Winter 12-4-2019

# Modeling the Effects of Distribution System Topology on Water Quality

Chun-Ying Chao

Washington University in St. Louis

McKelvey School of Engineering

Department of Energy, Environmental & Chemical Engineering

Thesis Examination Committee:

Pratim Biswas, Chair

Palghat Ramachandran

Rajan Chakrabarty

Jeff Yang

# Modeling the Effects of Distribution System Topology on Water Quality

By

Chun-Ying Chao

A thesis presented to the McKelvey School of Engineering of Washington University in St. Louis in partial fulfillment of the requirements for the degree of Master of Science

December 2019

St. Louis, Missouri

# Dedication

This thesis is dedicated to my parents

for their love, endless support

and encouragement.

# Table of Contents

# List of Tables

# List of Figures

# Acknowledgements

First and foremost, I would like to thank my thesis mentor, Dr. Pratim Biswas, for challenging me to take an active role in steering my experiments while guiding to keep me to stay on track or correct my course when things weren't working out. You are deeply passionate about, not just your work, but also the scientific development and well-being of your trainees. I am extremely grateful to have you as a mentor.

Second, I would also like to thank my thesis committee members: Dr. Palghat Ramachandran, Dr. Rajan Chakrabarty, and Dr. Jeff Yang, for providing valuable guidance and helping keep projects focused. Additionally, much of this work was made possible by Dr. Ahmed A. Abokifa. Thanks in particular to Dr. Abokifa for taking at least two hours for our weekly meeting.

I would like to thank all of the current and former members of the AAQR lab: Dr. Liang-Yi Lin, Dr. Jiayu Li, Dr. Huang Zhang, Dr. Guang Wang, Dr. Mojgan Kouhnavard, Sungyoon Jung, Sukrant Dhawan, Clayton Kacica, Girish Sharma, Hao Zhou, Poom Sittisomwong, and Onochie Okonkwo. Thanks in particular to Sungyoon, for supporting me in a rough patch. Thank you Guang, Huang, Sukrant, Sungyoon, Hao, and Onochie for reviewing my presentation during the Thanksgiving vacation.

Finally, I would like to thank Mr. James Ballard, for consulting on the writing of my thesis. Without his help, I could not finish this.

Chun-Ying Chao

*Washington University in St. Louis*

December 2019

# Abstract

By

Chun-Ying Chao

Master of Science in Energy, Environmental, and Chemical Engineering

Washington University in St. Louis, 2019

Professor Pratim Biswas

Inadequate treatment of drinking water causes the formation of disinfection by-products and the regrowth of harmful microbial species. Various studies have addressed the problem of water quality monitoring, but very few have employed topological analysis, a valuable mathematical tool widely applied in biological, business, and social research. This thesis examines the relationship between the topological properties of water distribution systems and water-quality models. In particular, the research proposes a novel framework for mapping network topological attributes to water-quality models. This research adopts topological metrics to assess the accuracy of the predictions of chlorine concentrations in dead ends. It examines four fundamental water-quality models: advection, advection-dispersion, bulk-advection, and bulk-advection-dispersion. The results show the bulk-advection-dispersion model has larger root mean square errors in networks with a grid structure, and that topological metrics are generally correlated with water-quality models, although more studies are required to develop this correlation in detail.

# Chapter 1: Introduction

## 1.1 Background and motivation

Providing sufficient drinking water of appropriate quality and quantity has historically been a challenge in the United States. As populations grew, user demands increased. In St. Louis, for example, the population burgeoned from approximately 5,000 to more than 160,000 in the 1850's [1]. Water purification played a major role in reducing the impact of St. Louis' typhoid and cholera epidemic of 1903, which claimed 287 lives. A new treatment system, finished in 1908, saved 1,900 lives from 1903 to 1915 [2].

Beyond water treatment, water distribution systems play major roles in the United States' economy and public health. Water distribution systems deliver clean and safe water from treatment plants to consumers' taps through a complex and extensive pipe network. The systems consist of pipes, junctions, pumps, valves, storage tanks, reservoirs and other hydraulic infrastructure. Rapid population growth and increased urbanization present major challenges, maintaining and upgrading the water infrastructure efficiently, from both operational and public health standpoints.

Aging water infrastructures contain pipes range from cast iron pipes installed during the late 19th century to ductile iron pipe and finally to plastic pipes introduced in the 1970s and beyond. Most water systems and distribution pipes will be reaching the end of their expected life spans in the next 30 years. Over 34 billion gallons of water are delivered annually by aging infrastructures [3]. This infrastructure is often operated manually, technological advances are limited to monitoring and management. There are an estimated 240,000 water main breaks per year in the United States, wasting over two trillion gallons of treated drinking water. The American Water

Works Association estimated that will cost at least $1 trillion to maintain and expand service to meet demands over the next 25 years [5].

The aging infrastructure also degrades water quality, which affects not only the environment and economy, but also public health. Based on the National Public Water Systems Compliance Report in 2013, systems with reported health-based violations served approximately 26.5 million consumers, and approximately 48 percent of the 27,056 public water systems in 2013 had at least one violation of the monitoring and the reporting requirements of the Total Coliform Rule[6]. Health-based standards include the allowable maximum contaminant level (MCL), maximum residual disinfectant level (MRDL), and a required treatment technique (TT) such as filtration or disinfection intended to prevent the occurrence of or deactivate contaminant in drinking water. If the water quality exceeds MCLs or MRDLs, or it lacks a satisfactory TT, the consumers face an increased possibility of health risks. According to the Centers for Disease Control, in 2014 there were 6,939 deaths and 477,000 emergency department visits as a result of drinking water contaminated with pathogens, including *Legionella* and *Pseudomonas* [7]. The emergency department visits resulted in $194 million in annual direct costs. Besides the contaminants in drinking water, disinfection byproducts (DBPs), such as trihalomethanes (THMs) and haloacetic acids (HAAs), are also potentially harmful agents, and long-term exposure to them can affect infants’ and young children’s nervous systems.

### 1.1.1 Disinfection in Water Distribution Systems

Most drinking water regulations focus on water quality at the treatment plant and not within the distribution system. In an ideal situation, the quality of drinking water should not change from the time it leaves the treatment plant to the time it is consumed. In reality, when drinking water leaves the treatment plant, the water quality deteriorates gradually during delivery through the distribution system, because complex physical, chemical, and biological reactions occur in the pipes after the water leaves the treatment plants.

Disinfection in water distribution systems eliminates pathogens that are responsible for waterborne diseases. Chlorination is the most widely used method for disinfecting water supplies in the United States, because of its convenience and highly satisfactory performance [8]. However, if the residual chlorine concentration is below the required value, chlorine-based disinfectants can react with natural organic matters (NOMs) that remains after water treatment, forming as trihalomethanes (THM's), disinfection byproducts (DBPs). Reactions with NOMs deplete of the disinfectant residual and leading to biofilm development in pipes. Although the general heterotrophs found in biofilms are unlikely to become a public health concern, the growth of biofilms in distribution systems still should be minimized. The biofilms can harbor opportunistic pathogens and increase their resistance to disinfection, leading to waterborne disease outbreaks. Furthermore, reactions with NOMs may contribute to corrosion, not only increasing maintenance costs, but also increasing the frequency of breaks, the discoloration of the drinking water [9], and the release of toxic heavy metals. The large-scale network means longer time for water to transport than small-scale network. The long residence time allows more interactions to occur and contributes dramatically to water quality deterioration.

In mean time the challenges of provide drinking water of appropriate quality and quantity, and maintaining and replacing aging infrastructures under limited budgets, incompetent management and improper operation should obviously be avoided. One recent famous case is the Flint, Michigan, lead poisoning outbreak. In 2014, thousands of people in the City of Flint, Michigan, were exposed to dangerously high levels of lead after the city officials changed to a more corrosive water source. The more corrosive water caused the treated water's quality to deteriorate fast, producing high levels of THMs [7, 8].

The tragedy in Flint captured public attention and emphasized the need for increased investments in the aging drinking water infrastructure across the U.S. replacing inadequate facilities will cost more than $1 trillion over the next 25 years [7]. Given this extraordinary expense, accurate models of water distribution become critically important. Appropriate models can decrease cost, predict the problem areas.

### 1.1.2 Graph Theory in Water Distribution Networks

Graphs that represent visual data, help us make better decisions. Like electric power lines, roads, and microwave radio networks, water systems can have a grid or branch network topology, or a combination of both. If any one section of a water distribution main fails or needs repair, that section can be isolated without disrupting all users on the network. In general, the physical layout design of water distribution networks is dependent on the natural source of the supply, the demand nodes, the location of physical barriers such as roads, terrain, rivers, and so on. Water distribution networks can be discussed in terms of their hydraulics, telemetry systems, history, user population, and topology [10].

To evaluate the topology of a complex water distribution system, graph theory simplifies the systems into nodes and links, and omits details of the physical systems and processes [11]. The graph illustrates the relationship or connections between the nodes of systems. Graph theory has been widely applied in different fields: in marketing analytics, it can be used to find the most influential people in a social network; in supply chains, it can optimize routes for delivery; in biological studies, it can measure the relationship between events and disease propagation [11, 12].

A topological attribute is defined as a property that is preserved under a homeomorphism, and topological attributes have strong impacts on system resilience [9].

The nodes in water distribution systems are typically sources, such as reservoirs, tanks, and storage facilities; control and distribution nodes, such as valves, and pipe junctions; and demand nodes or sinks. The links are transmission and distribution pipes of various martials, lengths, and diameters [11]. Water distribution systems can have a few hundred nodes or thousands to millions of nodes. Many complex water distributions systems have different degree of redundancy, with more alternative paths to reach a given node, and topological metrics are valuable in assessing this redundancy.

### 1.2 Ensuring Compliance with Water Quality Regulations

Regulations for ensuring the safety of drinking water are a powerful tool to protect public health in both developed and developing countries. In 1958, the World Health Organization (WHO) published the first edition of The Guidelines for Drinking-water Quality (GDWQ). The GDWQ are an international reference to establish national or regional regulations and standards for water safety [13]. They assist water quality and health regulators, policymakers, and their consultants to build national standards and regulations. The GDWQ derive maximum and minimum concentration guideline values for the various microbial, chemical contaminants that may be in the drinking water. In the United States, chlorination is the most widely used method for disinfecting water supplies because of its convenience and highly satisfactory performance. The GDWQ recommend that the concentration of free chlorine should be between 5.0 mg/L and 0.1 mg/ L. Based on their data, no countries set their maximum above 5.0 mg/L because of an increased risk of bladder cancer [13].

In the United States, Congress passed The Safe Drinking Water Act (SDWA) in 1974 to protect public health by regulating the nation's public drinking water supply. SDWA authorizes the United States Environmental Protection Agency (US EPA) to set standards for drinking water to protect against hazardous contaminants. Originally, SDWA focused primarily on treatment as the means of providing safe drinking water at the tap. Amendments made to SDWA in 1996 greatly enhanced the existing law by recognizing source water protection, operator training, funding for water system improvements, and public information as important components of safe drinking water. This approach seeks to ensure the quality of drinking water by protecting it from source to tap. The SDWA set the concentration of free chlorine between 4.0 mg/L and 0.2 mg/ L, and more 170,000 public water systems are responsible for meeting these standards, and most states follow them [14].

## 1.3 Research Gaps and Engineering Challenges

To meet regulatory requirements and people's expectations, scientists have continued to develop new models in support of the planning, design, and management of water distribution systems. EPANET is a software application that helps meet this goal. It predicts the dynamic hydraulic and water quality behavior within a drinking water distribution system operating over an extended period of time. EPANET, however, has limited accuracy in simulating chlorine decay and transport in problem zones, such as low-flow zones and mixing junction zones.

Water distribution systems (WDS) are traditionally built with topological redundancy to improve network reliability against mechanical and hydraulic failure. To reduce the complexity of a network, graph theory and topological metrics are widely used in hydraulic models. Little work, however, has been conducted on the impact of topological metrics on the selection of a water quality model.

This section highlight the research gaps and engineering challenges that stand in the way of successfully simulating water quality in complex water distribution systems. This thesis focuses on improving water quality simulation models and applying graph theory metrics to them.

