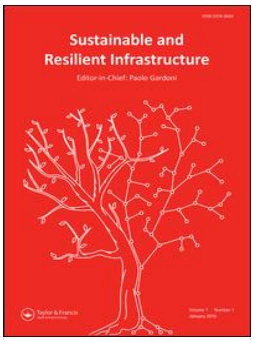

Sustainable and Resilient Infrastructure

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tsri20

A synthetic water distribution network model for urban resilience

Nasir Ahmad , Mikhail Chester , Emily Bondank , Mazdak Arabi , Nathan Johnson & Benjamin L. Ruddell

SUSTAINABLE AND RESILIENT INFRASTRUCTURE
https://doi.org/10.1080/23789689.2020.1788230

A synthetic water distribution network model for urban resilience

Nasir Ahmad [a], Mikhail Chester [a], Emily Bondank [a], Mazdak Arabi [b], Nathan Johnson [c] and Benjamin L. Ruddell [d]

[a]School of Sustainable Engineering and the Built Environment, Arizona State University, Tempe, AZ, USA; [b]Civil and Environmental Engineering, Colorado State University, Fort Collins, CO, USA; [c]The Polytechnic School, Arizona State University, Mesa, AZ, USA; [d]School of Informatics, Computing, and Cyber Systems, Northern Arizona University, Flagstaff, AZ, USA

ABSTRACT

Water distribution networks (WDN) are one of the most critical infrastructures, providing water for essential needs. However, the dearth of information on WDNs due to weak historical records, limited willingness to share data, and security concerns limit a researcher's understanding of the criticality, adaptability, vulnerability, and interdependencies of WDNs. To address this challenge, we develop a model entitled SyNF (Synthetic Infrastructure) for synthetic WDN generation. SyNF uses a roadway network, water demand, and water source locations to synthesize topology, diameter, and service year of pipes, and location and power requirements of pumps. To show SyNF's capabilities, we start with the City of Tempe and scale the model to Phoenix metro's seven major cities. We find a 6% average dissimilarity on pipe size distribution between the original and synthesized WDN in validating SyNF. We also discuss how SyNF can advance our understanding of the criticality, vulnerability, and resilience of WDNs.

ARTICLE HISTORY
Received 23 January 2020
Accepted 22 June 2020

KEYWORDS
Synthetic network; water distribution network; urban resilience; infrastructure

1. Introduction

Safe drinking water is vital for the protection of public health, making the infrastructure that supports the provision of water services critical. In the U.S., more than 80% of the population depends on approximately 153,000 public drinking water systems for potable water, industry, and fire fighting (Department of Homeland Security, 2014). Like any critical infrastructure, water supply systems are interconnected: the systems not only depend on other critical infrastructures, but other critical infrastructures depend on water infrastructures (Frederic Petit et al., 2015). As such, a failure in one infrastructure can affect others. Thus, understanding the complexity, dependencies, and interdependencies of water infrastructure is vital for enhancing their protection and resilience. However, a primary barrier to advancing our understanding of water supply systems is that there is limited public information on physical layouts (e.g., pipe layout, pipe diameter) and operational characteristics (e.g., pump operation) for researchers to study. The criticality and sensitivity of water distribution system information has led authorities to be protective of data about the design and operations of the systems, for instance via protection under the U.S. Protected Critical Infrastructure Information (PCII) categorization.

The dearth of precise data on the topology and specifications of water distributions systems has resulted in published academic engineering research that is too conceptual and that is questionably accurate. Analysis of real city water infrastructures remains limited. Several theoretical frameworks have been established to understand the resilience (Francis & Bekera, 2014; Park et al., 2013), criticality (Clark et al., 2018; NIPP, 2013, 2013), adaptability (Chester & Allenby, 2018), and interdependencies (Frederic Petit et al., 2015; Petit, 2014; Rinaldi et al., 2001) of critical infrastructures, along with the impacts of human activity, climate change, and emerging technologies (Allenby & Chester, 2018) on them. While these conceptual studies provide awareness about the potential for vulnerabilities in systems, they do not identify specific challenges in specific systems, and are questionably accurate. For example, existing water-power interdependency studies elucidate the connection between water and power systems in terms of how much resources of one the other requires. The conceptual studies do not address what the impacts would be for a specific water system from power outages, and vice versa (Frederic Petit et al., 2015; Petit, 2014; Rinaldi et al., 2001). Thus, critical lines of inquiry remain unanswered: How vulnerable are water components from internal or external causes? How vulnerable are the water system and other connected

CONTACT Nasir Ahmad sk.nasir.ahmad@asu.edu

infrastructure systems to the failure of components? What can be done to mitigate this vulnerability? To address these questions in any city's water system, an authentic and spatially explicit model of the network topology, specifications, and basic operating characteristics are needed.

Identification of component and system-level vulnerabilities requires understanding the dynamics of component failure and the dynamics characterizing the effects that component failures have on each other in independent and coupled infrastructure systems. Often the spatial location of components relative to each other underlies the dynamics of failure and cascade of failure across systems (Rinaldi et al., 2001). For example, in Baltimore, an underground cargo train derailed and severed nearby water main in 2001, resulting in substantial outgates (Pederson et al., 2006). Furthermore, nearby power transformers and fiberoptic cables flooded, causing city-wide power blackouts and disruptions of cellular communications (Pederson et al., 2006). Thus, spatially-explicit information about components and their interconnection – in the form of directed networks of dependencies – is critical for understanding vulnerability. For water distribution systems, it is therefore essential to know the locations of reservoirs, pipes, pumps, and demands, which allows for more accurate estimation of the impacts of failure on consumers (via water pressure loss) and thus can improve understanding of what actions could prevent failures and impacts.

The shortage of primary data on the locations of components along with their characteristics (e.g., pipe diameter, power requirement of a pump) of water networks can be partially overcome by estimating synthetic networks. Approximate and estimated network geometry can provide much of the information needed to characterize network vulnerability and resilience. Publicly available data on other infrastructure networks such as roads and power lines, combined with engineering principles of water system design and hydraulics, along with resource usage context in the form of census data, can be used to estimate 'synthetic networks.' This is made possible because water distribution network (WDN) topology and demand have strong correlations with street networks and population density, respectively (Mair et al., 2017; Sharvelle et al., 2017). Methods of estimation can be developed specifically to approximate the authentic details needed to understand the real network's vulnerabilities. In this paper, we explore the possibilities and limitations of developing synthetic WDNs for cities and broader metropolitan regions.

There is no existing method to generate a synthetic WDN which includes a combination of a realistic spatially-explicit pipe layout, demands based on city-specific demographic, environmental, and policy data, and locations of pumping units – especially for the purpose of authentically representing vulnerability-relevant details. The majority of existing engineering algorithms focus on optimizing the design, or the effects of additions to WDNs given a prescribed network geometry, such as demand locations and amounts, supply and storage locations, and capacities, and pipe locations and diameters (Fujiwara & Ganesharajah, 1993; Loganathan et al., 1995; Park & Liebman, 1993; Savic & Walters, 1997). These algorithms elucidate how cost-optimized networks should be configured. There have been a few studies that develop algorithms to generate representations of current WDNs (Mair et al., 2014; Möderl et al., 2011; Sitzenfrei et al., 2010b; Sitzenfrei R. et al., 2011; Robert Sitzenfrei et al., 2010a). These studies develop 'virtual' networks that represent the average network parameters of real WDN's according to graph theory principles. This kind of virtual WDN is useful for estimating aggregated or summary characteristics of networks, like the capital costs of expansion or maintenance (Sitzenfrei et al., 2010b), but without spatial details on interdependency, the methods cannot facilitate the characterization of spatially-explicit vulnerabilities.

