point and all other nodes. Betweenness centrality of a node evaluates the significance of a node based on its presence on the number of shortest paths. It is measured as the number of shortest geodesic paths between two given vertices that pass through that node divided by the total number of shortest geodesic paths between those two vertices. The value of central-point-dominance of real networks are smaller than 1 and larger than 0. Thus, there is no huge difference on node importance for WDNs. The efficiency of a WDN can be measured based on the connectivity between a root node and other nodes in the network by calculating route factor (q). It is the average ratio of Euclidean distance along the shortest path from a node to the root to its direct Euclidean distance to the root. Smallest possible value of route factor results in a star structure where all nodes are directly connected to the source. In real networks, the nodes are directly connected to each other because the route factor are larger than 1. The density of the cycles and loops in planar graphs can be measured by using meshedness coefficient (r_m). It is defined as the ratio of the number of total existing independent loops to the maximum possible loops. Algebraic connectivity and spectral gap were measured to distinguish between structural vulnerability and fault-tolerance of WDN. Algebraic connectivity (λ_2) is the second smallest eigenvalue of the normalized Laplacian matrix of the network. This metric is used to quantify the robustness and connectedness of the network against efforts to decouple parts of the network. Spectral gap ($\Delta\lambda$) is the difference between the first and second eigenvalues of the adjacent matrix of the network. Small spectral gap may result in low sparseness, connectivity and the presence of bottlenecks whose removal results in the split of the network into two or more large sections. In general, complex network metrics provide a basis for the analysis of water distribution systems and compare between WDNs and other networks, in terms of their structure, organization, network efficiency, redundancy and vulnerability.

2.4 Geographic Information System in Water Distribution Network

Geographic Information Systems (GIS) is becoming an important tool for both modeling and managing water distribution network. GIS lets us visualize, question, analyze, and interpret data to understand relationships, patterns, and trends [25]. Due to the growing potential of GIS, it is now possible to visualize, and model the entire cycle of water supply network from source to household [26]. Nearly 15 percent of water utilities are presently using GIS in their modeling and about 80 percent are planning to use GIS in the future [27]. GIS-based database management system (GIS-DBMS) is used in modeling, designing as well as management of water supply and sewer systems. The utility companies, especially water companies are using GIS to manage and assess their resources [28]. More than 80% of all information used by a water utility is geographically referenced [29]. GIS allows to link different social, economic and environmental factors connected to spatial entities of a water resources problem and they can be utilized in a decision-making process. In this way, GIS can incorporate spatial dimensions into the conventional water resource database and display an integrated view of the world [30].

Initial approaches of using GIS in water distribution network was limited to only visualizing the output of water resource models. However, predictive and associated critical abilities are more vital for designing and managing complex water distribution network [31]. GIS can combine spatial relationships and databases which allows a planner to superimpose soil data, repair data, and hydraulic modeling output to automatically allocate a condition rating to pipes. In GIS, non-spatial information about features is linked as attributes which means features in WDN such as pipes, tanks etc. can contain attributes. For example, the diameter of a pipe can be added as an attribute for that feature. GIS also offers a wide range of options for thematic mapping. Basic thematic mapping can be achieved by most water distribution modeling software but when it comes to visualization, GIS has better tools in its inventory. A good visualization helps in communicating the results of modeling to managers, regulators, the media, and the general public. If a WDN model is built using GIS, it can be modified in the future if needed by using latest GIS data. GIS can be used to find suitable locations for water system facility sites. It can also be used to identify prospective places for monitoring equipment [6].

Ayad et al. [32] used GIS as a tool to create a number of applications for water distribution network management. They performed tasks such as detecting valves to be closed in case of pipe break, service area for treatment plants, and network skeletonization using GIS. They also emphasized on the graphical side of GIS by displaying their results from hydraulic analysis and optimization models in GIS. Shamsi [33] discussed three methods of developing GIS-based modeling applications: interchange method, interface method, and integration method. No direct association between the GIS and the model exists in the interchange method and they run independently. Data is written to an intermediary file, where it is reformatted for the model if required. Then the data is read into the model. In the interface method, there is direct connection between GIS and the model. These connections are used to synchronize the model and the GIS. It consist of a preprocessor and a post processor. The preprocessor examines and exports the GIS data to model input files and the post-processor imports the model output and display it as a GIS layer. The integration method implies that both the GIS and the model are integrated into a single program. The integration of GIS and hydraulic model saves model building time. Zhang [34] integrated GIS and CARE-W to improve the management and restoration of water distribution networks. Computer Aided Rehabilitation on Water Network (CARE-W) is a research project whose purpose is to provide the most cost-efficient system of maintenance and repair of water distribution networks. Its objective is to ensure the security of water supply that meets social, health, economic and environmental requirements. The integration of GIS and CARE-W would provide the utility managers consistent and scientific support on water distribution network management and rehabilitation.

GIS maps may contain multiple features, each of which are displayed as a layer on the map. The appearance of the resulting map can be controlled by selecting which layers are displayed and the symbology of each features [6]. There are three types of method for geographic data representation: raster, vector and triangulated irregular networks (TIN). In raster representation, data is stored as discrete grids in which a single value for each attribute is associated with each grid cell. Each grid cell has an attribute associated with each grid cell. In vector representation, discrete features are stored as points, lines, or polygons. Each feature is linked to attribute values and a set of x-y coordinates [35]. Vector data are commonly used in hydraulic modeling. In GIS, nodes are represented as points, pipes as lines, and node service areas as polygons. Raster and TIN data are also useful in certain cases such as extracting elevation data or using an aerial photo as a background for the model. GIS database design for water utility involves three primary three major steps: (1) The water distribution facilities need to be represented cartographically, (2) An inventory of the network needs to be created, (3) The network need to be modeled using GIS. Creation of metadata is also an important step in database design process which is basically the compilation of information about the dataset. The facility to share data between organization over internet using eXtensible Markup Language (XML) makes metadata an important part in database design process [6].