Incomplete mixing in pipe junctions can play an important role in water distribution systems. The mixing behavior depends not only on the Reynolds number but also on the geometric configuration of junctions, such as double tee junctions and cross junctions [15, 16]. Most water distribution analysis software, like EPANET, basically assumes perfect mixing in junctions. It considers that incoming fluid streams with different containment concentrations are well mixing within the junctions and that the all concentrations at all outlets are equal. To refine this problem, new models consider bulk flow mixing in junctions. Mixing parameters or the degree of mixing have been used to describe bulk mixing [15, 16, 17, 18], and the results match both computational fluid dynamic modeling and laboratory experiments. These mixing parameters are applied in some models, but not for complex networks.

Booster chlorination is an approach to maintaining a chlorine residual by injecting the disinfectant at strategic locations within the water distribution system in smaller, more distributed, doses [19]. An adequate residual not only protects the public health but also reduces the formation of DBPs. EPANET is a valuable tool for identifying situations where booster chlorination is the most effective way to maintain a residual, but it does not work well for low-flow zones. Field data differs from simulation predictions [20]. The discrepancy causes potentially flawed results in solving network optimization problems, such as the placement and scheduling of booster chlorine stations and real-time boost-response schemes.

The concept of system resilience has been increasingly used to ensure that water distribution systems can rapidly recover from potential failures. Researchers have generally assumed that topology is correlated with the resilience of water distribution systems, and this correlation the basis of many studies on assessing and building resilience [11]. However, there is little work on mapping network topological attributes to water quality model performance.

## 1.4 Literature Review

This section discusses varying aspects of problem zones in water distribution systems, and focuses on graph theory as it is applied in the management of water distribution systems.

Biswas et al. [21] developed a generalized model to simulate chlorine consumption at the pipe wall and to determinate the average chlorine concentration under steady-state flow. The model accounted for mass transport in the axial and the radial directions by advection and dispersion or diffusion. It was the first model that explained chlorine decay both at the wall and in the bulk in the axial and the radial directions.

Rossman et al. [22] developed a film resistance model for predicting chlorine decay in drinking-water distribution networks. The film resistance model assumed that the consumption by the wall reaction could be represented by the mass-transfer coefficient. Based on this assumption, the 1-D advection-reaction model was incorporated with hydraulic and water-quality simulation models, becoming well-known software package EPANET. The water quality results of EPANET matched the chlorine measurements in most transmission mains. However, in low flowrate zones which the film resistance model could not explain well, the results were discrepant. These zones were mostly dead ends in the water distribution systems.

Axworthy et al. [23]first pointed out that a model that did not consider diffusion in low flowrate zones in water distribution systems would have large discrepancy in low flow simulations. Most of the time diffusion can be neglected because of the high flow rate in the pipes. During the night, however, the low flow domain of the water distribution system means that concentration changes are driven by diffusion. Thus, if the models do not account diffusion in the 24-hour simulations, they will underpredict the required concentration of chlorine at dead end locations. The dispersion can vary according to the length, diameter, and wall roughness of the pipes ,and this research presented here provides a means to quickly evaluate whether the advective transport model is appropriate for the simulation.

Tzatchkov et al. [24] developed an advection-dispersion model by using Green's functions to compute the numerical solutions for residual chlorine in water distribution systems. The model used 2-D advection-dispersion reactions to compute the chlorine concentration, and dispersion coefficient was used in the equations. The results of the advection-dispersion model were compared field measurements and simulations with the EPANET advection-reaction model. For velocities lower than 0.02 m/s, their model closely predicted the field measurement. For medium or high velocities, their model yielded the same prediction as EPANET. In this research, the advection-dispersion model improved the problem predicting concentrations at low flow, but to simulate dead ends appropriately, the unpredictable nodal demand also needs to be considered.

Abokifa et al. [20] developed a model for simulating disinfectant residuals in dead-end pipes. This model (Washington University Dead End Simulator – WUDESIM) is coupled not only with a stochastic flow demand generator but also with advection-dispersion reactions. In this research, they found the simulation error due to the different demand shares of the withdrawal nodes. To overcome the simulation error caused by spatial aggregation approximation, they sed three correction factors to adjust residence time, dispersion rate and wall demand. Compared with Tzatchkov's model and EPANET, their results showed better agreement with field-measured concentrations of free chlorine disinfectant, demonstrated that flow demands have significant dependence on spatial distribution.

Romero-Gomez et al. [17] investigated solute mixing phenomena at various flow rates within a cross junction, a type of connection commonly found in municipal drinking water distribution systems. They used computational fluid dynamics to simulate the solute concentrations leaving the junction at the Reynolds numbers larger than 10,000. They controlled one inlet to provide clean water while the other inlet carried a solute, which could be free chlorine or contaminant. The results indicated that mixing at pipe cross junctions was not perfect mixing, which most models assume. The incomplete mixing could cause chlorine concentration to be underpredicted or lead to an overdose of chlorine in drinking water distribution systems.

Ho et al. [18] focused on solute mixing and transport in cross junctions and developed a new model describing bulk advective mixing there. Based on mass balance, they added a mixing parameter to the previous model, EPANET, and implemented it with the hydraulic and water-quality models. They defined the mixing parameter to be 1 for complete mixing and 0 for bulk advective mixing. To confirm the value of the mixing parameter, they measured the mixing parameter in laboratory-scale experiments, and the results of field measure merits matched the simulations.

Shao et al. [15, 16] focused on mixing in laminar, transitional or uncompleted turbulent flow, and on double-tee junctions of unequal pipe sizes. The experimental results indicated that the average Reynolds number and the outflows the Reynolds number ratio having an influence on the mixing at the cross junction and double-tee junction. In pipe diameter experiments, it showed that junctions with larger diameters experienced more complete mixing. These results are important for contemporary water distribution models, most of which do not consider incomplete mixing.

Vanessa et al. [25] developed a model combining hydraulic and water-quality analyses to predict the changes in water quality after implementation of district metered areas (DMAs). To reduce leakage and break frequencies. DMAs were created many by setting valves to form closed boundaries, but this also created many dead ends. The overall water quality did not change following DMA implementation, but the water quality was degraded at dead ends because of the long residence time. The dead ends fostered high concentrations of trihalomethane and excessive growth of biofilms. The model closely pointed out the structural change of the network had an impact on water quality.

Yazdani et al. [10] used a link-node representation of water distribution systems and employed advanced network theory metrics to evaluate the building blocks of the systems and quantify their properties. In this research, the authors not only evaluated the performance of water distribution by using redundancy and fault tolerance, but also established relationships between the structural features and the performance of the water distribution systems. The research demonstrated that topological metrics are valuable tools for engineers and planners designing networks.

Nardo et al. [26] used novel topologic and energy metrics for water distribution network analysis, design, and partitioning. They used topologic metrics to analyze four existing networks and two literature networks and to identify general features of the water distribution networks. Based on the features, they optimized the network to minimize needless redundancy that might cause mechanical and hydraulic failures. However, the authors also mentioned that more studies are required for using appropriate metrics in designing water distribution systems.

Meng et al. [11] assessed resilience by using six topological metrics (connectivity, efficiency, centrality, diversity, robustness, and modularity). They measured system performance by six other metrics corresponding to system resistance, absorption and restoration capacities. They pointed out that the assumption that topological metrics have a great impact on water distribution systems was not justified and requires investigation. Only topological attribute metrics alone can guide the design of resilient water distribution systems, and other key details need to be considered.

Despite the abundance of modeling methods, there is no research establishing relationships between structural features and water quality in water distribution systems. The primary objective of this study is to use a new model to consider both the bulk-advection mixing and advection-dispersion transport in dead ends, thus establishing the relationship between topological metrics and water quality under the assumption that topology has a great impact on water quality.

# Chapter 2: Methodology

## 2.1 Water-Quality Models

In this thesis, four different water-quality models were used to simulate the water quality in dead ends. Topological metrics were used to find the relation between water-quality models and network structures. In this chapter, the mathematical backgrounds of four models: EPANET, EPANET-BAM, WUDESIM ,and WUDESIM-BAM; the topological metrics would be introduced. Also, the four scenarios of evaluating the correlation between water-quality models and topological metrics would be introduces.

### 2.1.1 EPANET

The governing equations for water quality models are based on the principles of conservation of mass and coupled with reaction kinetics. Although EPANET, it does not consider dispersion as a transport mechanism in pipe flow, free chlorine transport and decay in a dead-end pipe can be described by the mass balance. The disinfectant concentration in bulk flow $C(x,\ t)$ is written in a dynamic 1-D advective equation [21]:

$$\frac{\partial \mathrm{C}}{\partial \mathrm{t}} = -\frac{\partial}{\partial \mathrm{x}}(\mathrm{uC}) - \mathrm{k_b C} - \frac{\mathrm{k_f}}{\mathrm{r_h}}(\mathrm{C} - \mathrm{C_w}), \tag{1}$$

where, $x$ is the axial dimension coordinate (in meters), $t$ is the time (sec), $u$ is the average flow velocity in the pipe (m/sec), $k_b$ is the first order decay rate constant in the bulk flow (sec-1), $k$f is the mass-transfer coefficient, $r_h$ is the hydraulic radius of the pipe (1.5 times the pipe radius), and $C_w$ is the chlorine concentration at the pipe wall. The term $\frac{\partial \mathrm{C}}{\partial \mathrm{t}}$ represents the rate of change of the chlorine concentration within a differential section of pipe. The term $-\frac{\partial}{\partial \mathrm{x}}(\mathrm{uC})$ accounts for the advective flux of chlorine through the section, $-k_b C$ represents chlorine decay within the bulk flow, and $-\frac{k_f}{r_h}(C - C_w)$ accounts for the transport of chlorine from the bulk flow to the pipe wall.

The second term and the third term on right hand side of (1) can be simplified to

$$KC = (k_b + \frac{k_f k_w}{r_h(k_w + k_f)}) \times C \tag{2}$$

under the assumption that the reaction of chlorine at the pipe wall is first-order with respect to the wall concentration $C_w$ and that it proceeds at the same rate as chlorine is transported to the wall.[22]. K is an overall decay constant that contains the bulk decay constant, wall decay constant, hydraulic radius, and mass-transfer coefficient. Substituting (2) in equation (1) yields

$$\frac{\partial C}{\partial t} = -\frac{\partial}{\partial x}(uC) - KC. \tag{3}$$

In EPANET, mixing at pipe junctions follows the mass balance equations, assuming the mixing of fluid is complete and instantaneous. At junctions receiving inflow from two or more pipes, complete mixing happens in a short time and is quickly distributed to the next nodes. Thus, the concentration of a substance in water leaving the junction is the fraction of the total amount of chlorine from the inflowing pipes over the total amount of water from the inflowing pipes. The chlorine at the specific node can be described by

$$C_{out} = \frac{\sum Q_{in} C_{in}}{\sum Q_{out}}. \tag{4}$$

EPANET consider advective reaction in pipe flow and complete mixing in pipe junctions, and assumes that both obey mass balance and the first-order decay reaction of chlorine.

### 2.1.2 EPANET-BAM

EPANET-BAM is a model that incorporates the Bulk Advective Mixing (BAM) model. The BAM model considers momentum transfer and the separation of impinging fluid streams within a cross junction [17, 18]. This model follows the advective transport within pipes and governing equation (3).

At junctions, EPANET-BAM adds a mixing parameter to simulate incomplete mixing. The maximum mixing parameter is 1(complete mixing) and the minimum is 0 (no mixing). The concentration at the pipe junctions concentration can be expressed as follows:

$$C_1 = C_4 \tag{5}$$

$$C_3 = \frac{Q_2 C_2 + (Q_1 - Q_4) C_1}{Q_3}. \tag{6}$$

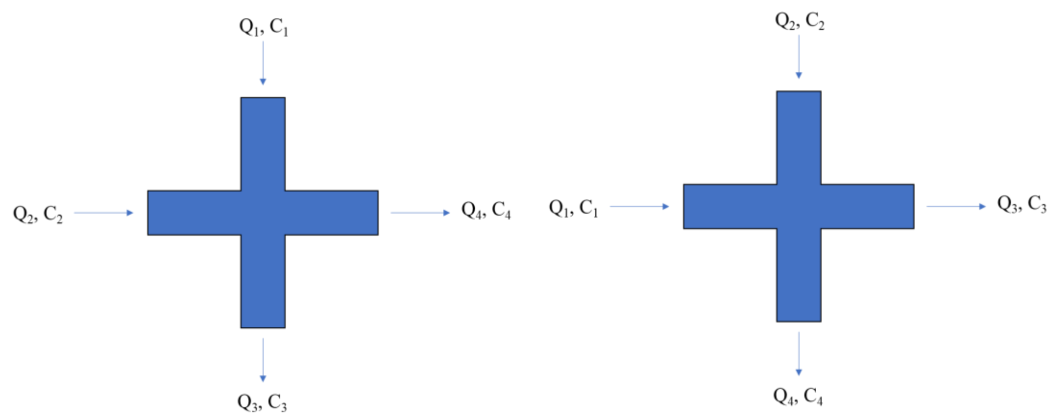


Figure 1. Incomplete mixing in pipe junctions. In the bulk advective mixing model, there are two flow configurations: a) the greatest momentum is in the vertical direction; b) the greatest momentum is in the horizontal direction.