Studies that develop realistic synthetic water distribution networks tend to focus on the infrastructure layout without consideration of demand. For example, Mair and colleagues (Mair et al., 2014) developed procedures to estimate synthetic spatial pipe layouts as a spanning tree, based on road networks (Mair et al., 2014). However, the study did not incorporate realistic hydraulic processes and cost-effective principles that influence layouts and sizing of water mains, nor did they consider the industry design standards associated with gridded and looped structures. Water demand also influences the design of WDNs. Previous studies have only considered the service area population to estimate demands and not the neighborhood-scale differences that drive design. Moreover, existing algorithms focus on informing network topology (i.e., pipe layout) and pipe diameter but do not estimate pump location and power requirements. A more complete system.

This paper develops a model and analysis tool for synthetic WDN generation that considers supply and demand information at neighborhood scales to estimate WDNs up to metro region scales. A new model is developed that not only estimates spatially explicit links (pipe location and diameter) in the network but also simulates network characteristics during operation, which satisfies the design standards for a minimum required pressure and minimum pipe diameter for fire flow. These characteristics include flow through pipes, probable locations of pumps, power requirements for pumps, and pipe initial year of construction. We title this model SyNF (Synthetic Infrastructure) and make it publicly available. SyNF provides the capacity to accurately estimate key geometries and sizes in the system, providing critical information for assessing perturbations, design changes, and adaptation strategies. Proof of concept studies is performed to examine the applicability of the model at the city level (City of Tempe, AZ) and regional scales (the Phoenix metro region, Arizona). The performance validity of the model is corroborated using data from other WDNs.

2. Model development

The model employs critical water supply and demand as well as transportation and power infrastructure information to create synthetic WDNs. Specifically, capacities are developed to incorporate key factors that represent roadway networks (assuming that pipes largely are deployed under roads), the location and capacity of water sources (e.g., WTP: water treatment plants, wells), where treated water enters the system, hydraulic principles constraining location and size of water mains, and temporally varying municipal water demand estimates. The model outputs include the topology of the WDN, flow direction, flow through pipes, pipe diameter, pump locations, and power requirement for pumps. Figure 1 illustrates the key components of the SyNF model.

The model and associated algorithm can be applied to any city; but is developed and tested with the City of Tempe, Arizona, to demonstrate its capabilities. We then scale to the Phoenix metro region to show how the model can be applied to a region of interconnected WDNs. Tempe is located just east of the City of Phoenix and has a land area of 104 km^2 and a population of 185,000 as of 2019. The city has a water demand of approximately 58 million gallons per day (MGD), and two 50 million gallons per day WTPs (*Water | City of Tempe, AZ*, 2019). Figure 2 shows the topographical map of the Phoenix metro area.

2.1. Network topology

The topology of WDNs encompasses information about the location and layout of pipelines. Water distribution pipes are assumed to be located underneath roads, and road networks are obtained from OpenStreetMap datasets. A Python library named OSMnx (Boeing, 2017) is used to retrieve the road network from OpenStreetMap and to convert that network to a graph, which symbolizes roads as links and intersections as nodes. Each node represents the aggregated water demand from buildings that receive water from the node. The graph is pruned by removing self-loops and parallel links; self-loops are edges that connect a node to itself, and parallel links are multiple links between two of the same nodes. Graph pruning is performed automatically in the model. Freeways and expressways are further removed to prune the graph with the assumption that there are no water

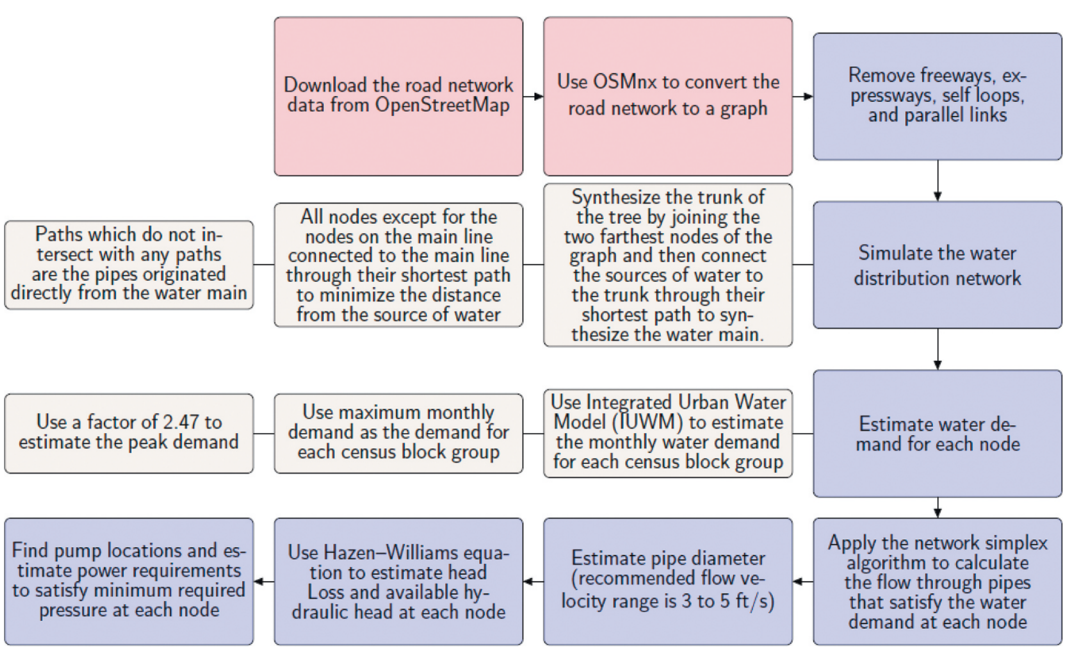

Figure 1. Synthetic water infrastructure generation process.