Eljamassi and Abeaid [36] showed how the GIS technology could be used to solve water distribution problem in Rafah city. They demonstrated how equal distribution of water to all zones could be achieved by using one of the zone as an example. In order to use GIS as the decision-making tool, the proposed system was included in the GIS software. A shapefile was created which contained features such as pipes, street, tank, manhole, land use, water distribution zone etc. Characteristics of the network such as pipe length, pipe diameter, water demand etc. was added to the feature classes as attributes. The main advantage of building the water distribution network in GIS was because of its efficient data analyzing tools. Guth and Klingel [37] developed a GIS tool to allocate demand in water distribution network model. The demand allocation involved two stages. The first step was the calculation of the catchment areas of the nodes and the second step was the determination of the water demand within the catchment area and its allocation to the corresponding node. The application was used as a dynamically linked library (DLL) which was integrated into ArcMap as a toolbar. Oliveira et al. [38] used GIS data in order to predict breakage rates and remaining service life of pipes taking into consideration the spatial interdependencies of breaks points. They studied the efficacy of GIS spatial analysis capabilities for the measurement of failure patterns and trends in water distribution networks. The existence of clusters in the water main breaks dataset supported the concept of the spatial dependences in the breakage phenomena. McKinney and Cai [39] applied an object-oriented method to create a tight connection between a GIS and a water resources management model for ideal water allocation in river basins. They prepared a conceptual model that combine data, model and user interactions in a GIS environment. They also created a prototype linkage between a GIS and a water resources management model. Pachri et al. [40] developed spatial data of hydrological modeling in Chitchik river basin. GIS tools were used to create the geo-spatial data for hydrological modeling. Gavekar and Nandavekar [26] prepared geographical layers for water distribution network. They created GIS layers for pipeline, valves, population density and contour maps.

Building a water network model is a lengthy, expensive and vulnerable step in a hydraulic modeling project. It was a specialized task for engineers until GIS has emerged as a suitable platform for hydraulic modeling. Engineers created model input files by collecting, joining and digitizing data from a range of hard copy source documents, such as water system maps, topographic maps and census maps. GIS can be used to create cheaper and more accurate model than the one created by an engineer for a number of reasons. The model created by GIS can contain a lot of details since GIS can manage large volumes of data. GIS can automate certain tasks which makes model building quicker and more useful. If the customer billing records are georeferenced, then they can be used to produce and assign water demands for the model. Some features of hydraulic modeling must be considered when GIS for this purpose. These features of the modeling are: network granularity, scenarios, time-series data, and ownership. In most of the cases of hydraulic modeling applications, every pipe in the actual system does not need to be included in the model network to obtain accurate results. A skeletonized version of the model good enough to make informed planning decisions. The skeletonized version of the model also improves the efficiency of the hydraulic modeling software. Different scenarios need to be considered in water system planning such as future demands, proposed pipes and system facilities, or facility outages for emergency response planning. GIS can be planned to include the various what-if conditions and phases of a water system facility. Hydraulic models and water quality models are usually dynamic so that they can be used to predict the response of the water system over an extended period of time. So, the GIS model needs to be designed to handle time-series data. In an enterprise-level GIS, most of the data is typically not 'owned' by the hydraulic modeler which raises several issues. So, the modeler needs to be in control of the data used for modeling [6].

CHAPTER 3

METHODOLOGY

3.1 Model Development for Complex Network

WDN is a complex system which has to develop over time to satisfy the increasing demand of population. The population demand is in a self-evolving manner. The social-economic environment affects the development and maintenance of the WDN at a variety of temporal and spatial scales. The dynamic growth of water demand (nodes) and the evolving of the pipeline networks (edges) need to be considered while modeling the complex behavior of WDN. The complex network model used in this study has two sections. The first section simulates the generation of new water demand based on urban development while the growth of WDN due to the new water demand is simulated in the second section. For a given study area, the model generates several new nodes at each time steps. Then the new nodes are connected to the existing network through edges forming a tree structure. The reliability of the network can also be increased by producing a looped network through extra edges. The generation of demand node itself is a complex process which makes it difficult to model real engineered complex systems. The topology of the water demand nodes follows the expansion and distribution of population in the city. Power law can be used to describe the expansion of a city. The distribution of any urban forms can be illustrated by using the following scaling function of the population (P) over the urban area (A) [41, 42]:

$$A \propto P^{\alpha} \tag{3.1}$$

where α is the scaling factor illustrating how urban area grows with population. The coefficient α varies as the location and time changes. This spatio-temporal change reflects different urban forms that are limited by the spatial and economic factors [43-46]. The water demand nodes generated in this study are developed following this scaling function. The demand node generation starts with one root node representing the reservoir located at the center of the given area. At each time step, multiple water demand nodes would be added at random following the scaling function. The number of nodes generated is proportional to the population if water demand of each node is assumed to be identical. Hence, new nodes are added with side length 2R, where R is measured by:

$$R \propto N^{\alpha/2} \tag{3.2}$$

where N is the total number of water demand nodes. The side length increases with total number of water demand nodes. It simulates the sprawling of urban area. The number of initial water demand nodes is N_0 with side length $2R_0$. The population growth rate governs the number of new water demand nodes generated at each time step. Based on the data from U.S. census bureau, the average yearly population growth rate in the United States was 1.3% in the 20th century. The growth of WDN is usually modeled for about 20 years [47]. So, the expansion of new water demand nodes is set to be 30% for each time frame in this study. Unrealistic demand node may be generated because of in-fill development when two nodes are too close to each other. To prevent this, an engineering constraint is introduced as the distance threshold r_0. If two nodes are produced within this threshold, they will be merged into one. If a new node is generated within the distance r_0 to an existing node in the network, the new node would be removed. The existing node is taken as a connected new water demand node.