In Figure 1, $C_1$ and $C_2$ are known concentrations at the inlet ,and $C_3$ is determinated by the amount of flowrate $Q_4$. If the $Q_4$ is equal to 0, the mixing is complete; if $Q_4$ is equal to $Q_1$, bulk mixing and $C_3$ is equal to $C_2$. The pipe junction concentration can be expressed with the mixing parameter s as follows:

$$C_{out} = C_{bulk} + s(C_{complete} - C_{bulk}), \tag{7}$$

where s has values ranging from 0 to 1. For s is equal to 1, the solution is completely mixed, and the result is the same as EPANET. Generally, a mixing parameter value between 0.2 and 0.5 yielded good matches with Ho's experimental data.[18]

### 2.1.3 WUDESIM

The new model (the Washington University Dead End Simulator, WUDESIM), is coupled with a stochastic demand generator based on a nonhomogeneous Poisson process to simulate residential water demand pulses on fine time scales [20]. The chlorine transport in dead ends can be appropriately modeled by a dynamic 2-D convection-diffusion equation in cylindrical coordinates to represent the mass balance on the disinfectant concentration C(x,r,t), which can be written as

$$\frac{\partial C}{\partial t} = -\frac{\partial}{\partial x}(uf(r)C) + \frac{\partial}{\partial x}\left(D\frac{\partial C}{\partial x}\right) + \frac{1}{r}\frac{\partial}{\partial r}\left(rD\frac{\partial C}{\partial r}\right) - KC, \tag{8}$$

where x and r are the axial and radial spatial coordinates, respectively (m); t is the time (sec); u is the average flow velocity in the pipe (m/sec); f(r) is the radial flow distribution parameter; D is the molecular diffusivity of the solute in water ($m^2$/sec); and K overall decay constant ($sec^{-1}$).

The term on the left-hand side is a net accumulation of chlorine in the control volume. The first term on the right-hand side is a chlorine concentration due to the inflow of the water into the control volume. The second term on the right-hand side is the chlorine concentration due to the diffusion of chlorine in the axial direction. The third term on the right-hand side is the chlorine concentration due to the diffusion of chlorine in a radial direction. The last term on the right-hand side is the loss of chlorine concentration due to chemical reactions. In the pipe reaction, the diffusion in the radial direction can be neglect. Thus, equation (8) is a 2-D unsteady convection diffusion equation that can be simplified to a 1-D unsteady advection-dispersion equation by using the effective longitudinal dispersion coefficient ($m^2$/sec). Using the simplified equation below can avoid intensive calculation and increase the computational efficiency:

$$\frac{\partial C}{\partial t} = -\frac{\partial}{\partial x}(uC) + E\frac{\partial^2 C}{\partial x^2} - KC, \tag{9}$$

where the K is overall decay constant and E is the effective longitudinal dispersion coefficient. In practice, using dimensionless quantities can describe the equation more clearly. One concern when simplifying Equation (8) into equation (9) is the error caused by neglecting the combined effects of radial molecular diffusion and the flow velocity profile in the radial direction [20]. Choosing an appropriate dispersion coefficient is crucial for the success of the simulation. In WUDESIM, two dispersion coefficients can be chosen: Taylor's coefficient (10) and Lee's coefficient (11)

$$E_T = \frac{a^2u^2}{48D}, \text{ and} \tag{10}$$

$$\overline{E}_k = \frac{1}{(t_k - t_{k-1})}\int_{t_{k-1}}^{t_k} E_k(t)dt. \tag{11}$$

Here a is the pipe radius, u is the average flow velocity in the pipe, D is the molecular diffusivity of the disinfectant in water ($m^2$/sec), where, $E_k$ is the instantaneous dispersion coefficient for pulse $k$, $t_{k-1}$ is the ending time of pulse $(k-1)$, $\bar{E}_k$ is averaged dispersion coefficient during any pulse (k).

Equation (9) can use the Péclet number (Pe) and the Damköhler number (Da). The Péclet number represents the ratio of the convection rate over the diffusion rate, and the Damköhler number is the ratio of the chemical reaction rate over the diffusion rate. Equation (9) can be transformed to

$$\frac{\partial C*}{\partial t*} = -\frac{\partial C*}{\partial x*} + \frac{1}{Pe}\frac{\partial^2 C*}{\partial x*^2} - DaC*, \qquad (12)$$

where C* is the dimensionless concentration, $C/C_0$; t*; is the dimensionless time, $t/t_0$; x* is the dimensionless distance, x/L; Pe is the axial Peclet number, uL/E; and Da is the Damkohler number. $C_0$ is a reference concentration, usually taken as the inlet concentration (mg/L), while $t_0$ is the characteristic residence time L/u (sec), and L is the pipe length (m).

The Péclet number indicates whether the solute transport is dominated by advection or diffusion. For axial mass transfer in a pipe geometry, the Péclet number can be replaced by the characteristic time of diffusion, $\tau_d = \frac{L^2}{E}$, when the molecular diffusion is negligible compared to both the laminar dispersion and the characteristic time of diffusion over the characteristic time of advection. The characteristic time of advection is $\tau_a = \frac{L}{u}$, and. Thus, the Péclet number is $\tau_d/\tau_a = \frac{Lu}{E}$. When the Péclet number is large, the diffusion time is much larger than the advection time, so the solute transport is dominated by advection. The second term on the right side in equation (10) can be neglected, and the equation (10) then will become a dimensionless form of advective equation (3). Most of the time chlorine transport in a water distribution system follows the advective equation. In dead ends, however, it does not. In such a low flow situation, $\tau_d$ is much smaller than $\tau_a$, and the Péclet number is zero, i.e., transport is mainly dominated by diffusion. In dead ends, laminar flow conditions prevail, so the advection-dispersion equation (12) is an appropriate representation because it considers both advection and diffusion.

The Damköhler number is the reaction rate divided by the advective rate. For a chlorine decay reaction, a first order reaction rate is used. The relation between the reaction rate and advective rate can use a time scale to represent the Damköhler number: $D_a = \frac{\tau_a}{\tau_r}$ and $\tau_a = \frac{L}{u}$, $\tau_r = \frac{1}{K}$. Thus, the Damköhler number is $\frac{KL}{u}$. When the reaction rate is very small or the advection rate is very large, the Damköhler number is close to zero.

For mixing at a pipe junction, WUDESIM makes the same assumptions as EPANET: the mixing obeys mass balance equations and is complete and instantaneous, following equation (4).

### 2.1.4 WUDESIM-BAM

WUDESIM-BAM considers the advection-dispersion reaction,

$$\frac{\partial \mathrm{C}*}{\partial \mathrm{t}*} = -\frac{\partial \mathrm{C}*}{\partial \mathrm{x}*} + \frac{1}{\mathrm{Pe}}\frac{\partial^2 \mathrm{C}*}{\partial \mathrm{x}*^2} - \mathrm{DaC}*,$$

in pipes and incomplete mixing at cross junctions,

$$\mathrm{C_{out}} = \frac{\sum \mathrm{Q_{in}C_{in}}}{\sum \mathrm{Q_{out}}}.$$

For water distribution systems, the two equations are crucial, and their omission can cause larger errors.

## 2.2 Topological Metrics

Topology has been assumed to have a great impact on the resilience of water distribution systems, and it is the basis of many studies on assessing and building resilience [11]. A water distribution network can be represented as nodes and links, a process which can reveal important weaknesses and defects. We call the properties of the network topological attribute metrics. Using metrics for these topological properties, we can evaluate the resilience of networks. However, this fundamental approach has not been applied to assessing water-quality models. In this research, we assume that topological analysis can be a useful tool for evaluating water quality models in different kinds of networks. In this section, four topological metrics are introduced: the link density, the mean node-degree, the meshedness coefficient, and the dead-end fraction.

Table 1 Topological Attribute Metrics

| Definition | Metric | Parameter |
|---|---|---|
| Link Density | $\frac{2m}{n(n-1)}$ | m = pipes<br>n = nodes |
| Mean Node-Degree | $\frac{2m}{n}$ | m = pipes<br>n = nodes |
| Meshedness Coefficient | $\frac{m-n+1}{2n-5}$ | m = pipes<br>n = nodes |
| Dead-End Fraction | $\frac{n_d}{n_{all}}$ | $n_d$ = dead-end nodes<br>$n_{all}$ = total nodes |

The link density for a network is given by $\frac{2m}{n(n-1)}$. This fraction relates the total links present and the maximum possible links to indicate the sparseness or density of the connectivity in the network layout. It provides information about the general level of connection between the nodes of a graph in terms of “inclusivity” [10, 11, 26].

The mean node-degree is given by $\frac{2m}{n}$, the average number of connections per node. It provides immediate information on the organization of the network, representing the total number of “connections” that the network has on average [10, 11, 26].

The meshedness coefficient is given by $\frac{m-n+1}{2n-5}$, the fraction relating the actual number of loops and the maximum possible number of loops. It is used to describe the structural organization of water distribution systems. The value of the meshedness coefficient can indicate the structure of the networks. If the number is large, the network is a grid structure. If the number is small, the network is treelike. In general, grid structures facilitate equalized distribution of flow and pressure under varying demand rates and locations in water distribution systems, and to a limited extent, the meshedness coefficient illustrates the hydraulic efficiency of the network [10, 11, 26].

The dead-end fraction is given by $\frac{n_d}{n_{all}}$, the fraction relating the total numbers of nodes and dead-end nodes. A dead end is a point in the network that is fed only from one end. A typical example would be at the end of a cul-de-sac in a residential area. The fraction of dead ends is normal higher in treelike networks where the main pipeline supplies the entire community, which is quite risky.

## 2.3 Scenarios Design for Topological Metrics and Models

To build an index for choosing an appropriate water-quality model, this thesis presents four simulation scenarios for dead-end mains. The present analysis is first applied to simulate the concentrations of free chlorine with different four models in the dead-end within one network. The simulators update the residual chlorine concentrations every hour. The four models all use the same duration, water quality timestep, reaction rate, and initial concentration of chlorine. Table 2 for analysis details for each simulation. The reaction time is set for 72 hours to make sure that each model is steady-state. For dead-end branches, is most accurate when it considers the incomplete mixing in pipe junctions and the advection-dispersion reaction in pipe transport. Using WUDESIM-BAM as a reference, we preform regression analysis on the results of the other three models.

Table 2 Parameters of each Simulation

| Parameters | Value |
|---|---|
| Duration | 72 hours |
| Water Quality Timestep | 1 min |
| Global Bulk Coefficient | -0.5 $days^{-1}$ |
| Global Wall Coefficient | -0.5 m/day |
| Order of Wall Reaction | 1 |
| Order of Bulk Reaction | 1 |
| Initial Chlorine Concentration | 4.0 mg/L |

Table 3 Simulation Scenarios

| Simulation Number | Model | Pipe Junction Mixing Assumption | Pipe reactions Mixing Assumption |
|---|---|---|---|
| 1 | EPANET | Complete Mixing | Advection |
| 2 | EPANET-BAM | Incomplete Mixing | Advection |
| 3 | WUDESIM | Complete Mixing | Advection-Dispersion |
| 4 | WUDESIM-BAM | Incomplete Mixing | Advection-Dispersion |

## 2.4 Network Database

To evaluate the correlation between the water-quality models and network properties, the thesis examines 10 different networks, including either hypothetical water distribution systems or actual systems from the ASCE Task Committee on Research Databases for Water Distribution Systems. The actual systems include two widely used network, the Cherry Hills/Brushy Plains service area of the South-Central Connecticut Regional Water Authority and the Example 3 Network of EPANET. Both networks are benchmarks for many studies [20]. The ten networks, whose system components are listed in Table 4, have been used by many researchers to verify water quality models in distribution systems [10,16,18]. Diagrams of each network are shown in Figure 2.