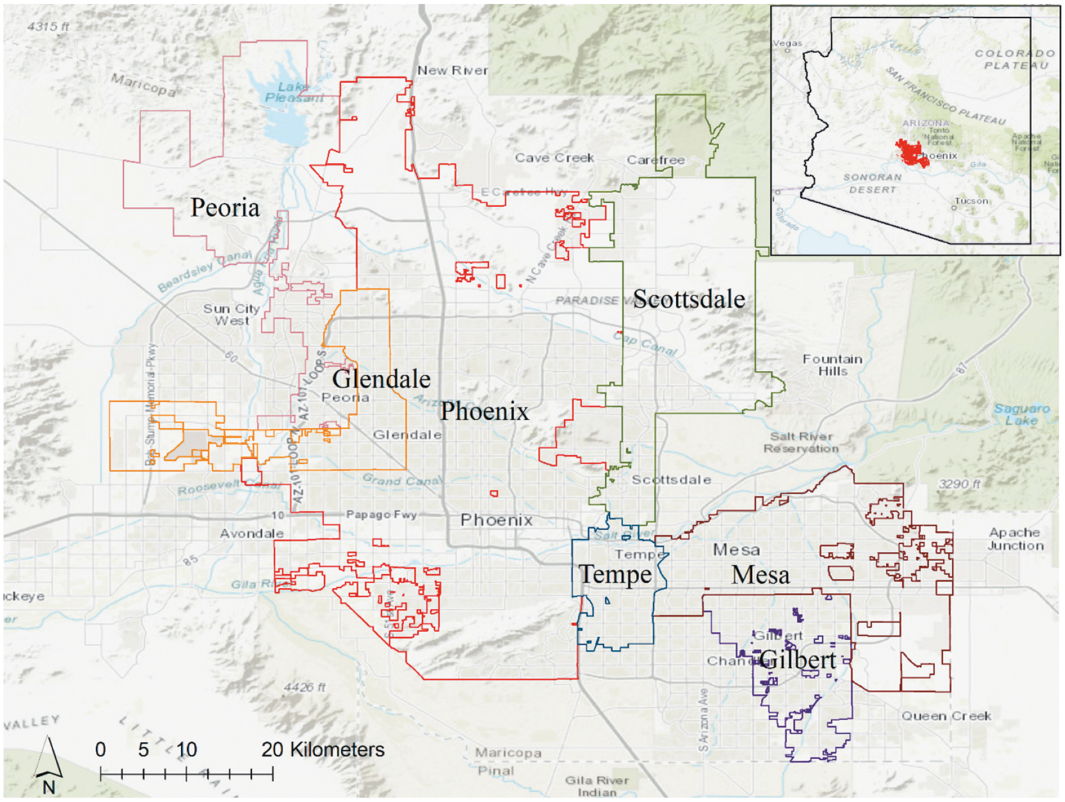

Figure 2. Map of the Phoenix metro area.

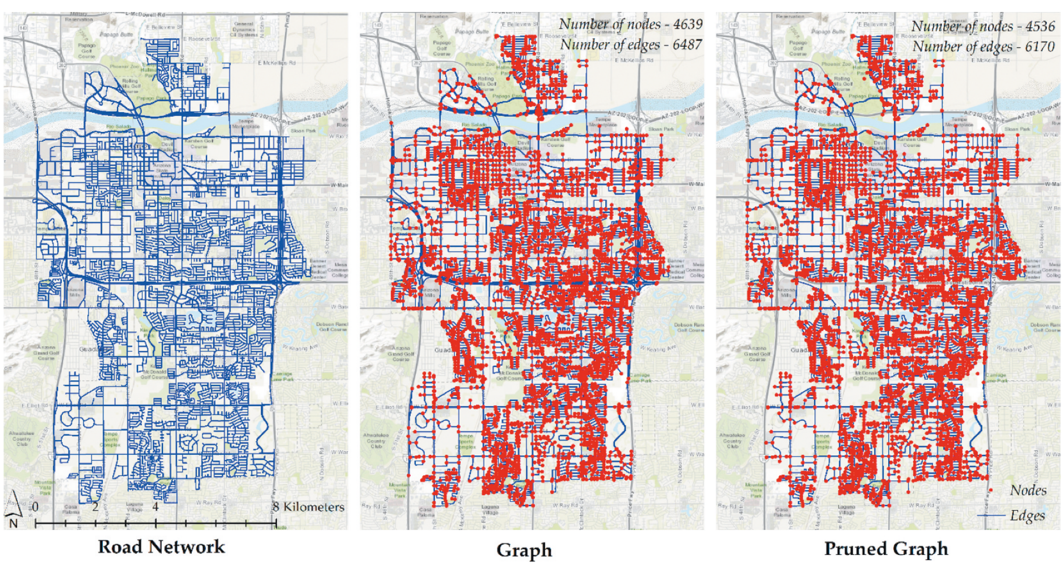

Figure 3. Road network, graph, and pruned graph for Tempe, AZ.

pipes under higher-level function transportation assets. Figure 3 shows the road network, graph, and pruned graph.

2.2. Synthesis of a water main

In a typical WDN, water flows from a source (i.e., WTP) to the water main, and from there to different parts of the service area (i.e., municipality) like a tree network, where the water main is the trunk of the tree (Mays, 2000). The primary objective of this synthesis is to optimize the total length of the WDN so that every node can receive water from its nearest possible source. Therefore, each pipeline, including the water main, needs to cover the maximum possible service area by traversing the shortest distance possible. Since the water main is required to cover the entire service area and originates from main sources (e.g., WTP), the shortest path between the two farthest nodes on the graph is used as the trunk of the tree, and the WTPs are then connected to the trunk through their shortest path to synthesize the water main. Figure 4(a) shows the water main for Tempe, AZ.

2.3. Synthesis of branching pipes

After synthesizing the water main, we simulate the branching pipelines of the WDN. In this simulation, a pipe can originate directly from the water main or other pipelines, like a tree. There are two steps to optimize the total length of the WDN. First, the shortest path is found from each node to the water main, except for the nodes on the water main. Then, the largest shortest paths are obtained that do not intersect with other shortest paths towards the water main. These paths are the pipelines originating directly from the water main. Figure 4(b) shows the pipes stemming from the water main for Tempe. The same procedure is used to simulate the branching pipes where newly synthesized pipes are the source for the next branches of the tree. This iteration continues until all nodes are connected. Figure 4(c) shows the water main and other pipes for Tempe, where red pipes (Subline 1) originated from the water main, blue pipes (Subline 2) originated from red pipes, orange pipes originated from blue pipes (Subline 3), and black pipes (Other Sublines) are the remaining pipes. The iterations are later assumed to correspond to different pipe sizes, as they represent smaller and smaller service areas.

2.4. Water demand

Water demand is used to estimate the infrastructure and operational characteristics of WDNs. We use the Integrated Urban Water Management Model (IUWM) (Sharvelle et al., 2017) to characterize demand across the city. IUWM estimates urban water demand at varying spatial scales, from building to municipal levels at daily time steps. The model includes capacities to assess urban water management strategies (e.g., conservation and reuse) and climate change impacts using census data, climate data, and land-use land-cover information. Although IUWM doesn't identify major water users like manufacturing plants, it provides a realistic urban water demand for synthesizing a WDN. Web services are available to obtain IUWM water demand estimates for any region within the U.S. ('Urban Water Demand Forecasting').

Monthly water demand was estimated using the 2010 U.S. census (US Census, 2010), 2011 USGS national land use land cover (NLCD) (Homer et al., 2004), and the PRISM (Mesinger et al., 2006) climate model. Table 1 shows IUWM input values. IUWM was used to estimate monthly water demand with seasonal changes between 1981 to 2018. Demand nodes are placed at junctions in roadways. The maximum monthly demand is then distributed to each node within the corresponding census block group. For example, for a census block group that has 100 nodes and a maximum demand of 80 million gallons per month (MGM), each node of that census block group would have a demand of 0.80 MGM. Figure 5 shows the census block group and nodal demand for Tempe, in addition to population density. As water demand varies hourly,

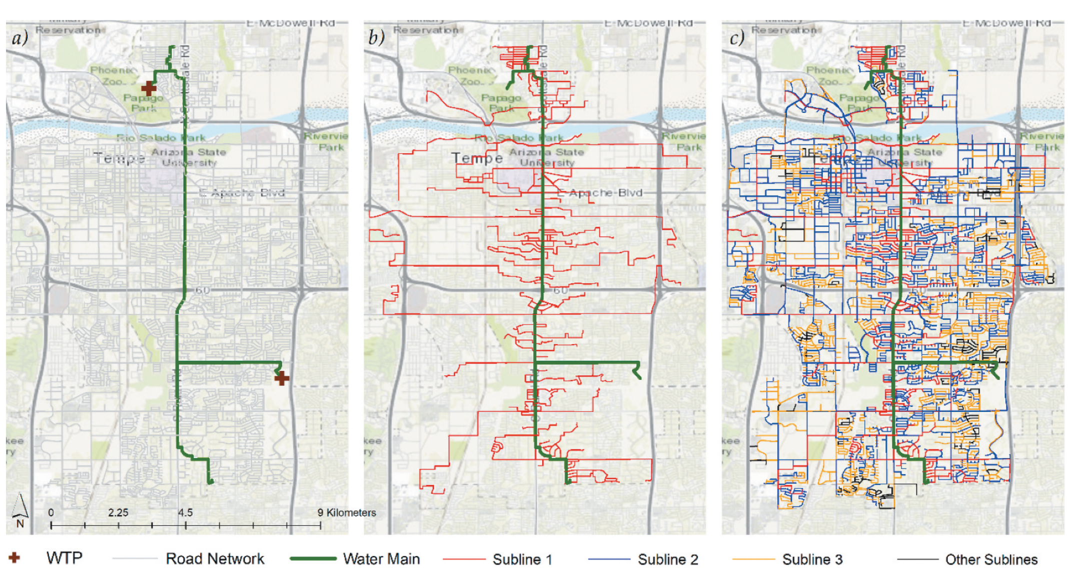

Figure 4. Synthesized water main and other pipes for Tempe, AZ.