After generating the nodes, the edges need to be generated to connect the nodes. The optimization of the network can be achieved by considering both the total length of pipes and shortest path distances from the root to demand nodes. An increase in total length of pipes increases the total capital cost. The capital cost can be minimized by using a minimum spanning tree (MST) [48] which produces a star structure that connects each node directly to the root. The shortest path distance stands for operational energy demand and water age. A tradeoff between the total length of pipes and shortest path distance is usually considered in real-world network. Gastner and Newman [5] proposed a growth model with local minimization process considering both total length of edges and shortest path distances from root node to other nodes. The edge generation mechanism used in this study is designed from the local minimization process developed by Gastner and Newman [5]. They considered both total length of edges and shortest path distances from root node to other nodes in their spatial distribution network model. In our model, an unconnected new water demand node *i* is connected to an edge in the existing network following the objective function:

$$min(d_{ij} + \beta d_{j0}) \tag{3.3}$$

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. β can be used to control whether the network will be a minimum spanning tree or a star structure. The higher the value of β, the more the possibility of creating a star structure. For lower values of β, the generated network will be a tree structure.

In real world, water distribution system may suffer from disconnection or disorder. The edge generation process mentioned above will produce a tree structure which might lead to a break down in the whole network in case there is a problem at some part of the network. So, such kind of network is not good in terms of reliability. Real world WDN contain loops which increases the reliability of the network. Another optimization mechanism is used in our model to take care of this issue. For the new water demand node i, an extra edge can be added to connect it to another edge through node k by minimizing the distance:

$$min(d_{ik}) \tag{3.4}$$

WDNs are usually planar network [19], so the extra edge shall not be intersecting with other edges. Network graph and engineering rules are used to govern the generation of edge to form a looped structure. The distance between node k and node i should be less than a given maximum distance d_{max}. The value of the parameter d_{max} can be used to control the amount of looping in the network. A higher value of d_{max} will increase the level of looping in the network. Loops in WDNs are mostly quadrilateral [19]. So, the new node i should be restricted from directly connecting to one edge repeatedly which might result in a triangular cycle. Hence, the geodesic distance between node k and new node i should be larger than 2. A real-world WDN is usually less connected than a grid network [23]. So, the degree of new node i shall be lower than 3. Tree-structured networks usually have a lower degree than the looped network. The angle between the incident edge and extra edges is an important factor in edge generation. This angle is measured by using the parameter θ. The arrangement of WDNs follows certain street pattern [49] and in practice, the angle between two roads are not usually very small. So, θ should be greater than certain minimum angle θ_{min} and θ_{min} is set to 45° in our study to obtain a more realistic street pattern [50].

3.2 Integration of GIS into the model

The integration of GIS into the model involves three primary steps: (1) processing the shapefile of the study area in ArcMap; (2) importing the shapefile into MATLAB; (3) modifying the model in MATLAB so that it can read the shapefile and draw the water distribution network within the boundary provided by the shapefile. First of all, the shapefile of the study area is collected from an available data source. There are many data source for GIS shapefiles such as: U.S. Census, USGS, Geospatial Data Gateway etc. Then, all the features in the study area need to be aggregated into a single feature. The dissolve tool in ArcMap is used to merge all the features. ArcMap has an editor toolbar which allows editing the features of any layer. This toolbar is used to draw restricted areas on the study area. The determination of restricted areas is an important step in this study and is performed by following the dasymetric mapping technique. Dasymetric mapping is a method of thematic mapping which uses discrete areal data to visually represent a statistical surface [51]. The use of areal interpolation allows to include additional information about the study area and enables to visualize data distributions which would otherwise be obscured [52, 51]. This mapping method allows displaying a more realistic distribution of the data by reducing spatial discontinuities in the data produced by the boundary units of the collected information [53]. This mapping technique was first developed by Russian geographers who defined dasymetric maps as density measuring maps [54]. Dasymetric mapping is used in this study to provide a more realistic estimate of how the water demand nodes and edges may be distributed within the study area. This is achieved by using ancillary data to produce a new group of areal units that improve the representation of the data surface. Limiting variables and related variables are the two types of ancillary data variables used in general. Limiting variables can be used to remove areas where data variable (water pipeline) could not exist [51]. Typically, water pipelines cannot pass through any waterbodies such as lakes, streams, rivers etc. So, waterbodies is an example of limiting variable in this study and all the water bodies within the study area are selected as restricted areas. Related variables have some kind of association or relationship with the data variable which can be used to predict the distribution of the data variable [51]. Population density and road density are used as related ancillary attributes in this study. Population is the basis for water demand [3]. There is usually higher water demand in locations of high population density. The growth in populations in an area increases water demand for domestic, industrial and municipal uses. In particular, population growth is primarily responsible for the increase in domestic water demand [55]. Now, the demand for transportation is higher for a given population in a large unit than in a small one [56]. So, there will be more roads per unit area in places where the population per unit area is higher [57].