Table 4. System Components

| Network | Number of Pipes | Number of Junctions | Number of Pumps | Number of Tanks | Number of Reservoirs | Number of Dead Ends |
|---|---|---|---|---|---|---|
| KY 1 | 903 | 777 | 8 | 4 | 3 | 188 |
| KY 2 | 595 | 513 | 3 | 3 | 2 | 163 |
| KY 3 | 344 | 939 | 5 | 3 | 3 | 36 |
| KY 4 | 1,118 | 409 | 2 | 4 | 1 | 255 |
| KY 5 | 498 | 763 | 11 | 3 | 4 | 114 |
| A-town | 43 | 22 | 3 | 2 | 1 | 0 |
| C-town | 429 | 389 | 11 | 7 | 1 | 70 |
| Net 2 | 41 | 36 | 0 | 1 | 0 | 4 |
| Net 3 | 117 | 92 | 2 | 3 | 2 | 11 |
| Jilin | 34 | 27 | 0 | 0 | 1 | 4 |

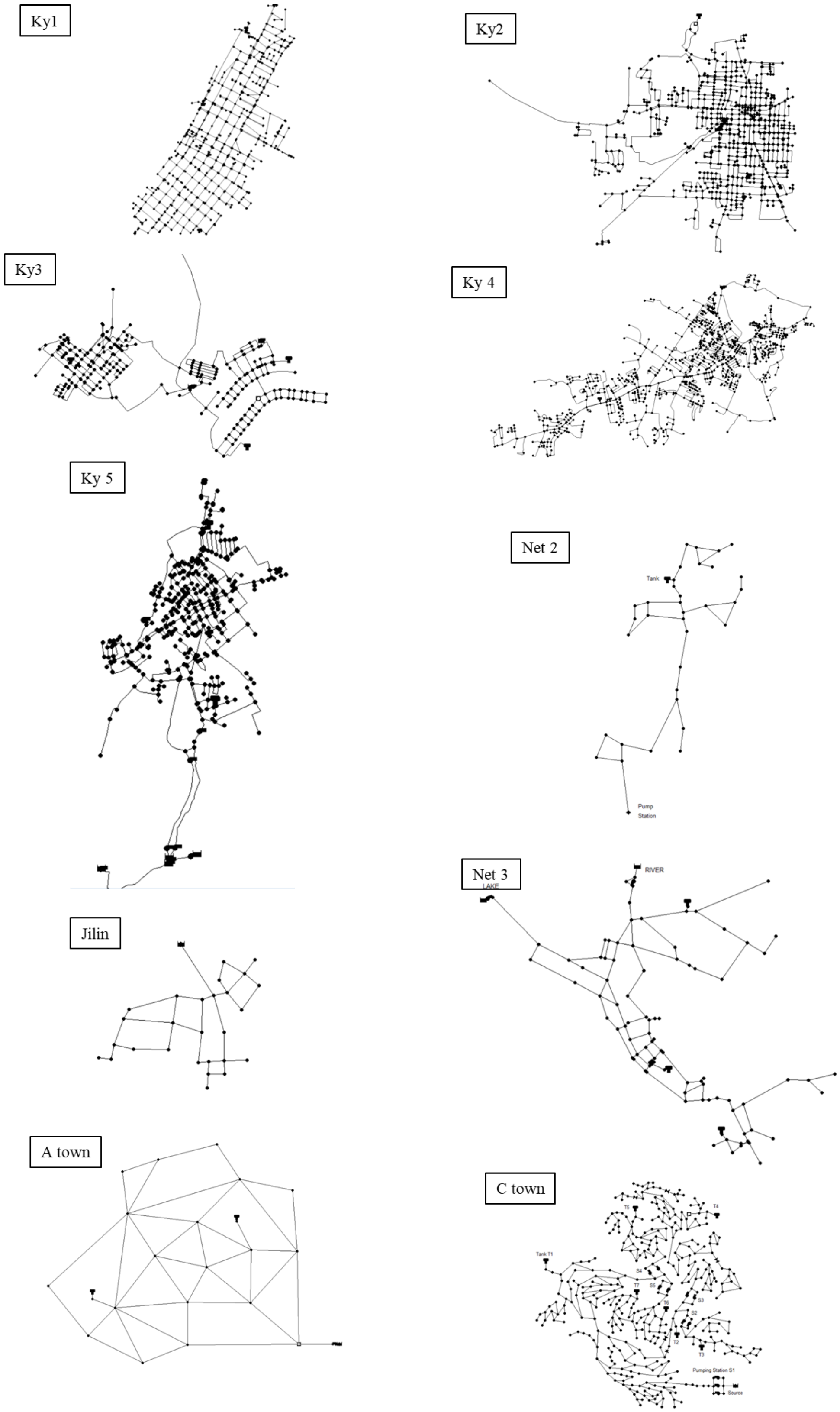


Figure 2. The structure of each network

# Chapter 3: Results and Discussion

## 3.1 Model Evaluation

To evaluate the quality of a model's predictions, outcomes can be designed with a variety of scoring parameters. For our models, the scoring parameters include the mean absolute error, RMSE, and explained variance. The scoring follows the convention that lower error values are better than higher error values. The R language is widely used among statisticians for calculating the scoring parameters of a model. In this thesis, we used R with the "scorer package" to calculate the mean absolute error, root mean squared error, and explained variance.

The mean absolute error (MAE), $MAE = \frac{\sum_{n=0}^{n} |y_{ref} - y_{pre}|}{n}$, is a measure of the difference between the reference and the prediction. Here $y_{ref}$ is a result from WUDESIM-BAM and $y_{pre}$ is a result from one of three models: EPANET, EPANET-BAM, or WUDESIM. In Figure 3, the y value of each node is the average mean absolute error of all dead ends. WUDESIM has the smallest MAE of all the networks except Net 3.

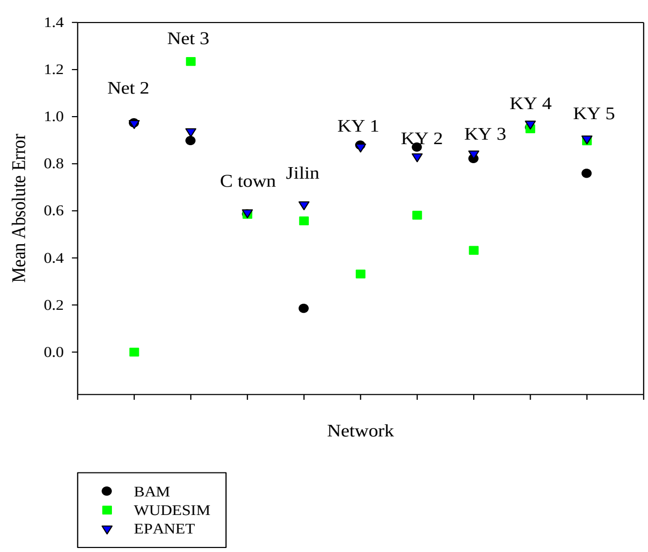


Figure 3. Mean absolute errors of EPANET-BAM, WUDESIM, and EPANET

The root mean squared error (RMSE), $\mathrm{RMSE} = \sqrt{\frac{\sum_{n=0}^{n}(y_{ref,t} - y_t)^2}{n}}$, is a quadratic scoring rule that measures the average magnitude of the error. It is the square root of the average of the squared differences between the predicted and actually observed values. In Figure 4, the y value of each node is the average RMSE value of the dead ends. We observe that EPANET-BAM has the largest RMSE. Since the errors are squared before they are averaged, the RMSE gives a relatively high weight to large errors, which means the RMSE should be more informative when large errors are particularly undesirable.

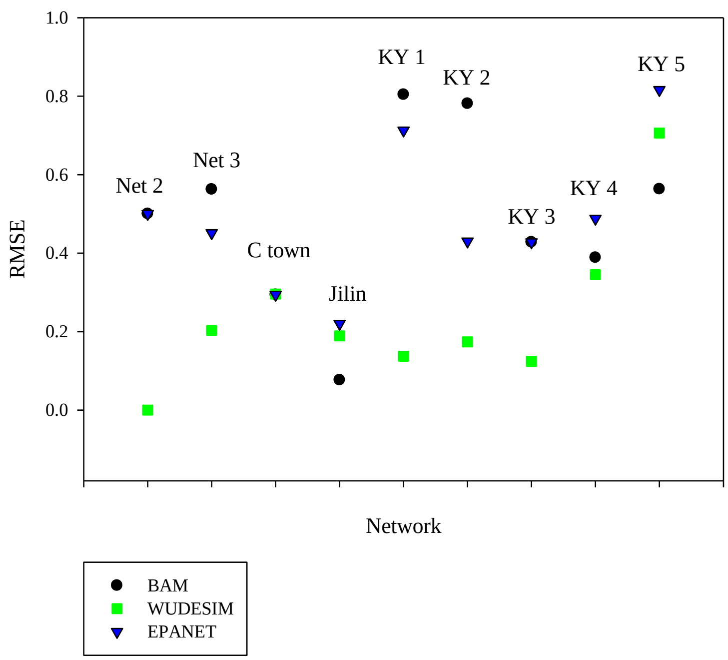


Figure 4. Root mean squared errors of EPANET-BAM, WUDESIM, and EPANET

The explained variance score measures the degree to which a mathematical model accounts for the variance of a given data set. It compares the variance within groups of the data set to the variance between the groups. In Figure 5, the y value of each node is the accumulated value of the explained variance of all dead ends. We observe that for the simple networks Net2, Net3, and Jilin, the explained variance scores are much smaller than for the more complex networks. Also, for complex networks, the variances of EPANET-BAM become larger, especially in KY 2, where the variance of EPANET-BAM is around 80% higher than that of EPANET.

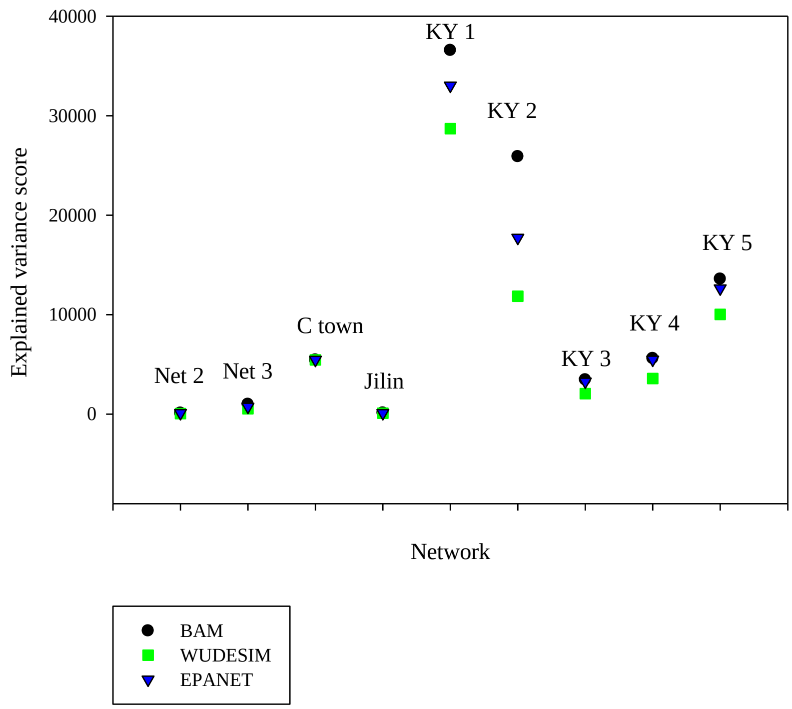

Figure 5. Explained variance scores of EPANET-BAM, WUDESIM, and EPANET

## 3.2 Water Quality Analysis in Dead Ends

Regression analysis is widely used for estimating the relationships between a dependent variable and independent variables. The goal of regression analysis is to build a mathematical equation that yields an accurate estimate. Linear regression is the most simple and popular technique for predicting a continuous variable. It assumes a linear relationship between the dependent variable and the independent variables.