Table 1. IUWM parameters and values used in the study.

Parameter	Value
Home Profiles	
Average Homes (before 1999), 177 gphd (%)	0
Average Homes (2016), 138 gphd (%)	100
High Efficiency Homes, 112 gphd (%)	0
User-Defined Home Profile (%)	0
Percent of indoor demand that is consumed (%)	10
Landscape Irrigation Demand and Conservation	
Temperature above which residents irrigate (deg. C)	13.3
Open Space	60
Low Density Development	65
Medium Density Development	35
High Density Development	5
Graywater	
Portion of Faucet Water that is not Kitchen Wastewater (%)	50
Wastewater	
Min. percent blended with treated raw water supply (%)	60
Stormwater	
Percent of precipitation that turns into runoff (%)	90

peak hourly demand as a percentage of average daily demand (P) is estimated using the Goodrich (Brière, 2007) empirical equation,

$$P = 180 * t^{-0.10} \tag{1}$$

where t is the length of period for which peak demand is required (days). From Equation 1 the hourly peak demand was 247% of average daily demand, thus, multiply the average daily demand by 2.47 to obtain the hourly peak demand. As synthesizing storage tanks or storage requirements for fire flow is not a scope of SyNF; thus, we do not explicitly consider fire flow demand for this model. However, the general peak demand factor of 2.47 is well above the factor (i.e., 1.70) recommended by the city of Phoenix (City of Phoenix, 2020b). Finally, the network simplex algorithm (Király & Kovács, 2012; Orlin, 1997) is applied to calculate the flow through pipes that satisfy the requirement at each node. The network simplex algorithm is the graph-theoretic version of the simplex linear optimization algorithm. In many WDN studies, flow in pipes is calculated by modeling a network in EPANET (US EPA, 2014) the state-of-the-art software model developed by the United States environmental protection agency. EPANET requires water pipe diameter as a fundamental input into the model, which can be obtained from the SyNF.

2.5. Pipe diameter

The diameter of each pipe is calculated using the flow through pipes and standards of flow velocity. The continuity equation for a circular pipe is used to calculate the diameter of each pipe:

$$d = \sqrt{\frac{4 * Q}{\pi * V}} \tag{2}$$

where d is the diameter of a circular pipe in meters, Q is the flow in cubic meters per second, and V is the velocity in meters per second. The flow velocity in a typical WDN is typically maintained within the range of 3 ft/s (0.91 m/s) to 5 ft/s (1.52 m/s) (City of Phoenix, 2020b) to avoid accumulation of debris due to low velocity or to prevent water hammering due to high velocity (Viessman & Viessman, 2009). Similarly, A minimum pipe diameter may be specified to accommodate for minimum flow requirements under fire design scenarios (ADEQ Arizona Department of Environmental Quality, 1978). Moreover, for simplicity, we do not set a maximum

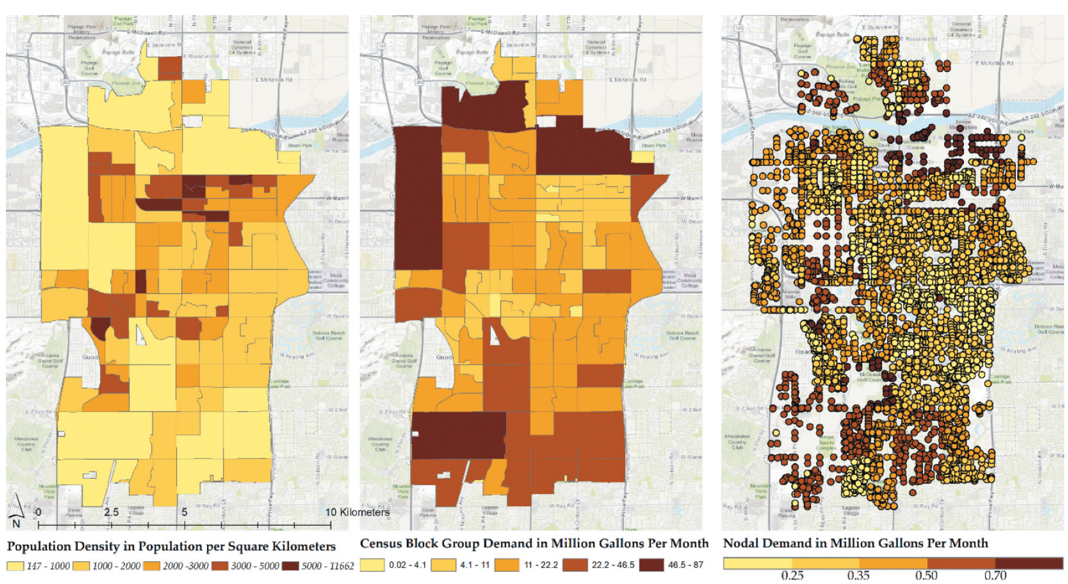

Figure 5. Water demand per census block group (from IUWM) and nodal demand of Tempe, AZ.

pipe size, so the model will use one big pipe instead of multiple smaller pipes.

2.6. Hydraulic head

Every point of a WDN has a hydraulic or piezometric head because of the difference in elevation from the source. Moreover, headloss is also associated with the flow, resulting primarily from friction, and secondarily from bends, fittings, and valves. The Hazen Williams equation (Williams & Hazen, 1908) is used to calculate the head loss in each linear pipe segment,

$$h = \frac{L * 4.52 * Q^{1.852}}{c^{1.852} * d^{4.87}} \tag{3}$$

where h is the head loss in pounds per square inch (psi), L is the length of pipe in feet, Q is the flow in gallons per minute (gpm), c is the pipe roughness coefficient, and d is the diameter in inches. SyNF can calculate L, Q, and d accurately but require input for c to estimate the head loss. With detailed information on pipe materials (i.e., the value of c), a more accurate scenario for head loss can be developed by using SyNF. As the scope of SyNF is not to simulate different scenarios and information on pipe materials are not publicly available, thus, we use 150 as the value of c to estimate the head loss for the Phoenix metro region, which is a relatively new metro region in the USA. Therefore, we assume that majority of the water pipes are new and made of polyvinyl chloride (PVC). Furthermore, a node is assumed to receive water from the WTP (or other water supply systems like wells and storage tanks) that has a minimum head loss. The American Water Works Association (AWWA) recommends a static pressure of 40 to 75 psi for a typical WDN (*American Water Works Association,* 2019). To compute the hydraulic head for each node, it is assumed that water is delivered from the water supply system (e.g. WTP) at a pressure equivalent to a piezometric head of 75 psi. The hydraulic head or available pressure for each node is calculated as,

$$P_n = 75 + 0.433 * (E_{WTP} - E_n) - H \tag{4}$$

where P_n is psi of the hydraulic head of node n, E_{WTP} is the elevation of the WTP in feet, E_n is the elevation of the node in feet, and H is the total head loss in all linear pipe segments from the source to node n. We apply equation 4 for every node of the network.