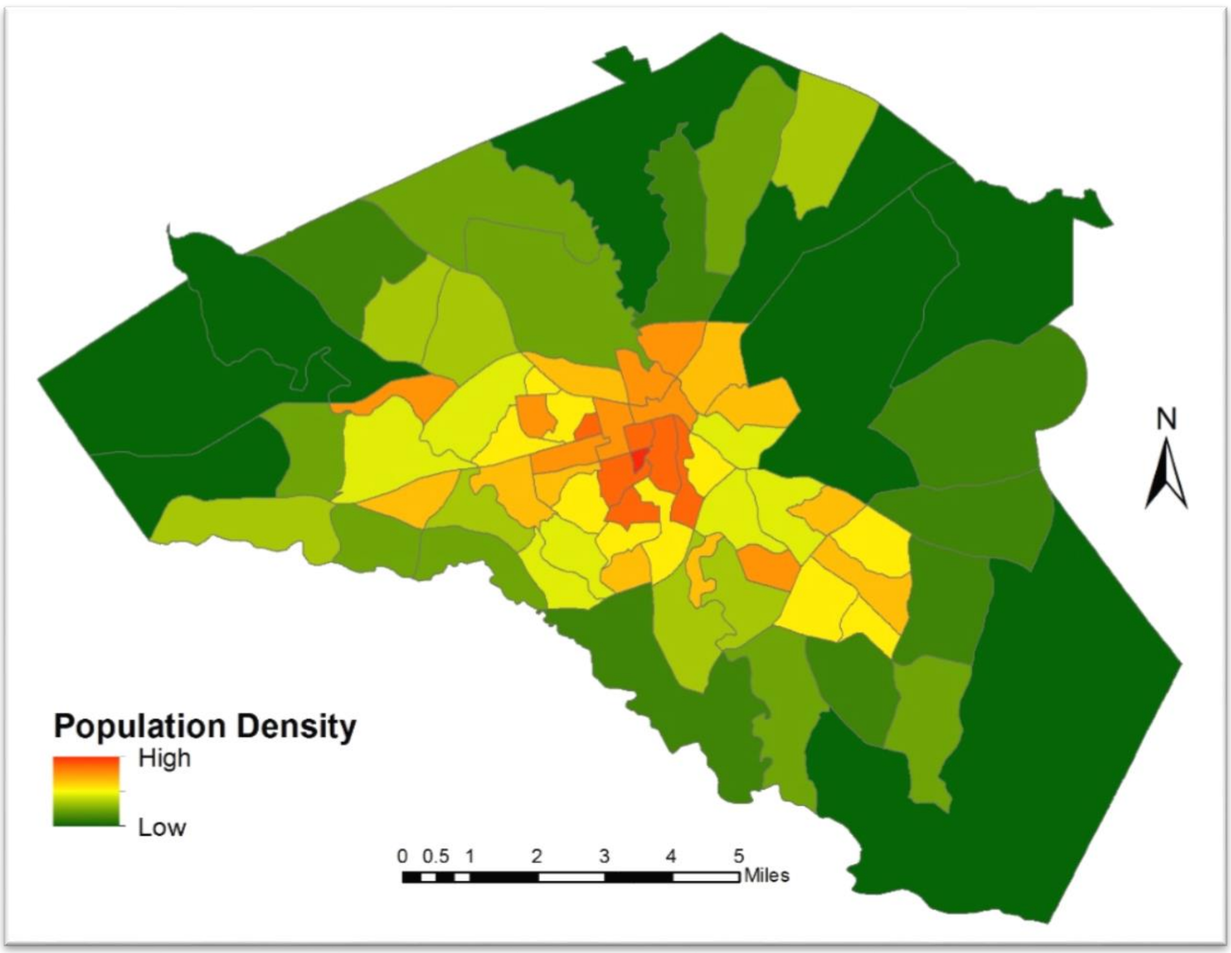


Fig 3.1. A map showing the variation of population density in Athens Clarke County

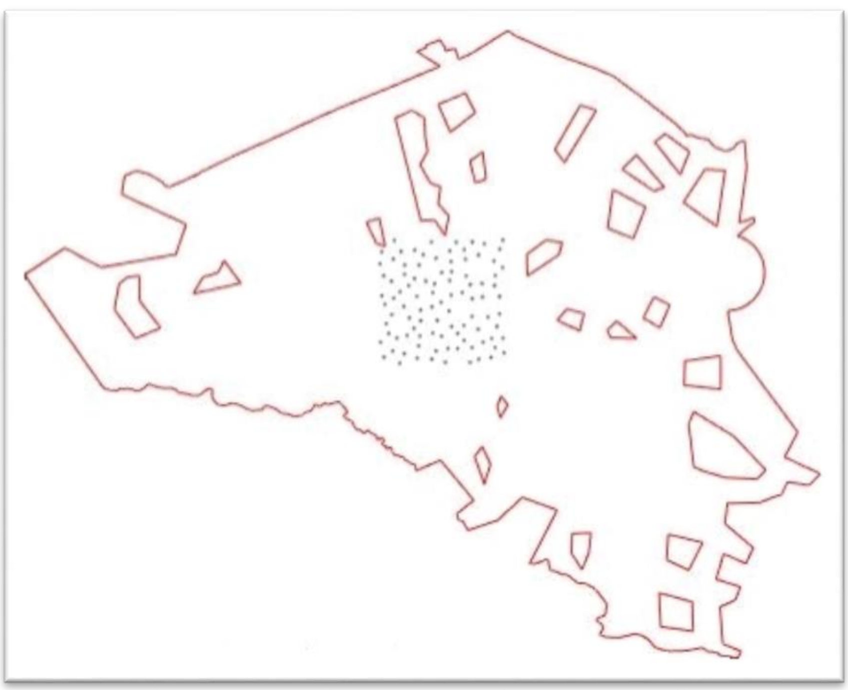
Fig 3.2. Road network in Athens Clarke County