In this section, ten networks are analyzed for each of four models: EPANET, EPANET-BAM, WUDESIM, and WUDESIM-BAM. There are two special cases for the four models. First, if the network does not have cross junctions, the results of EPANET and EPANET-BAM are the same. Because the results of EPANET and EPANET-BAM are used as the boundary conditions for WUDESIM and WUDESIM-BAM, if the results of EPANET and EPANET-BAM are the same, the results of WUDESIM and WUDESIM-BAM are also the same. Second, if the network does not have dead ends, the results of EPANET and WUDESIM are the same, and EPANET-BAM and WUDESIM-BAM also have the same results. Using WUDESIM-BAM as a reference, we compare the RMSE values of the models for each of the 10 networks. The RMSE indicates the absolute fit of the model to the data, that is, how close the observed data points are to the model's predicted values. Lower RMSE values indicate better fit. Because the RMSE is a good measure of how accurately the model predicts the response, it is the most important criterion for assessing goodness of fit where the main purpose of the model is prediction [27].

For each dead end, the measured duration time is 72 hours, and the chlorine concentration is reported every hour. Thus, the value of n is 72, $y_{ref}$ is the concentration of WUDESIM-BAM at time t, and $y_t$ is the concentration in the other three models at time t. From Figure 6 to Figure 15, WUDESIM has the smallest RMSE for each simulation because the dead-end mains are low flow and the mass transport is dominated by diffusion in the pipes. We also found that considering only bulk advection mixing in networks may cause a larger RMSE than otherwise. So, for a network with many dead ends, EPANET-BAM may not be suitable. In EPANET-BAM, incomplete mixing is applied to a cross junction which has a high Reynolds number (Re). In Ho and O’Hern’s research, they found that for large vertical pipes (~52 mm), incomplete mixing happened at Re from 3,000 to 12,000, and for small vertical pipes (~26 mm), incomplete mixing happened at Re from 3,000 to 9,000. Using the results from the bulk advective mixing model in dead ends, the errors accumulate in dead-end mains because in dead ends, the Re is very low, sometimes even close to 0 [20]. Additionally, the verifications of EPANET-BAM used laboratory-scale experimental data to fit the EPANET-BAM model. But, in a complex network, we also found mass imbalances.

Table 5. RMSEs of EPANET-BAM, WUDESIM, and EPANET

| Network | BAM | WUDESIM | EPANET |
|---|---|---|---|
| Net2 | 0.500 | 0.000 | 0.500 |
| Net3 | 0.563 | 0.203 | 0.451 |
| Ctown | 0.294 | 0.296 | 0.294 |
| Jilin | 0.077 | 0.189 | 0.220 |
| Ky1 | 0.804 | 0.137 | 0.713 |
| Ky2 | 0.781 | 0.174 | 0.430 |
| Ky3 | 0.428 | 0.124 | 0.428 |
| Ky4 | 0.389 | 0.345 | 0.488 |
| Ky5 | 0.563 | 0.706 | 0.816 |