2.7. Pumps and power requirements

The primary objective of a WDN is to supply potable water with sufficient quantity and pressure. Thus, each node of a WDN must have a specific hydraulic head to ensure adequate water pressure (e.g., 40 psi in this case). Hydraulic pumps are used to elevate the water pressure for a node that lacks sufficient hydraulic head. As water flows from the primary source (e.g., WTP) to the water main, nodes on the water main are secondary sources for the synthesized WDN and must have adequate hydraulic head to ensure required pressure for other nodes of the network. Therefore, nodes with an insufficient hydraulic head are identified and pumps are added to ensure a minimum pressure of 40 psi for the entire network. As the water main is the secondary source of water, to maximize the performance of pumps nodes on the water main which require pressure elevation are identified as the locations for pumps. Energy inputs for the pumps are also computed. To increase the pressure of a node with a demand of 100 gallons by 100 psi, 73 Wh energy is required, which we then use to estimate the power requirement of the pump.

2.8. Service year

In addition to physical and operational characteristics, the year in which pipes are first installed is also estimated. The service year is useful for a number of potential analyses, including assessment of rehabilitation schedules and the likelihood of failure (Bondank et al., 2018). It is assumed that a WDN network starts operating when properties in the surrounding neighborhood (census block group) were first built. Therefore, first the construction year of each parcel within a census block group is obtained from the county's Assessor's database (*GIS – Parcel Database – Maricopa County Assessor's Office,* 2019). Then, one standard deviation below the mean is used as the service year for all pipes in that neighborhood. The result is a pipe service year that corresponds with the construction of the first critical mass of buildings in the neighborhood.

2.9. Model testing

To validate the outcomes we compare the findings with a real WDN. As previously discussed, information about real WDNs remains sparse. As such, we use the North Marin, California WDN, a publicly available model used extensively by the water resources community (*North Marin Water District,* 2019). The WDN of North Marin (NMWD) does not provide complete geospatial information as the locations of nodes are provided only in Cartesian format. While we are unable to compare the networks spatially, we can assess a number of other variables, specifically pipe size and distribution characteristics. To compare the pipe size distribution, first, we apply the SyNF to generate the WDN for the city of

Novato, which is a part of the NMWD. Then, we distribute the pipe diameters for both networks (i.e., EPANET network of NMWD and synthetic WDN) into six bins: 1) less than 8 inches, 2) 8 to 12 inches, 3) 12 to 16 inches, 4) 16 to 20 inches, 5) 20 to 24 inches, and 6) greater than 24 inches, where, the range of pipe diameter is from 6 to 30 inches. Finally, we estimate the length of the network covered by each bin. For example, the overall length of the synthetic WDN is 103,648 ft (31,600 m), where, the total length of pipes with a diameter between 8 to 12 inches is 28,152 ft (8583 m), thus 27% of the synthetic WDN has a diameter of 8 to 12 inches. Validation findings are reported after the results.

3. Results

City of Tempe results are first reported, based on the two 50 MGD WTPs and a flow velocity of 5 ft/s (1.52 m/s). Figure 6 shows the results where the width and color of the pipes are based on their diameter, where population density in population per square kilometers is portrayed as a background.

Next, the hydraulic head was estimated for each node of the network. Maximum pressure of 108 psi and a minimum pressure of −31 psi were computed for the nodes, before installing any pumps or valves for the peak water demand. Figure 7(a) shows the pressure at

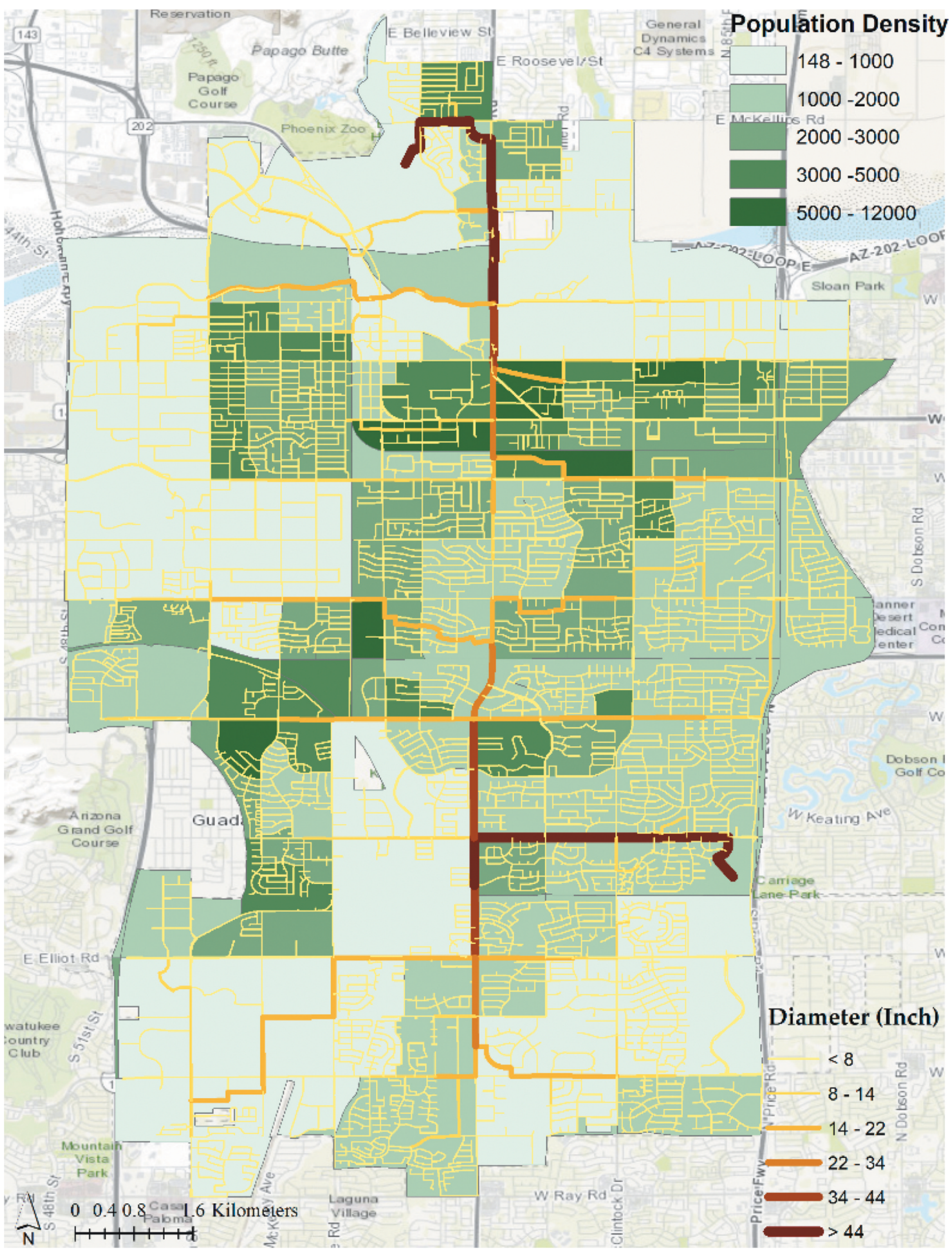

Figure 6. Synthetic water distribution network of Tempe, AZ.

each node. Figure 7(b) shows the pump locations where WTP pumps are marked with ‘X’ and other pumps are marked with ‘+’. The synthetic water distribution network has two WTP pumps and 24 other pumps. Finally, we estimate the power requirement in kW for each pump. Both WTPs have a power requirement of 1,079 kW as they can deliver 50 MGD. Moreover, the power requirements for the other 24 pumps range from 2 to 568 kW, which are shown in Figure 7(b).