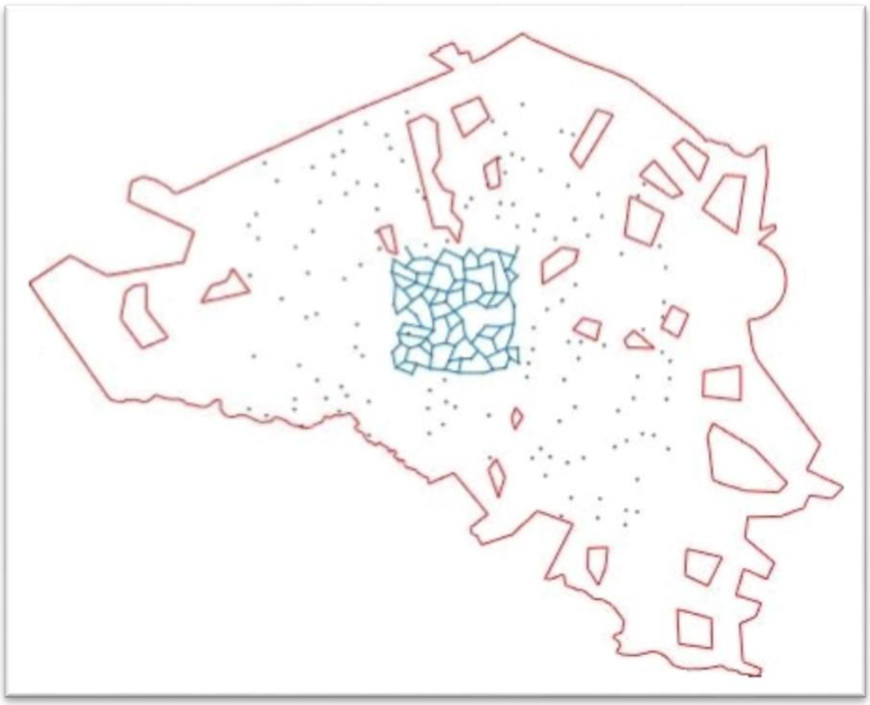
It has already been mentioned that the water demand is higher in places of high population density and now it has also been deduced that higher population density is a major cause of high road density. So, a direct relationship can be considered between water demand and road density. There is higher water demand in locations of high road density. In real world, water distribution network generally follows road network and more dense water network is found in places of dense road network. In fact, the WDN for our case study area in Athens Clarke County matches closely with its existing road network.

Fig 3.3. Water distribution network in Athens Clarke County

So, all the areas where no roads are present are also included in our restricted zone. This concludes the identification of the restricted zones using dasymetric mapping technique. After identifying all the restricted areas, they are drawn on the layer as a separate feature. Then those features are used to clip the study area to create holes which resemble the restricted areas. This concludes the processing of the study area in ArcMap. The next step is to import the study area into ArcMap. MATLAB can read vector features and attributes from shapefile by using the function "shaperead". This function reads vector features and attributes from a shapefile and returns a geographic data structure array. A successful import of the shapefile allows generating the WDN within the given study area.

The final step involves generating the WDN within the unrestricted areas. The WDN 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. The model generates WDN in a dynamic two-stage process. The first stage simulates the generation of new water demand following the growth of the urban network. At each time step *t*, multiple new water demand nodes would be generated for the given study area. The model detects restricted areas within the polygon of the study area as holes and does not create any demand nodes within those restricted areas. The second stage simulates the expansion of WDN due to the new water demand. As the demand increases, pipes are added into the system to connect new water demand nodes to the existing network. During the link generation, the program can identify whether a line joining two nodes would pass through the restricted areas and prevents generation of edges through the holes. In this way, the model can be used to generate WDN within the study area following the road network and avoiding waterbodies.





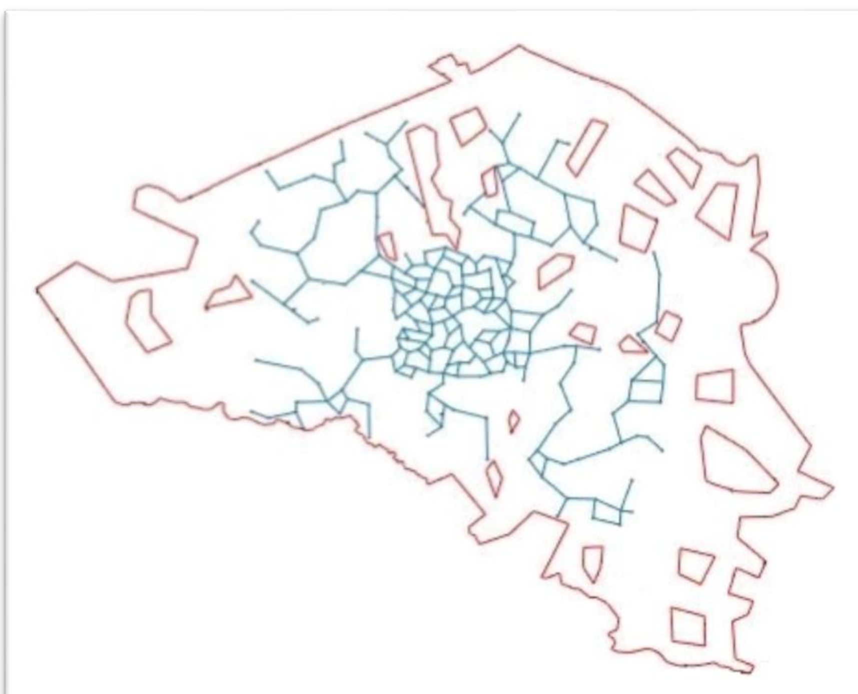


Fig 3.4. Dynamic generation of water demand nodes and edges. Water demand nodes generated avoiding the restricted areas (top). The links generated to connect the existing nodes and more new nodes are generated (middle). The links generated also follows the geographic restriction (bottom).