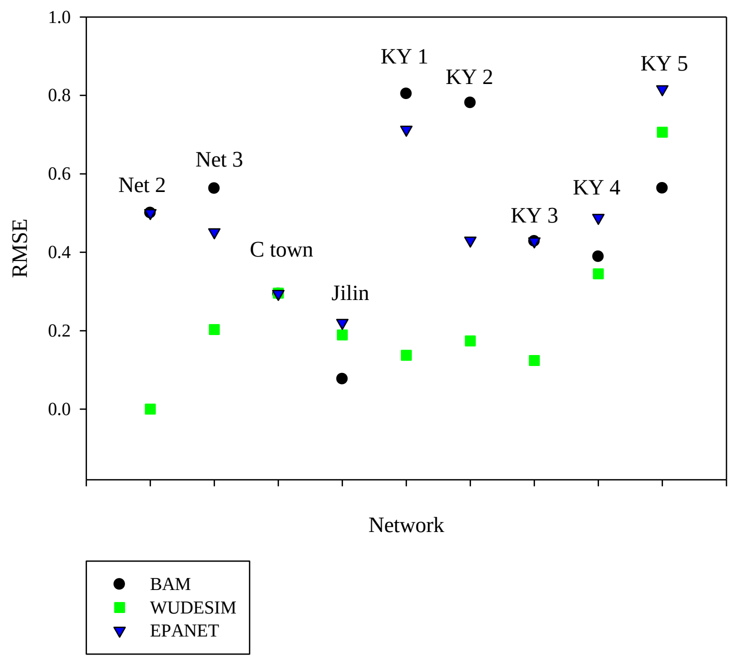


Figure 6. RMSE of EPANET-BAM, WUDESIM, and EPANET

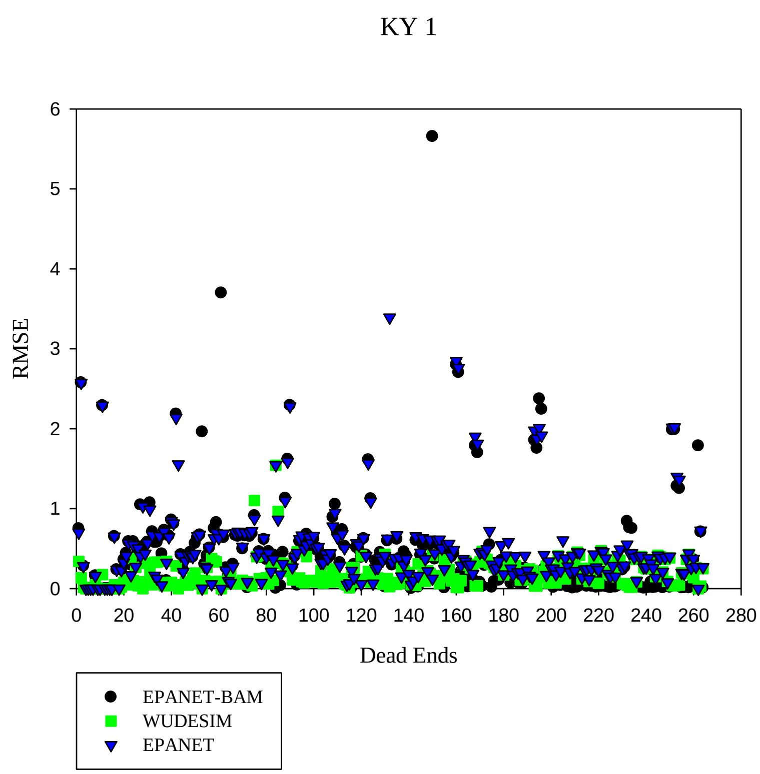


Figure 7. RMSE of Network KY 1 for Each Dead End

KY 2

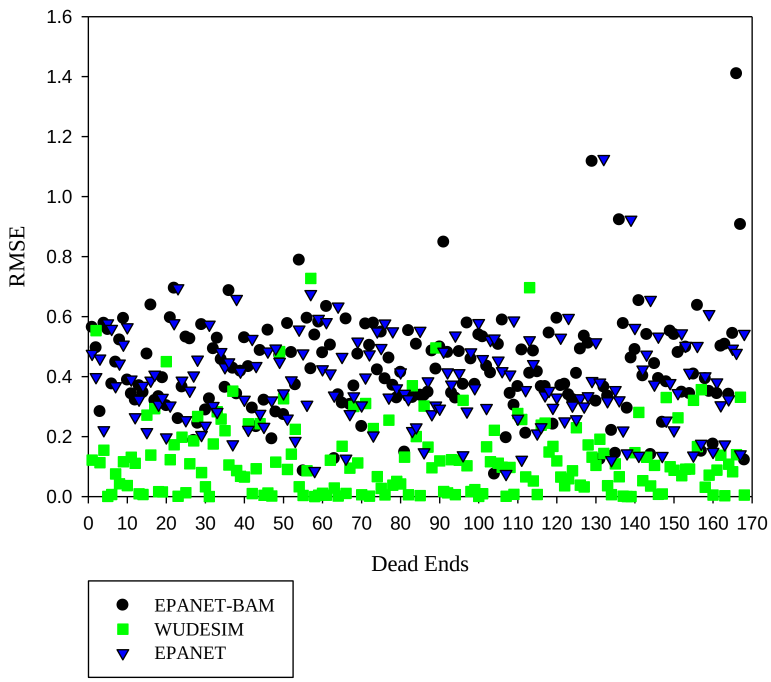


Figure 8. RMSE of Network KY 2 for Each Dead End

KY 3

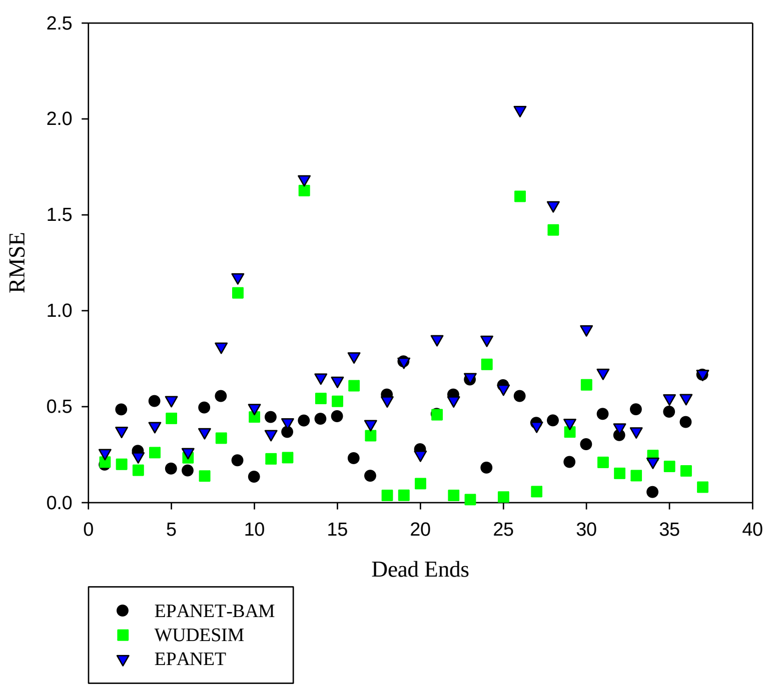


Figure 9. RMSE of Network KY 3 for Each Dead End

KY 4

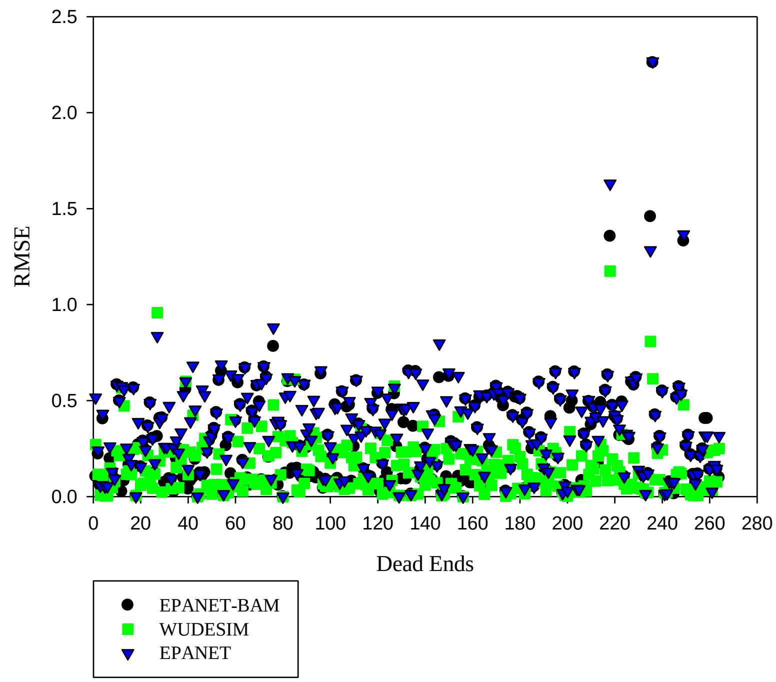


Figure 10. RMSE of Network KY 4 for Each Dead End

KY 5

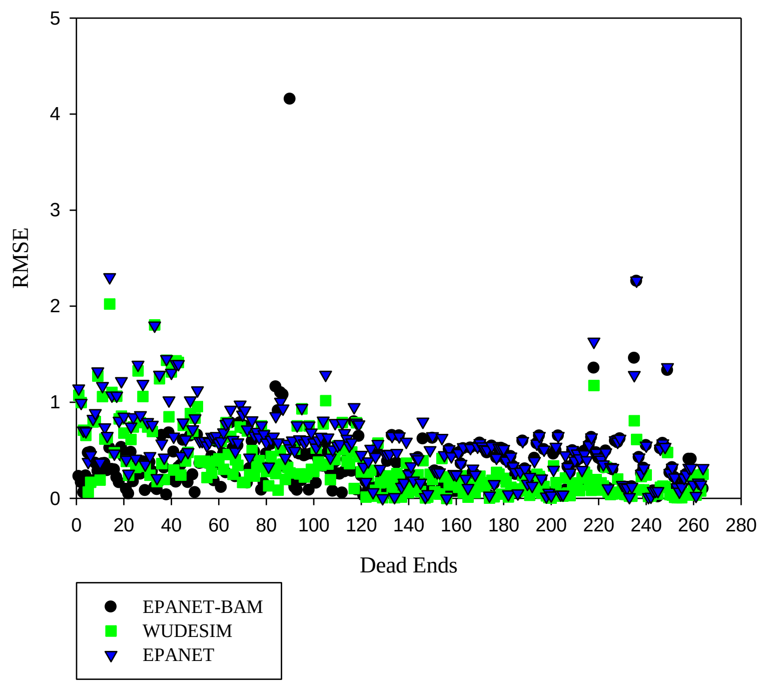


Figure 11. RMSE of Network KY 5 for Each Dead End

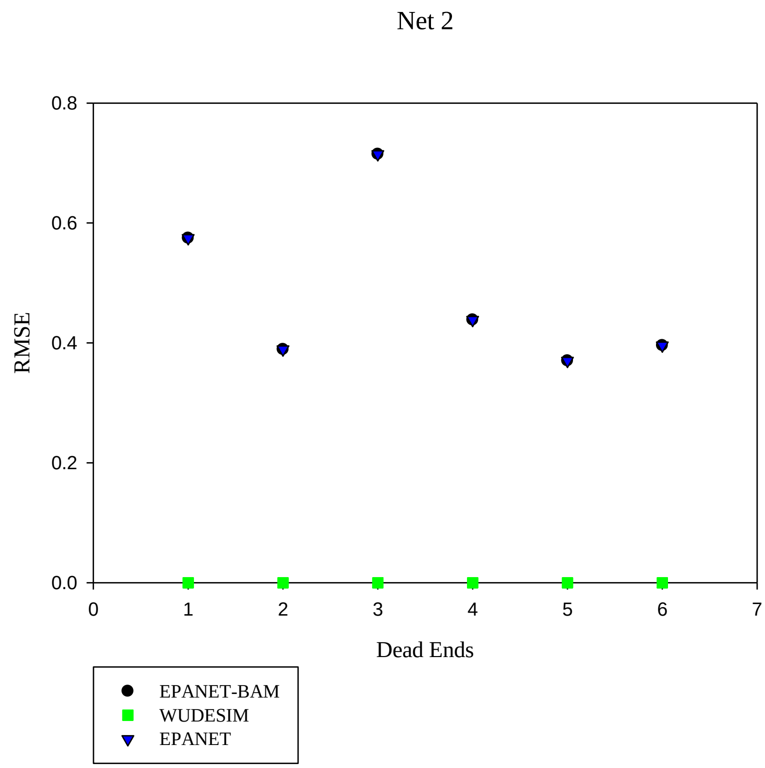


Figure 12. RMSE of Network Net 2 for Each Dead End

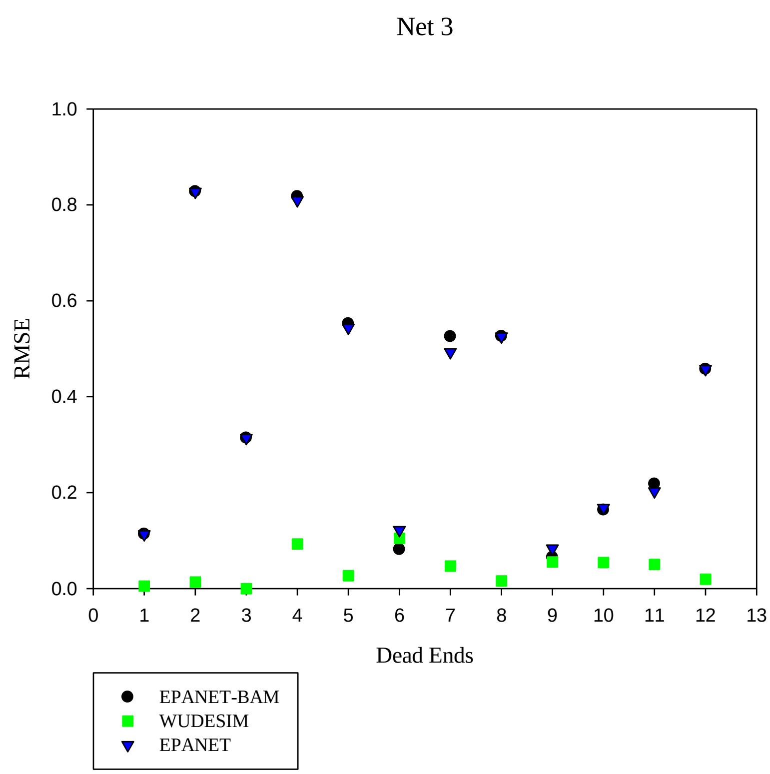


Figure 13. RMSE of Network Net 3 for Each Dead End

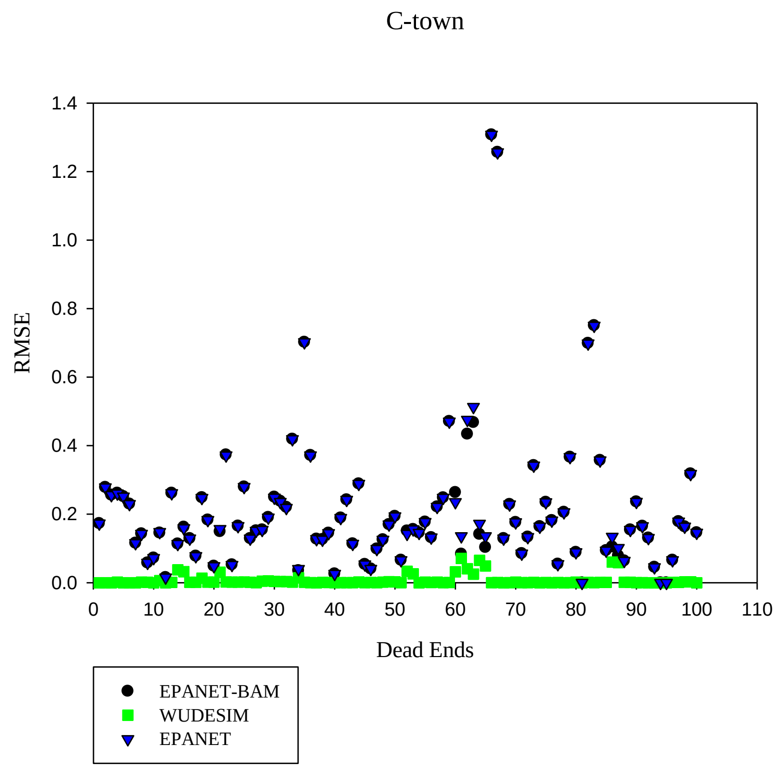


Figure 14. RMSE of Network C town for Each Dead End

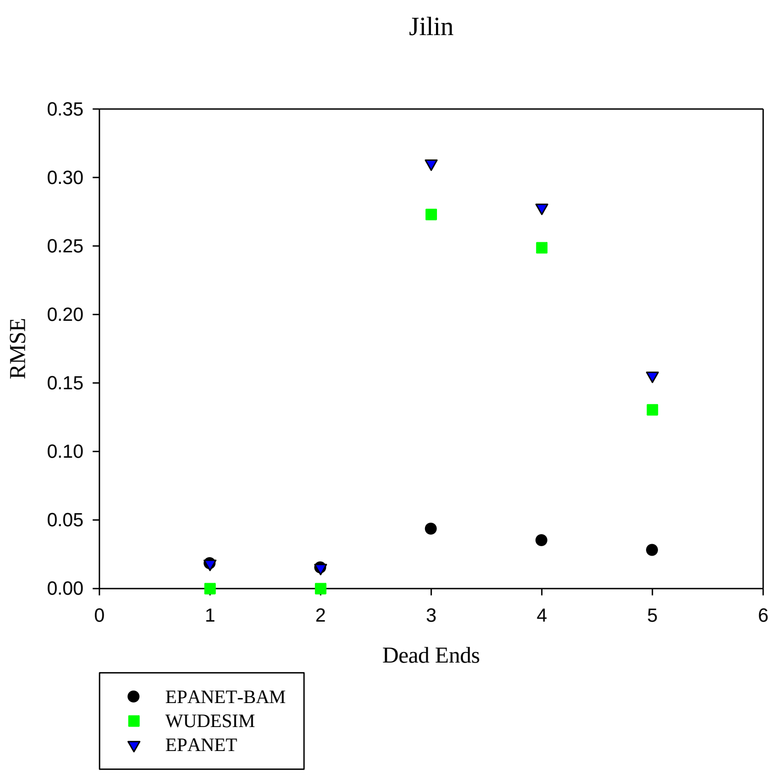


Figure 15. RMSE of Network Jilin for Each Dead End

## 3.3 Topological Properties

Topology has been assumed to have a great impact on the resilience of water distribution systems, and it is the object of many assessments [11]. Based on its structural patterns and network building blocks, a water distribution network can be represented as nodes and links, a process which can quantify the organizational properties of the network,. Using metrics for these topological properties, we can evaluate the properties of networks.

As illustrated in Figure 2, this thesis analyzed ten water distribution networks: seven American networks, one Chinese network (Jilin), and two imaginary networks from the literature (A town and C town). The ten networks were chosen from a variety of networks with different sizes and structures, providing some diversity their analyses. Their calculated topological properties are presented in Figures 13 and 14 and Table 6. As for the link density, KY 1, KY 2, KY 3, KY 4, KY 5, and C town demonstrated low link densities. The link density measures how close the number of links in a network is to the maximum possible number for a given number of nodes. If the value is close to 0, the networks are sparse. Thus, KY 1, KY 2, KY 3, KY 4, KY 5, and C town can be regarded as sparse networks [11, 28, 29].

As for the meshedness coefficient, the A town network has the highest value, 0.564. The meshedness coefficient ranges from 0 for trees to 1 for maximal planar graphs. It reflects the overall topological similarity of the network to perfect grids or lattice-like structures [11, 28, 29]. The shape of A town, a simple network without any dead ends, explains this high coefficient. Other networks have values from around 0.1 to 0.2. C town has the lowest meshedness coefficient, 0.053, which means the structure is closer to a treelike structure.

As for the average degree, the A town network has the highest value, 3.909. The average degree is a measure of connectivity, reflecting the overall topological similarity of the network to perfect grids or lattice-like structures. In our simulations, the maximum nodal degree is 4, and it means that A town is very close to a grid structure. In other networks, the average degrees are around 2 because most of the pipes are not cross junctions or tees.

As for the dead-end fraction, KY 5 has the highest value, 0.273. The dead-end fraction reflects how many dead ends are in a network. More dead ends in a network increase the possibility of unsuitable drinking water. The A town system has the lowest value, 0, as there are no dead ends.