The synthetic generation of an entire metropolitan region could provide insights into the composite impact of a widespread threat to all water systems. The applicability of the model at the regional scale was assessed for the Phoenix metro region, which encompasses WDNs for seven major cities in the region. Phoenix, Mesa, Gilbert, Glendale, Peoria, and Scottsdale are modeled in addition to Tempe. The modeling approach in Figure 1 was used to generate the WDNs, which are different in scale from each other. The water networks in the region are owned and managed separately per municipality. Thus the city networks were largely generated independently from each other. The model at regional scale illuminates spatial interdependencies (e.g., probable interconnecting pipes between neighboring cities) among critical natural and built factors that influence WDN reliability planning. For example, there is a direct connection between Mesa and Phoenix as both cities get water from the Val Vista Water Treatment Plant, which is owned jointly (*Water Treatment Process | City of Mesa*, 2019). Furthermore, some utilities have interconnecting pipes between cities to improve reliability (US EPA, 2019). Data on the locations of interconnections in metro regions would allow for an analysis of the water flows between cities in the case of emergencies. A synthetic WDN for a city with a larger road network (e.g., Phoenix) requires more computational time than a city with a smaller road network (e.g., Tempe). Table 2 shows the scale (i.e., number of nodes, overall length), largest diameter, and approximate computational time for seven cities in the Phoenix metro region. Figure 8 shows the results of the multi-city simulation. We use road network to synthesize WDN, and thus, in Figure 8, we can observe some sections of the cities that don’t have any pipes (e.g., western edge of the map) because there are no roads in those sections (e.g., Lake pleasant regional park, which is located northwest of Phoenix, Arizona and within the municipal boundaries of Peoria, Arizona). From Table 2 we can observe

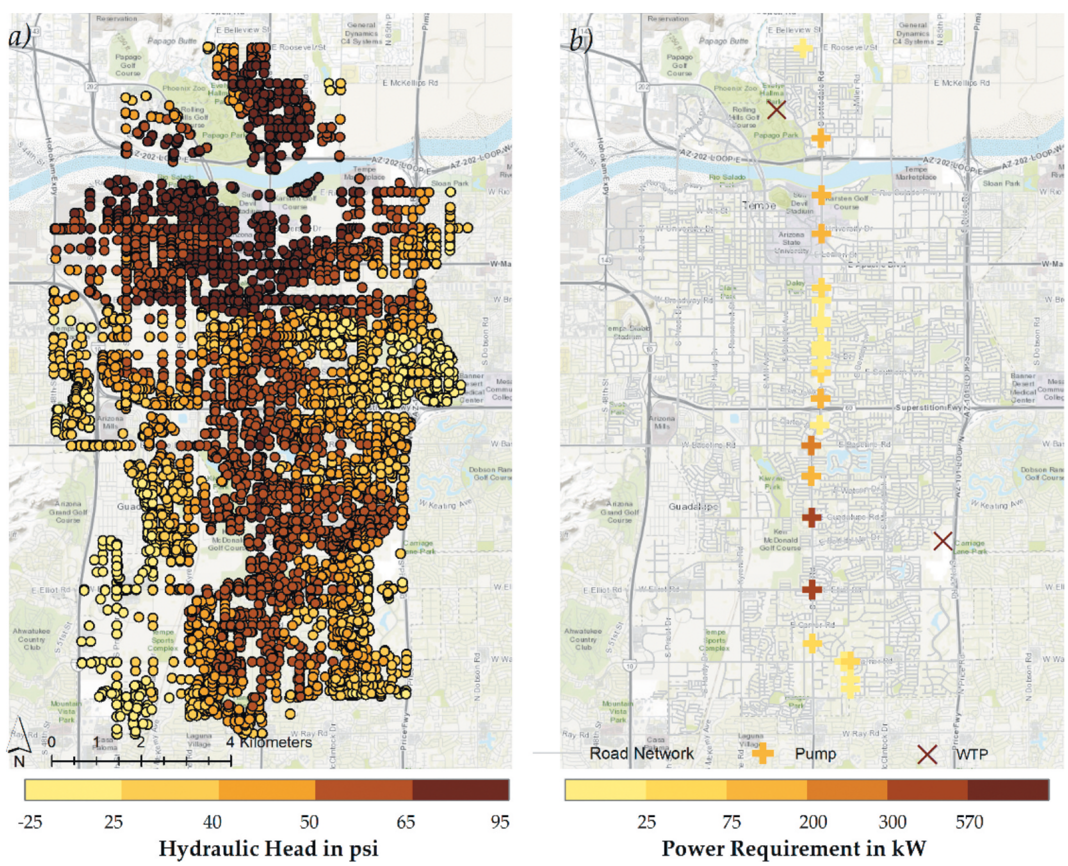

Figure 7. Hydraulic head of each node, pump locations and power requirements of pumps for Tempe, AZ.

Table 2. Scale and computational time for seven cities in the phoenix metro region.

City	Population Served	Water Demand in Million Gallons per Day	Number of Nodes	Overall Length in km	Largest Diameter in inch	Computational Time in hours
Phoenix	1,579,000	466	47,120	8558.16	173	~ 12
Gilbert	247,600	91	9,329	1693.30	67	~2
Glendale	234,766	71	7,168	1318.05	58	~1.5
Mesa	466,000	147	13,655	2440.00	92	~ 4
Tempe	165,000	57	4,536	877.56	46	~1
Peoria	135,975	55	5,633	1111.40	60	~1
Scottsdale	230,000	125	14,078	2487.32	93	~ 4