3.3 Flow Diagram of the WDN generation process

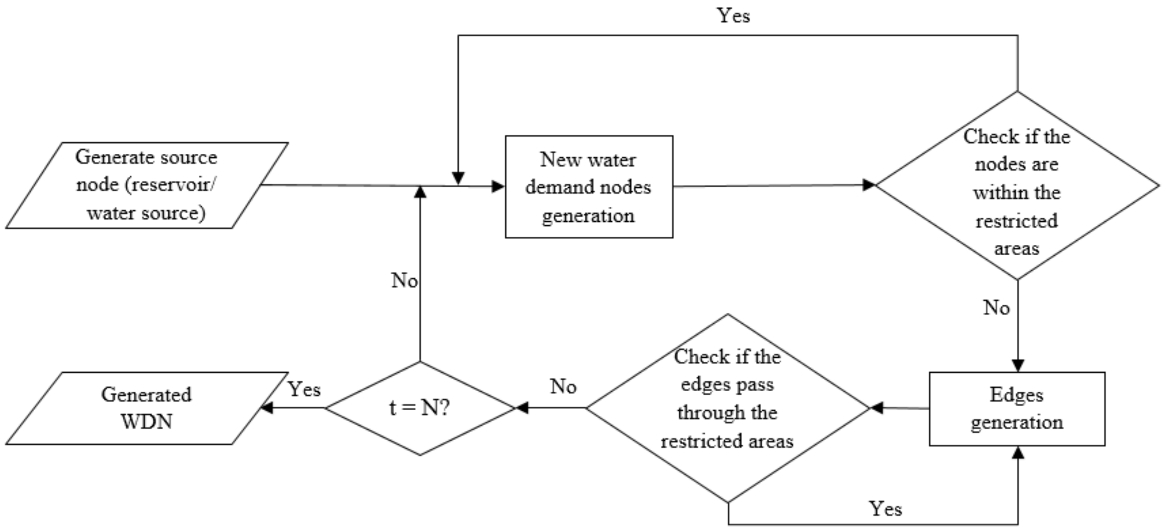


Figure 3.5. Schematic flowchart.

CHAPTER 4

RESULTS AND CONCLUSION

4.1 An example of generated WDN

After setting up all the criteria, the model is prepared to run for a particular study area. In this research, Athens Clarke County is chosen as the proposed study area. First, the shapefile of the Athens Clarke County is processed in ArcMap to include all the restricted areas. Then the layer is imported into MATLAB to generate water distribution network for Athens Clarke County. There are several modifiable parameters in the model to regulate the growth of the network. The scaling factor α can be used to simulate different urban forms. The larger the value of α, the more sprawled the water demand nodes will be. Since Athens is a sprawled city, so a higher value of α is used in this study. The weighting coefficient β allows regulating the weight of the shortest path distance to the root. There is a tradeoff between the total length of pipes and path distance from water source to consumer. So, a small value of β is used in this model. The value of the parameter d_{max} can be used to control the amount of looping in the network. A higher value of d_{max} will increase the level of looping in the network. Real world WDN are usually partially looped. So, the value of d_{max} was adjusted to achieve a partial looping of the network. After specifying all the parameters, the model has been run to generate WDN for Athens Clarke County. Figure 6. Shows the generated WDN which is a nearly fully looped structure around the center due to high population density.

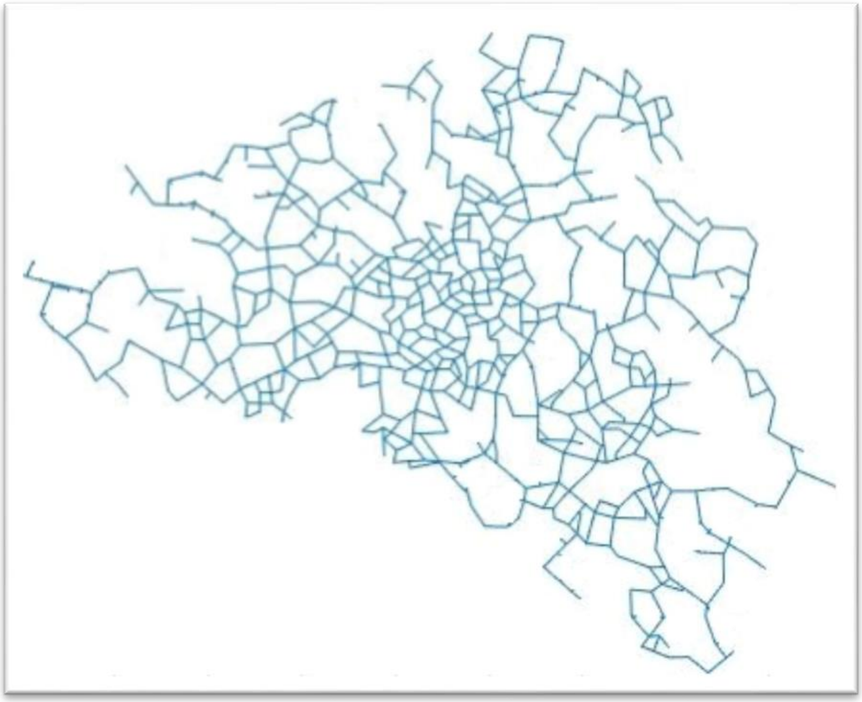


Fig 4.1. Generated WDN for Athens Clarke County

4.2 Comparison with real-world WDN

After generating the WDN, the next step is to validate the generated network. One of the ways to validate the network is to compare the structural properties of the generated WDN with the structural properties of real-world networks. The structural properties of the eight real-world networks (Figure 3.2) along with the generated WDN for Athens Clarke County (Figure 4.1) are listed in Table 4.1. The metrices used by researchers to explain the structural properties of WDN are described in section 2.2.