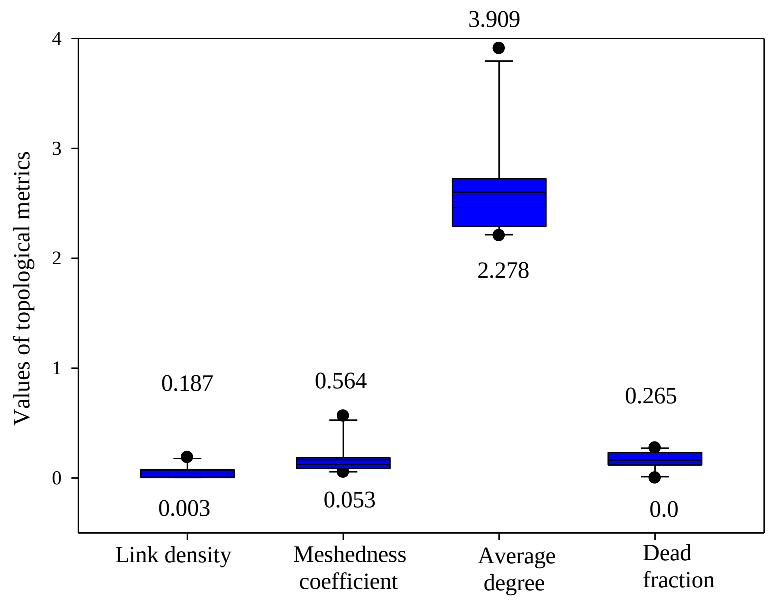


Figure 16. Values of four topological attribute metrics of the ten water distribution systems

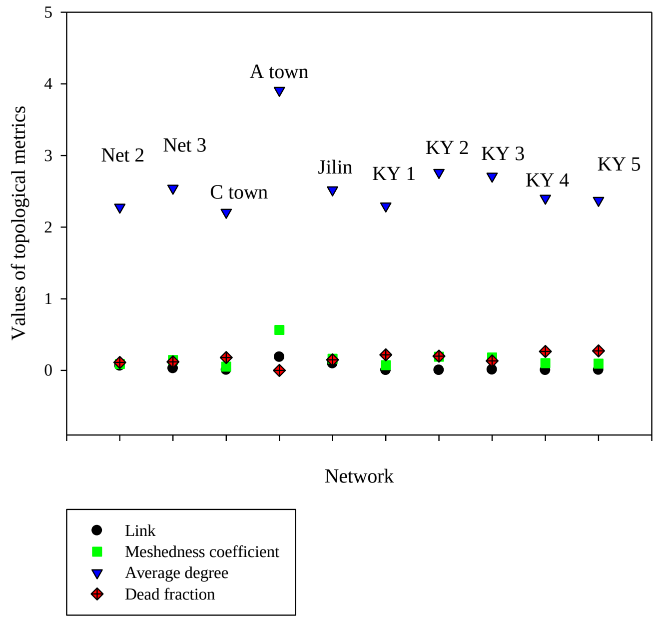


Figure 17. Topological attribute metrics of the ten water distribution systems

Table 6. Topological Attribute Metric Values for The Ten Networks

| Network | Link density | Meshedness coefficient | Average degree | Dead fraction |
|---|---|---|---|---|
| Net 2 | 0.065 | 0.090 | 2.278 | 0.111 |
| Net 3 | 0.028 | 0.145 | 2.543 | 0.120 |
| C town | 0.006 | 0.053 | 2.206 | 0.180 |
| A town | 0.186 | 0.564 | 3.909 | 0.000 |
| Jilin | 0.097 | 0.163 | 2.519 | 0.148 |
| Ky 1 | 0.003 | 0.074 | 2.294 | 0.219 |
| Ky 2 | 0.003 | 0.192 | 2.762 | 0.200 |
| Ky 3 | 0.010 | 0.181 | 2.711 | 0.133 |
| Ky 4 | 0.003 | 0.102 | 2.403 | 0.265 |
| Ky 5 | 0.006 | 0.095 | 2.373 | 0.273 |

### 3.4 Topological Metrics and Water-Quality Models

We can evaluate the properties of networks by using metrics for their topological properties. The goal of this section is to build a mathematical equation of topological metrics that yields an accurate prediction of water quality.

In this section, regression analysis is used to estimate the relationships between the water-quality models and their topological attribute metrics. The RMSE is used to correlate the link density, meshedness coefficient, average degree, and dead-end fraction. Figure 18 to Figure 21 are scatter plots for each model's RMSE versus the values of topological attribute metrics. It is hard to see a correlation between the RMSE and link density, meshedness coefficient, and average degree for EPANET-BAM, WUDESIM, and EPANET. The $R^2$ ranges from 0.02 to 0.05. For EPANET-BAM, the $R^2$ is 0.0195, reflecting only a small correlation between the dead-end fractions and RMSE. For EPANET, we find the $R^2$ is 0.2976, representing a medium correlation between the RMSE and the dead-end fractions. For WUDESIM, the $R^2$ of 0.5598 shows a strong correlation between the RMSE and the dead-end fractions [27]. Because a strong correlation means more dead-ends, it may affect WUDESIM more than WUDESIM-BAM. Complex water distribution systems also have more cross-junction pipes that may cause incomplete mixing, introducing further discrepancies between WUDESIM and WUDESIM-BAM. However, we still need more information about cross junction fractions to prove our hypothesis that topological properties affect water quality.

From the results, the topological properties do not appear strongly correlated with water-quality models. There could be several causes. First, the chosen topological metrics in this thesis concern only structural basic properties. They do not describe the connectivity properties of the water distribution system [28]. Second, ten networks is too small a number to fully represent the correlations. Meng et al. used 80 networks to propose the correlation between topological properties and resilience, although fewer than 80 might have been sufficient.

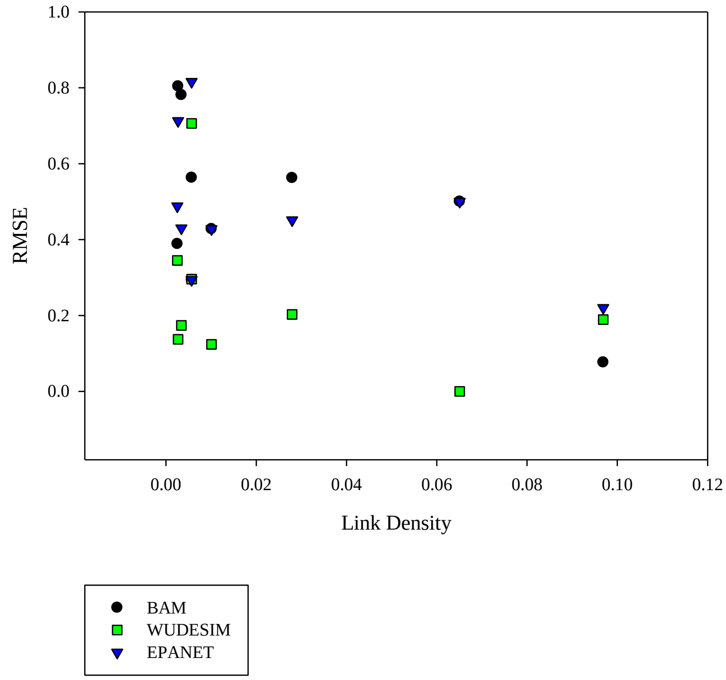


Figure 18. RMSE versus link density

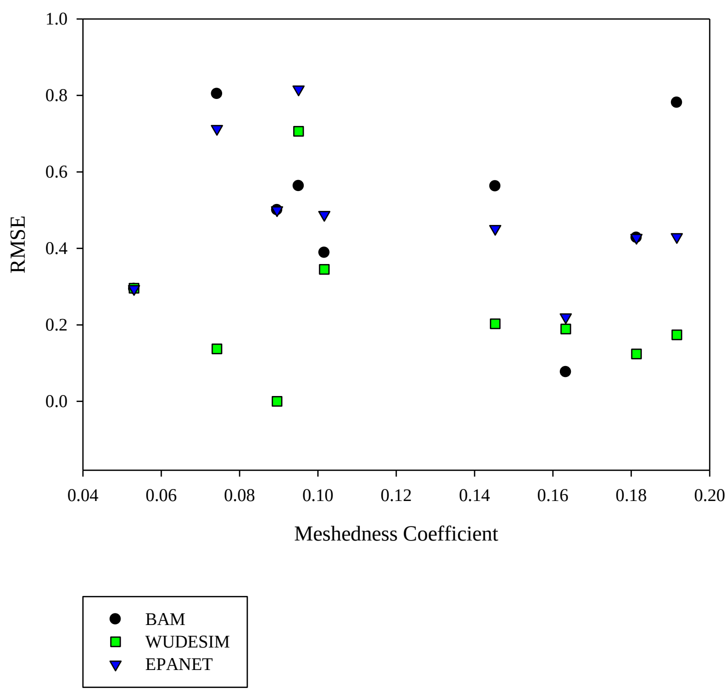


Figure 19. RMSE versus meshedness coefficient

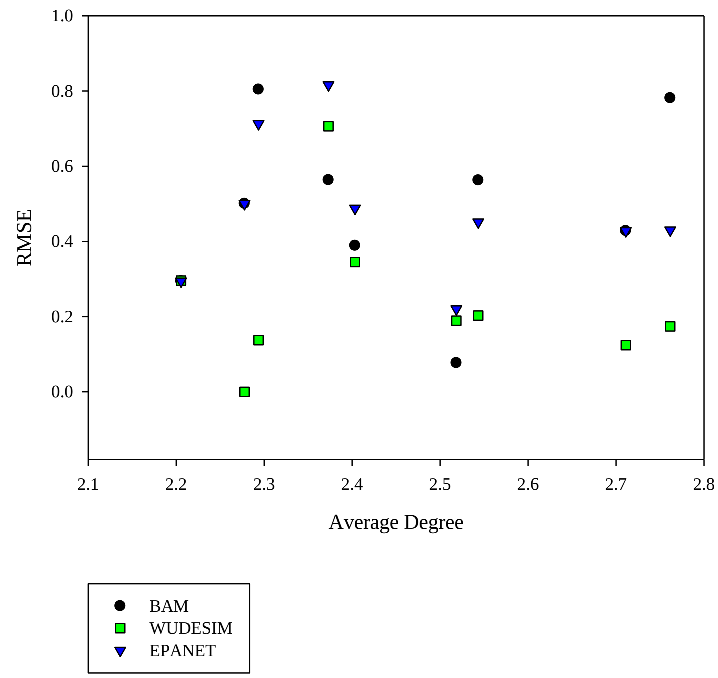


Figure 20. RMSE versus link average degree

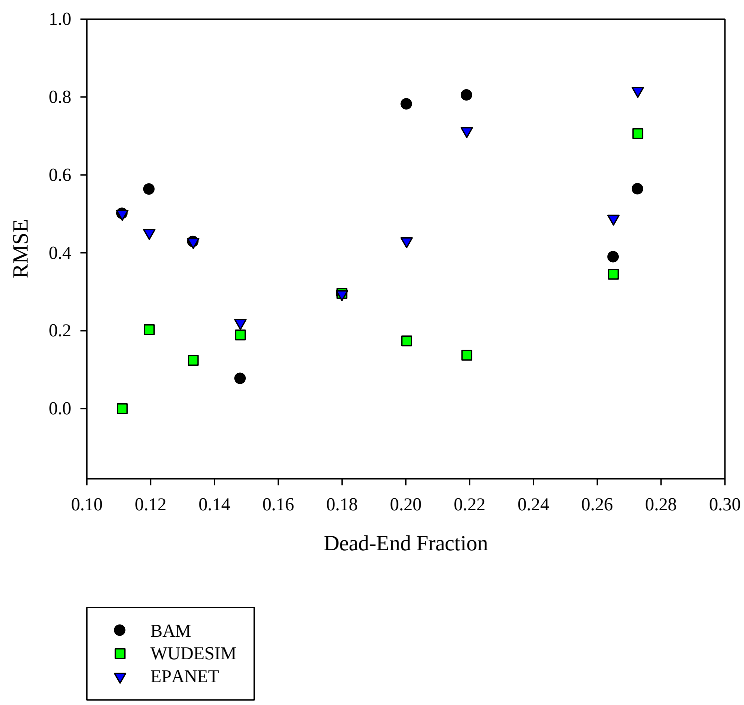


Figure 21. RMSE versus dead-end fraction

# Chapter 4: Conclusions and Future Work

In this work, a numerical model, WUDESIM-BAM, was developed to simulate chlorine residuals of dead ends in water distribution systems. To illustrate the relationship between water-quality and topological properties, the topological properties of four water-quality models we reexamined: EPANET, EPANET-BAM, WUDESIM, and WUDESIM-BAM.