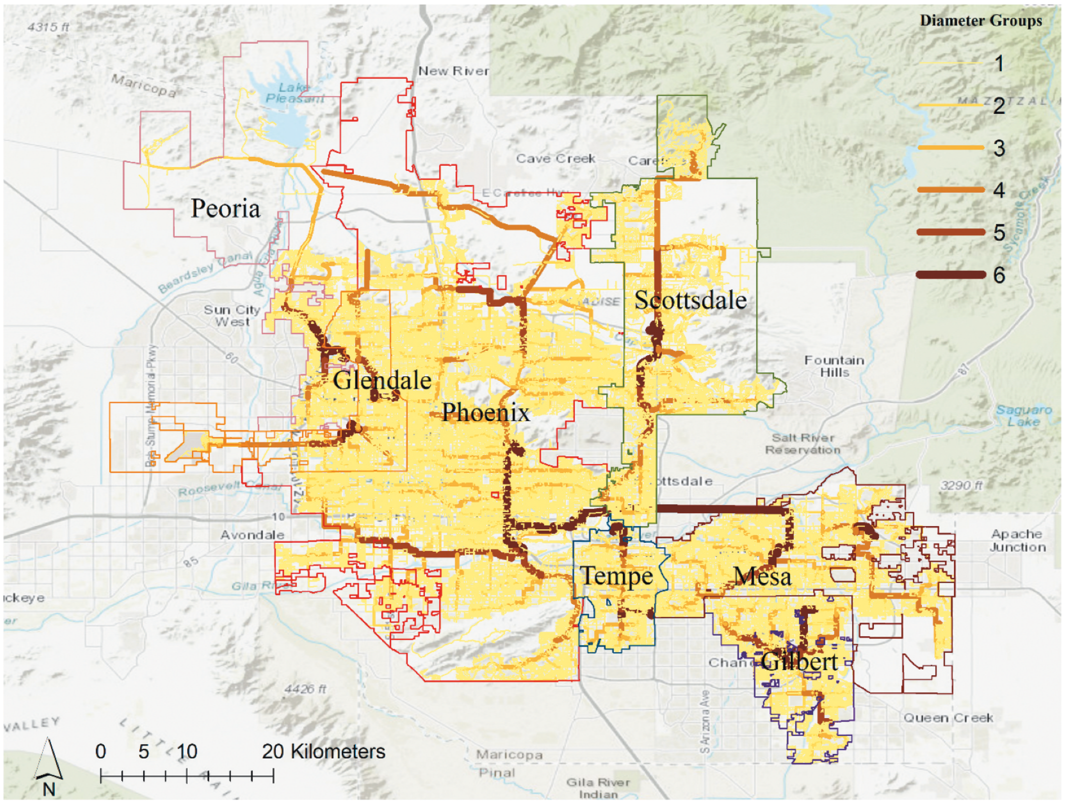

Figure 8. Synthetic WDN of Phoenix, Mesa, Tempe, Gilbert, Glendale, Peoria, and Scottsdale (pipe diameters for each city are divided into six diameter groups for visualization).

that the diameter varies significantly from city to city because of the scale, thus, for visualization we divide pipe diameters for each city into six groups (shown in Figure 8).

Finally, the service year of pipes was estimated for the seven cities. Figure 9 shows the growth of the network since the mid-1900s when metro Phoenix began a significant acceleration of development(City of Phoenix, 2020a). The older neighborhoods in the metro region are found in the center and growth has trended outward from there indicating that older pipes are more likely to be found in the core of the region and newer at the fringe. The growth of the WDN is most heavily represented by the 1980 on time period, when Phoenix experienced major growth. Figure 9 also shows estimates of the initial construction year of pipes but not necessarily their age, as some pipe segments will have been replaced. However, the results suggest that differences in maintenance and rehabilitation practices are needed and are useful for future reliability analyses.

4. Model testing

The NMWD system encompasses 65 kilometers of pipe segments, and the City of Novato has a road network of 298 kilometers. Thus, the WDN contains partial information on major pipes since the length of a WDN is usually greater than that of the corresponding road network (Mair et al., 2017). Figure 10(a–b) show the reported and synthesized WDNs of the Novato. As we

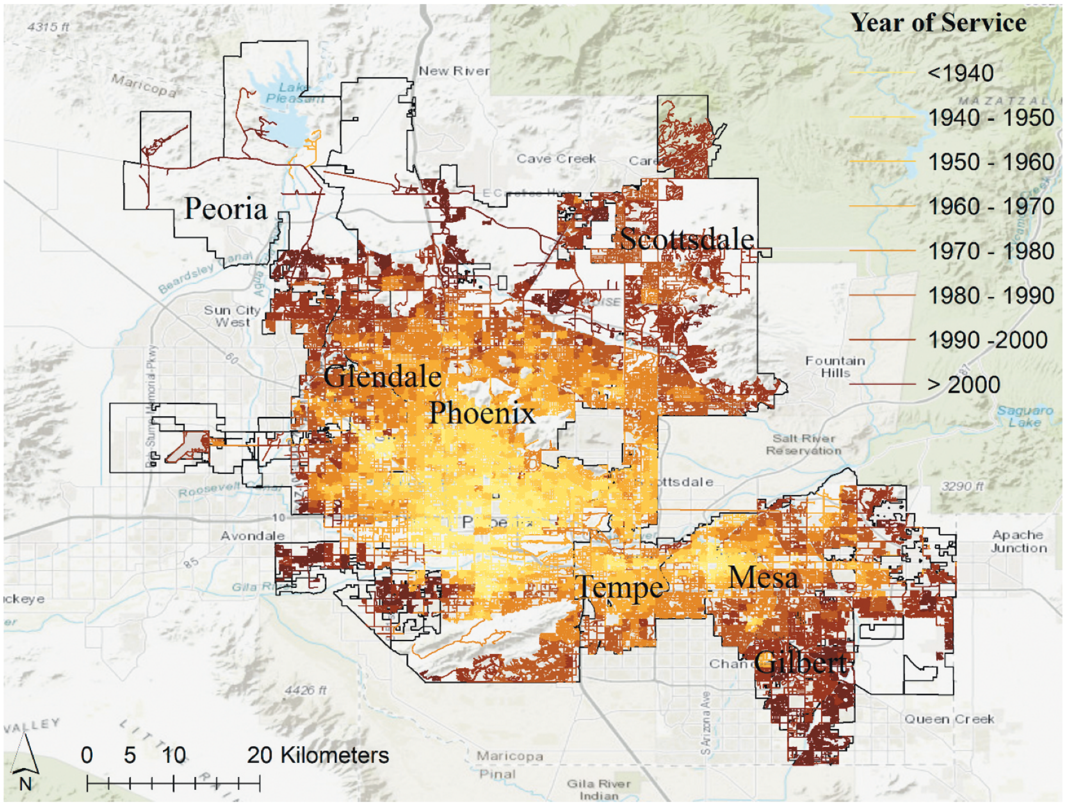

Figure 9. Service year of pipes for Phoenix, Mesa, Tempe, Gilbert, Glendale, Peoria, and Scottsdale.

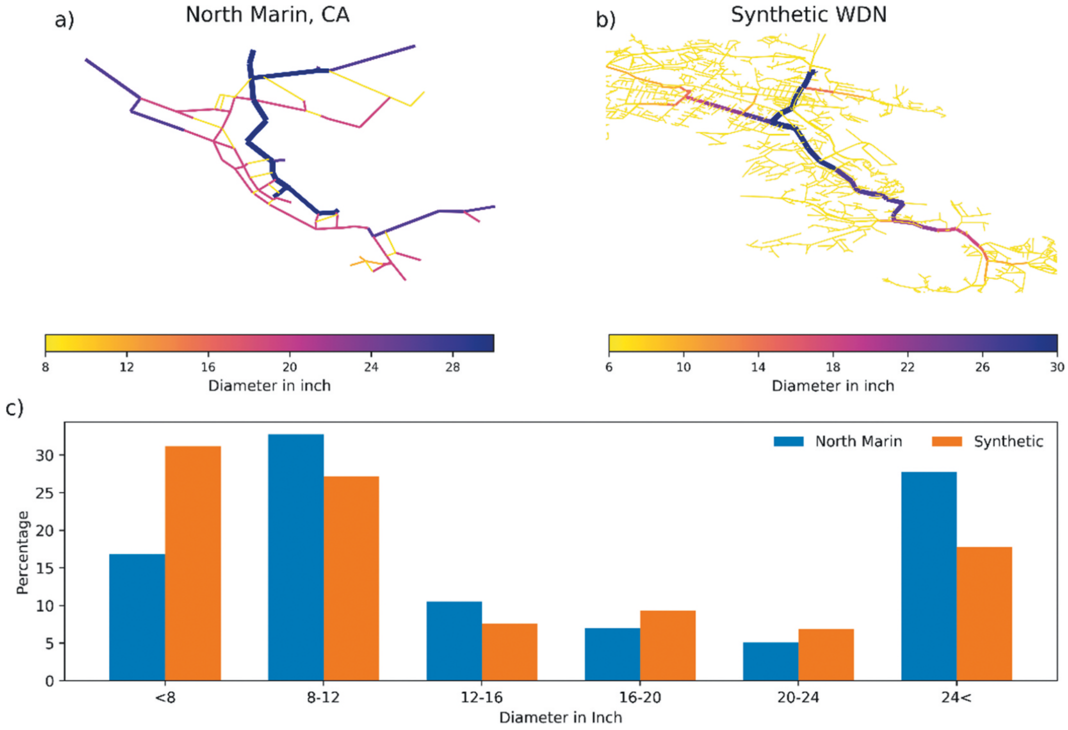

Figure 10. Comparison of the WDN of NMWD and the synthetic WDN.