Table 4.1. 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; r_m, meshedness; λ_2, Algebraic connectivity).

	n	*m*	*e*	<*k*>	k_{max}	*q*	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	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	0.05	6.09×10^{-5}
Blacksburg[b]	31	35	7.53×10^{-2}	2.26	4		0.09	4.19×10^{-2}
Fossolo[b]	37	58	8.71×10^{-2}	3.14	4		0.32	7.37×10^{-2}
Pescara[b]	71	99	3.94×10^{-2}	2.79	5		0.21	3.26×10^{-3}
Modena[b]	272	317	8.60×10^{-3}	2.33	5		0.09	4.06×10^{-3}
Athens, GA	756	1012	3.55×10^{-3}	2.68	5	1.31	0.16	1.28×10^{-3}

[a]Data from Yazdani and Jeffrey's works [19]
[b]Data from De Corte and Sörensen's work[23,42]

The generated WDN for Athens Clarke County has 756 nodes and 1012 edges. The number of nodes and the number of edges in the generated WDN is smaller than Colorado Springs and Kumasi while more than all other networks. This is expected because both Colorado Springs and Kumasi are larger than Athens and the total population in both of these cities are also higher. So, both of those cities have higher water demand and larger network is required to cover all the areas. The edge density of the generated WDN is small which is similar to all the real world WDN. This shows that the generated WDN is a sparse network similar to the real world networks. The generated WDN has an average node degree of 2.68 which is larger than 2 but much smaller than 4. This is also evident in the results of the real world networks which implies that the generated WDN is much less connected than the grid network. Similar to some real-world network, the generated network has a maximum node degree of 5 and a minimum node degree of 1. The generated network has a similar but slightly smaller route factor than the real world networks which means that the simulated network is more efficient in distributing water. The number of loops for the generated network is smaller than Fossolob and Pescara but larger than other reported networks. This is due to the fact that the generated network has a fully looped structure at the center of the network. The algebraic connectivity for the generated WDN is higher than the real cases reported by Yazdani and Jeffrey and lower than those reported by De Corte and Sörensen. This implies that the simulated network is more robust than the real world networks reported by Yazdani and Jeffrey and less robust than those reported by De Corte and Sörensen. Overall, there is a similarity in general with the generated water distribution network and the real world networks. The generated WDN is also compared with the real WDN for Athens Clarke County. The structural properties of the real WDN for Athens Clarke County along with the generated WDN for Athens Clarke County (Figure 4.1) are listed in Table 4.2.

Table 4.2. Structural properties of real-world WDN for Athens Clarke County and the generated WDN (*n*, nodes; *m*, edges; e, edge density; <k>, average node degree; k_{max}, maximum node degree; *m/n*, edge-per-node ratio; r_m, meshedness).

	n	*m*	*e*	<*k*>	k_{max}	*m/n*	r_m
Real-world WDN	29063	29701	7.03×10^{-5}	2.04	4	1.02	0.01
Generated WDN	756	1012	3.55×10^{-3}	2.68	5	1.34	0.16

The number of edges and the number of nodes is higher in the real network. There are many very small pipes which results in a large number of edges and nodes. The edge density of the real-world WDN is smaller than the generated network. This shows that the real-world WDN is sparser than the simulated network. The real-world WDN has an average node degree of 2.04 which is larger than 2 but much smaller than 4. This is also evident in the generated WDN. The real-world network has a maximum node degree of 4 and a minimum node degree of 1 whereas our generated network has a maximum node degree of 5. The edge-per-node ratio of the real-world network is similar to the simulated network. The number of loops for the generated network is larger than the real-world network indicating that the simulated network is more reliable.

4.3 Conclusion

The primary purpose of the research is to bridge GIS and complex network analysis. A complex network model is used to model the growth of water distribution network (WDN). The model is developed following the combination of local optimization rules and engineering considerations. The demand nodes generation is dynamic and is governed by the scaling law of urban growth. A case study is performed to generate water distribution network (WDN) for Athens Clarke County. Along with the scaling law of urban growth, the growth of the network is also guided by geographic restrictions to simulate a real-world network. Dasymetric mapping method is used to determine restricted zones for the simulation of the WDN. This mapping technique provides a more realistic estimate of how the water demand nodes and edges may be distributed within the study area. Waterbodies are used as limiting variables whereas road density and population density are used as related variables in this study. Based on some structural properties, the developed model for Athens Clarke County is compared with real-world WDN's reported by other studies and with the real-world network for Athens Clarke County. Comparison with other real-world networks showed that the structural properties of the generated WDN comply with the values reported for real-world WDNs. This indicates that a practical WDN can be simulated by introducing geographic considerations into complex network approach.

REFERENCES

1. Selvakumar, A., R. M. Clark, and M. Sivaganesan, *Costs for water supply distribution system rehabilitation*. Journal of Water Resources Planning and Management, 2002. 128(4), 303–306.

2. American Water Works Association, *Buried no longer: Confronting America's water infrastructure challenge*, 2012.

3. Boulos, P.F., K. E. Lansey, and B. W. Karney, *Comprehensive Water Distribution Systems Analysis Handbook for Engineers and Planners*, USA, CO, Denver: American Water Works Assoc., 2006.

4. Rouse, W.B., *Complex engineered, organizational and natural systems.* Systems Engineering, 2007. 10(3): p. 260-271.

5. 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

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

7. Vespignani, A., *Predicting the behavior of techno-social systems*. Science, 2009. 325: 425–428.

8. Bar-Yam, Y., *Dynamics of Complex Systems*1997: Addison Wesley, Reading, MA.

9. Hmelo-Silver, C. E., and M.G. Pfeffer, *Comparing expert and novice understanding of a complex system from the perspective of structures, behaviors, and functions*. Cognitive Science, 2004, 28 (1): 127–38.

10. Wilensky, U., and M. Resnick, *Thinking in levels: A dynamic systems approach to making sense of the world*. Journal of Science Education and Technology, 1999, 8, 3–19.