The new model, WUDESIM-BAM, is initially assumed to accurately predict the water quality in dead ends, measured by residual free chlorine. Using WUDESIM-BAM as the reference, scoring parameters are used to rank the accuracy of the three other models' predictions in terms of mean absolute error. The results demonstrate that WUDESIM has the smallest mean absolute error difference from the reference. The results further demonstrate that EPANET-BAM has the largest RMSE. Because EPANET-BAM also has the largest explained variance, and WUDESIM has the smallest explained variance with regard to dead end predictions, WUDESIM is the most accurate of the four models. EPANET-BAM is not appropriate because low flow in a dead end can cause misleading results for this bulk flow model.

In a topological attribute metric analysis, we find little correlation between water-quality models and topological properties, for several possible reasons. First, the chosen topological metrics concern structural properties. They do not measure the connectivity of the water distribution system [11, 28, 29]. Second, ten networks may be too few to properly represent the correlations.

This thesis creates an index for a water-quality model. However, more research is needed to fully define the relationship between topological properties and water-quality models. In the future, people can use just a simple index and easily choose an appropriate model. Similarly, the Environmental Protection Agency could more easily determine with appropriate standard.

# References

[1] J. L. Brown, Drinking the Mississippi: The Chain of Rocks Water Purification Plant, Civil Engineering Magazine Archive 85 (2015).

[2] City of St. Louis Water Division, History timeline, http://www.stlwater.com/about/history (2019).

[3] National Research Council, Drinking Water Distribution Systems: Assessing and Reducing Risks, (2006).

[4] American Society of Civil Engineers (ASCE), Infrastructure Report Card: Drinking Water,(2017).

[5] American Water Works Association, Buried No Longer: Confronting America's Water Infrastructure Challenge, (2012).

[6] K.P. Fedinick, M. Wu, M. Pandithartne, E. Olson, Threats on Tap: Widespread Violations Highlight Need for Investment in Water Infrastructure and Protections, NRDC-Report. (2017).

[7] Centers for Disease Control and Prevention, Current Waterborne Disease Burden Data & Gaps, https://www.cdc.gov/healthywater/burden/current-Data.html (Accessed 11/01/2019).(2017).

[8] National Research Council. 1980. Drinking Water and Health: Volume 2. Washington, DC: The National Academies Press. https://doi.org/10.17226/1904.

[9] M.R. Islam, and M.H. Chaudhry, Modeling of Constituent Transport in Unsteady Flows in Pipe Networks, J. Hydraul. Eng. ASCE. 124 (1998) 1115 – 1124.

[10] A. Yazdani and P. Jeffrey, Complex Network Analysis of Water Distribution Systems, AIP. Chaos 21 (2011), 016111.

[11] F. Meng, G. Fu, R. Farmani, C. Sweetapple, D. Butler. Topological Attributes of Network Resilience: A Study in Water Distribution Systems, Water Research 143 (2018) 376–386.

[12] M. Rubinov, O. Sporns. Complex Network Measures of Brain Connectivity: Uses and Interpretations, NeuroImage. (2010), 52, 1059–1069

[13] World Health Organization, A Global Overview of National Regulations and Standards for Drinking-Water Quality. (2018). Licence: CC BY-NC-SA 3.0 IGO.

[14] United States Environmental Protection Agency (USEPA), Understanding the Safe Drinking Water Act, (2004).

[15] Y. Shao1, L. Zhao, Y. J. Yang, T. Zhang, T. Yu5, Experimental Observation on Solute Mixing of Laminar and Transitional Flow at Junctions in Water Distribution Systems, Advances in Civil Engineering. (2019), Article ID 3686510

[16] Yu, T., H. Qiu, J. Yang, Y. shao, L. Tao, Mixing at Double-Tee Junctions with Unequal Pipe Sizes in Water Distribution Systems. Water Science and Technology: Water Supply. IWA Publishing, (2016), 16(6):1595-1602.

[17] P. Romero-Gomez1, C. K. Ho, C. Y. Choi, Mixing at Cross Junctions in Water Distribution Systems.I: Numerical Study, J. Water Resour. Plann. Manage. 134 (2008) 285-294

[18] C.K Ho, and L. O'Rear, Evaluation of Solute Mixing in Water Distribution Pipe Junctions, American Water Works Association, 101 (2009)

[19] M.E. Tryby, D.L. Boccelli, M.T. Koechling, J.G. Uber, R.S. Summers, L. Rossman, Booster Chlorination for Managing Disinfectant Residuals, J. Am. Water Works Assoc. 91 (1999) 95–108.

[20] A.A. Abokifa, Y.J. Yang, C.S. Lo, P. Biswas, Water Quality Modeling in the Dead End Sections of Drinking Water Distribution Networks, Water Res. 89 (2016) 107–117.

[21] P. Biswas, C. Lu, R.M. Clark, A Model for Chlorine Concentration Decay in Pipes, Water Res. 27 (1993) 1715–1724.

[22] L.A. Rossman, R.M. Clark, W.M. Grayman, Modeling Chlorine Residuals In Drinking Water Distribution Systems, J. Environ. Eng. 120 (1994) 803–820.

[23] D.H. Axworthy, B.W. Karney, Modelling Low Velocity/High Dispersion Flow in Water Distribution Systems, J. Water Resour. Plan. Manag. 122 (1996) 218–221.

[24] V.G. Tzatchkov, A.A. Aldama, F.I. Arreguin, Advection-Dispersion-Reaction Modeling in Water Distribution Networks, J. Water Resour. Plan. Manag. 128 (2002) 334–342.

[25] V.C.F. Dias, M.C. Besner, M. Prévost, Predicting Water Quality Impact After District Metered Area Implementation in a Full-Scale Drinking Water Distribution System, J. Am. Water WorksAssoc. 109 (2017) E363–E380.

[26] A. D. Nardo, M. D. Natale, C. Giudicianni, D. Musmarraa, J.M. Rodriguez Varela, G.F. Santonastaso, A. Simone, V. Tzatchkov, Redundancy Features of Water Distribution Systems, Procedia Eng. 186 ( 2017 ) 412 – 419

[27] scikit-learn Machine Learning in Python, sklearn.linear_model.LinearRegression, https://scikitlearn.org/stable/modules/generated/sklearn.linear_model.LinearRegression.html (2017).

[28] A. Yazdani, R. A. Otoo , P. Jeffrey, Resilience Enhancing Expansion Strategies for Water Distribution Systems: A Network Theory Approach, Environmental Modelling & Software 26 (2011) 1574 – 1582

[29] E. Todini, Looped Water Distribution Networks Design Using a Resilience Index Based Heuristic Approach, Urban Water 2 (2000) 115 – 122

# CHUN-YING CHAO

212 S. Meramec Ave., Apt. 1106, Clayton, MO 63130 | 314-224-9641 | c.chao@wustl.edu

## SUMMARY

I would like to develop a new air quality index with topological analysis to address the air quality issues. My research particularly focuses on calibrating air quality data from low-cost particulate matter (PM) sensors and PM monitor stations. My research will examine whether and how graph theory can calibrate an air quality index.

## EDUCATION

**Washington University Department of Energy, Environmental & Chemical Engineering,** St. Louis, MO
MA.Sc., expected 2019

*Honors:* Washington University Scholarships and Grants – awarded USD $60,000 for academic excellence

*Thesis:* *Modeling the Effects of Distribution System Topology on Water Quality*

*Publications:* **Integrating Low-cost Air Quality Sensor Networks with Fixed and Satellite Monitoring Systems for Enhanced Accuracy, Reliability, and Applicability,** Jiayu Li, Huang Zhang, **Chun-Ying Chao**, Chih-Hsiang Chien, Chang-Yu Wu, Cyuan Heng Luo, Ling-Jyh Chen, Pratim Biswas, Environmental Science & Technology Letters (Submitted April 2nd, 2019)
**User Manual for Washington University Dead-End Simulator**, **Chun-Ying Chao**
**Catalytic Removal of Trace Oxygen from Oxy-Coal Combustion Flue Gas over CoOx Catalysts: The Significant Role of Oxygen Vacancy**, Liang-Yi Lin, Sungyoon, Jeon, **Chun-Ying Chao**, Pratim Biswas (in progress)

*Activities:* Taiwanese Graduate Student Association, Vice President
Research Assistant, Professor Pratim Biswas

**National Cheng Kung University,** Tainan, Taiwan
BSc. in Chemical Engineering, June 2015

*Publications:* **Corrosion Resistance of Electroless Ni-W-P Alloy Prepared in Different Conditions and Optimized by Response Surface Methodology,** Chao, Chun-Ying and Bing-Hung Chen, Chinese Institute of Engineers Students' Thesis Competition, 2015.

*Honors:* Mrs. May Jen Scholarship – awarded USD $1,700 for academic excellence (2011 – 2015)
Kinmen International Invention and Design Fair (KIIDF), Bronze Medal (2014)
National Cheng Kung University Engineering School Invention Award, Golden Medal (2014)
Honor student; ranked 2nd among 48 students (2011)

*Activities:* National Cheng Kung University Mixology Association, President
5th Career Coaching Program, Vice President
Taiwan-United States Sister Relations Alliance Summer Scholarship Program, School Ambassador

## EXPERIENCE

**Aerosol and Air Quality Research Laboratory,** Washington University, St. Louis, MO
*Research Assistant* Sep.
2018 – Present
Integrated low-cost PM sensor data with satellite aerosol optical depth (AOD) data and fixed station data. Optimized water distribution models using C++ and evaluated the models' robustness using R, including regression and error analysis. Applied topological attributes in water distribution systems. Synthesized oxygen removal catalysts in a furnace aerosol reactor.

**Emerging Contaminants Laboratory,** National Cheng Kung University, Tainan, Taiwan
*Research Assistant* Jan.
2017 – Jul. 2018
Integrated water quality data with agricultural activity data to assess the total phosphorus in a reservoir water source for EPA. Analyzed emerging contaminants in reservoirs and groundwater using HPLC-ICP/MS. Published an article: **Evaluation of the Impacts of Human Activity on Water Quality: A case Study in a Reservoir Catchment in Southern Taiwan**, Chia-Chen Chung, **Chun-Ying Chao** and Wan-Ru Chen, 16th IEEE, Tainan, Taiwan, 2017.

**586 Armor Brigade,** Taichung, Taiwan
*Corporal and Ombudsman of Artillery Aug.*
2015 – Oct. 2016
Disaster relief during typhoon season. Responsible for enforcing military law and maintaining military discipline.

**Surfactant Laboratory,** Tainan, Taiwan
*Independent Study Student* Sep.
2013 – June 2014
Synthesized Ni-W-P anti-corrosion alloy by electroless plating and optimized the precursor by response surface methodology. Tested electrochemical corrosion with potentiostat.

## AWARDS AND COMMUNITY INVOLVEMENT

**Dharma Drum Mountain Buddhist Association, Pure Mind Center,** St. Louis, MO
*Volunteer*
Aug. 2018 – Present
Organize and plan annual events, including fundraising; co-organize meditation practice

•

## LANGUAGES

Native Mandarin and Taiwanese; Fluent English

## TECHNICAL SKILLS

**Software and Language**

EPANET: A software of hydraulic and water-quality behavior within pressurized pipe networks.
Minitab: A statistics package provides statistical analyses for experiment, including design of experiments (DOE).
Chromeleon™ Chromatography Data System: A software for unifying chromatography instruments and data processing.
C++: Using C++ for modifying water-quality model WUDESIM.
R: Using R for regression analysis and assessing the robustness of the model.
MATLAB: Using MATLAB for numerical and matrix calculation and solving partial differential equations.

**Instrument**

Furnace Aerosol Reactors: An aerosol reactor with nebulizers which used for nanoparticle synthesis.
High Performance Liquid Chromatography (HPLC): Analyzing sub-ppb level contaminants from environmental samples.
Ionic Chromatography (IC): Analyzing contaminants having cations and anions form the steel corporation’s wastewater.
Inductively coupled plasma mass spectrometry (ICP-MS): Using to analysis the ppb level arsenic from ground water.
Spectrometers: Measuring the intensity of light relative to wavelength.
Potentiostat: Analyzing the Tafel equation to estimate the anti-corrosion properties of Ni-W-P alloys form.
Total Kjeldahl Nitrogen (TKN) analysis: Determining both the organic and the inorganic forms of nitrogen.

**Others**

Field Sampling: Design the sampling in reservoir catchment area during typhoon season.