11. National Academies, *Engineering the health care system*2006: National Academy Press, Washington, DC.

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

13. Barabási, A.L. and R. Albert, *Emergence of scaling in random networks*. Science, 1999. 286(5439): p. 509-512.

14. 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*, Springer, 2002. p. 110-122.

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

16. 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.

17. Ostfeld, A.: Water distribution networks. In: Kyriakides, E., Polycarpou, M. (eds.) *Intelligent Monitoring, Control, and Security of Critical Infrastructure Systems*. SCI, 2015, vol. 565, pp. 101–124. Springer, Heidelberg.

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

19. 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.

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

21. Shen, Y., and K. Vairavamoorthy, *Small World Phenomena in Water Distribution Network*. International Conference on Computing in Civil Engineering (2005), 2005. pp. 1333–1339.

22. 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.

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

24. De Corte, A. and K. Sörensen, *Optimisation of gravity-fed water distribution network design: A critical review*, Eur J Oper Res 228, 2013, 1–10.

25. What is GIS? Retrieved from http://www.esri.com/what-is-gis

26. Gavekari, V.S. and V. D. Nandavdekar, *Mapping of water distribution network through geographical information system (GIS)*, International Journal of Civil Engineering (IJCE), 2013. ISSN(P): 2278-9987; ISSN(E): 2278-9995 Vol. 2, Issue 5, Nov 2013, 33-44

27. Anderson, J. L., M.V. Lowry, and J.C. Thomte, *"Hydraulic and Water Quality Modeling of Distribution Systems: What are the Trends in the US and Canada."* AWWA Annual Conference, Washington, DC., 2001.

28. Bartolín, H., F. Martínez, *Modelling and Calibration of Water Distribution Systems. A new GIS approach*, 33rd ESRI International User Conference. San Diego (EEUU). CD-ROM Proceedings, 2003.

29. Shamsi, U.M., *GIS and Water Resources Modeling: State-of-the-Art*, Journal of Water Management Modeling, 1995, R204-05. DOI:10.14796/JWMM.R204-05.

30. Csillag, F., *Variation on hierarchies: toward linking and integrating structures*1996. In: Goodchild, M.F. et al. (Eds.), GIS and Environmental Modeling: Progress and Research Issues. GIS World Publication, Fort Collins, CO, pp. 433–437.

31. Walsh, M.R., *Toward spatial decision support systems in water resources*. J. Wat. Resour. Plan. Mgmt, 1992. 109 (2), 158–169

32. Ayad, A.R., H.A. Awad, A.A. Yassin, *Geographic Information Systems in Water Distribution Networks*, International perspective on water resources & the environment conference, Vol. 5, 2012.

33. Shamsi, U.M., *GIS Applications for Water Distribution Systems*, Journal of Water Management Modeling, 2004, DOI:10.14796/JWMM.R220-21.

34. Zhang, T., *The Application of GIS and CARE-W on Water Distribution Networks in Skärholmen Pressure Zone, Stockholm, Sweden*, Pipeline Technology Conference, 2006.

35. Kennedy, H., *ESRI Press Dictionary of GIS Terminology*. ESRI Press, 2001

36. Eljamassi, A. and R. A. Abeaid, *A GIS-Based DSS for Management of Water Distribution Networks (Rafah City as Case Study)*, Journal of Geographic Information System, 2013, 5, 281-291.

37. Nicolai Guth, N. and P. Klingel, *Demand Allocation in Water Distribution Network Modelling – A GIS-Based Approach Using Voronoi Diagrams with Constraints*, In M. A. Bhuiyan (Ed.), Application of geographic information systems, 2012, DOI: 10.5772/50014.

38. De Oliveira, D.P., L. Soibelman, Garrett, H.J., *"GIS applications for spatial analysis of water distribution pipeline breakage and condition assessment data"*, IEEE. Maximizing Performance of Our Pipeline Infrastructure v321 2008 Proceedings of Pipelines Congress 2008.

39. McKinney, D. C., and X. Cai, *Linking GIS and water resources management models: An object-oriented method*. Environmental Modelling & Software, 17(1), 413–425, 2002

40. Pachri, H., Y. Mitani, *Development of water management modeling by using GIS in Chirchik river basin, Uzbekistan*. International Conference on Earth Science and Technology Proceedings. Tashkent, 2012, pp. 169-176.

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

42. 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.

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

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

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

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

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

48. 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.

49. 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.

50. 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.

51. DeMers, M., *Fundamentals of Geographical Information Systems*2000: John Wiley and Sons, New York.

52. Flowerdew, R., and M. Green, *Data integration: Statistical methods for transferring data between zonal systems.* In Masser, I., and Blakemore, M. (eds.), Handling Geographical Information, 1991. pp. 38–54. Longman, New York.

53. Langford, M., and D. J. Unwin, *Generating and mapping population density surfaces within a geographical information system*. Cartogr. J., 1994. 31: 21–26.

54. Wright, J. K., *A method of mapping densities of population with cape cod as an example*. Geogr. Rev., 1936. 26: 103–110.

55. Schute, C.F., and W.A. Pretorious, *Water demand and population growth*, 1997. Water SA, 23: 127–133.

56. Taaffe, E. J., R. L. Morrill, and P. R. Gould., *Transport Expansion in Underdeveloped Countries: a Comparative Analysis,* 1963, Geographical Review, 53, pp. 503–29.

57. Glover, D. and J. L. Simon, *The effect of population density on infrastructure: the case of roadbuilding*. Economic Development and Cultural Change, 1975, vol. 23 (3), PP. 453-68.

58. Watts, D.J. and S.H. Strogatz, *Collective dynamics of `small-world' networks*. Nature, 1998. 393(6684): p. 440-442.