set the minimum diameter to 6 inches for the minimum requirement, the majority of the pipes in the synthesized WDN are assigned this diameter. Then, the pipe size distributions wasere compared between the two WDNs, where only major pipes or pipes with a minimum diameter of 6 inch were considered for the synthesized network, as the NMWD only contains information on major pipes. Figure 10(c) shows the histogram of pipe size in the actual and synthesized networks. Approximately 32% of the actual NMWD pipes have a diameter of 8 to 12 inches, whereas the synthetic algorithm estimates 27% for its smallest diameter pipes. Similarly, 10%, 7%, and 5% pipes of the NWMD have a diameter of 12 to 16 inch, 16 to 20 inch, and 20 to 24 inch, which are estimated as 8%, 9%, and 6% for the synthesized WDN. Thus, these four bins differ on average by 2.5% between the two WDNs. However, 16% and 30% of the NMWD have a diameter of less than 8 inches and greater than 24 inches, which is estimated at 31% and 19% for the synthesized WDN for equivalent diameters. Therefore, these two outer bins differ on average by 13% between the two WDNs. The mean absolute error (MAE) was used to compute the difference between two continuous variables:

$$MAE = \frac{\sum_{i=1}^{n} \left| y_i - \hat{y}_i \right|}{n} \tag{5}$$

where y and $\hat{y}$ denote the frequency of actual and synthesized pipe diameters in each histogram bin, respectively, and n represents the total number of observations. The MAE was 6.18 between the actual and predicted pipe size distribution. Thus, the pipe size distributions of the NMWD (actual) and the synthetic WDN (predicted) differ on average by 6.18 among the six bins shown in Figure 10(c) and as the pipe size distribution is reported in percentage, therefore the actual and predicted pipe size distribution are on average 6.18% dissimilar. Moreover, as nodes in the synthetic WDN are intersections of road, which are not same for a WDN and the difference in length between the original and the synthetic WDN is significant, thus, we do not compare number of nodes while testing the model.

The comparison indicates that predicted results from the model differ from actual by a reasonably low percentage except for the terminal bins, and as such the model provides reasonable results. It is difficult to gauge broader accuracy estimates given the scarcity of water infrastructure and operational data. While it is reasonable (and even likely) to assume that most water pipes are under roads (outside of large conveyance lines), there is little to no information from actual networks as to their performance and operational characteristics (e.g., pump operation). We anticipate that further validation will be possible as new datasets emerge.

5. Discussion

The resulting synthetic model provides critical new capabilities for understanding the layout, behavior, and vulnerabilities of WDNs. To date, synthetic models have not offered the capability of estimating water network characteristics and flow patterns, especially not at large scales for metropolitan areas. Although the scope of this model is not to simulate different scenarios, the advancements described here provide infrastructure, sustainability, and resilience scientists with insights that create new assessment possibilities including the movement of water, the criticality of pipe segments, the criticality of elements in the broader network, the vulnerability of water components from interconnected infrastructure,and the relationships and dynamics between WDNs in a metro region. The comparison between the NMWD and the synthetic WDN of Novato corroborates that, the synthetic network algorithm can simulate a network similar to the original WDN with an average dissimilarity of 6.18% between the pipe size distributions of the original and the synthetic WDN. However, the NMWD distributes water to other areas in the West Marin apart from Novato, which contributes to this dissimilarity, as the NMWD WDN has a higher percentage of pipes with a diameter of 24 inches or above than the synthetic WDN of Novato, where, these pipes with a larger diameter are usually used to convey water to distant areas and therefore result in a lower percentage of pipes with a diameter of 8 inches or less. Nevertheless, this dissimilarity of 6.18% is low, primarily because of the minimal information used and assumptions made to generate such a synthetic WDN.

The increasing interest in the resilience and vulnerability of critical infrastructure systems creates impetus to fully understand the dynamics and behaviors of WDNs. WDNs have been shown to be susceptible to a number of hazards including accidental damage, decay (AWWA, 2012), climate change (Bondank et al., 2018), and terrorist attacks (Mays, 2004). As resources are invested in improving the resilience of critical infrastructure, an understanding of the underlying characteristics and behaviors will be foundational in developing robust resilience strategies. To this end, the model presented here provides important new capabilities towards ensuring the reliable and equitable distribution of potable water into the long-term. Due to the spatial explicit and hydraulically realistic layout, the model can help identify locations and durations of outages resulting from a failure in the WDN. Due to the robust modeling of demand, the model can help identify the impact these outages will have on the public and economic health of specific neighborhoods. This feature also allows for an exploration of how changes to factors of demand (i.e., new conservation policies or increased extreme events) either exacerbate or mitigate vulnerability in specific neighborhoods, and how this affects the performance of the overall system.

The synthetic approach offers new potentials for estimating the relationships between coupled water and other infrastructure, namely power. There has been much interest recently in understanding the relationships between water and power systems, often termed the water-energy nexus. It is clear that both systems rely heavily on each other, and given each's criticality understanding the dynamics across systems is doubly important. Yet as described gathering actual system data for any one system is challenging enough, let alone two. As there is a much more mature body of literature (Gegner et al., 2016; Schweitzer et al., 2017; Wang et al., 2010) on synthetic power distribution system analysis, the coupling of synthetic water and power systems represents an important new frontier for research and insights. Future work should focus on the integration of two or more synthetic infrastructure systems to begin elucidating the complexities between our ever-changing infrastructure. The algorithm developed in this study to generate synthetic networks provides the spatially-explicit information that would be necessary for modeling the sets of dependencies between the water and power systems including water pumps and power loads (Cascos et al., 2004), water SCADA and power loads (Mays, 2000), water pipes and power transformers (Guinyard et al., 2015; Idaho National Laboratory (INL), 2006), and water demand nodes and power cooling generation(Bartos & Chester, 2014).

6. Conclusion

Little and limited data on water infrastructure is one of the main obstacles to understand its criticality, vulnerability, and resilience. To help address this issue we developed SyNF model to generate a synthetic WDN at a city scale, where, a road network, modeled water demand, and location of water sources are used as input to synthesize the topology of the pipe network, diameter of pipes, location of pumps, and power requirement for pumps, where, minimum hydraulic pressure and minimum pipe size for fire flow are maintained to comply with design standards. We applied the model to generate synthetic WDN for seven major cities of the Phoenix metro region and use the city of Tempe to demonstrate the procedure of the model. Moreover, we use the North Marin, California WDN, which, is the hallmark model for the EPANET community, to validate our model and find that it performs reasonably well, as we find an average dissimilarity of 6.18% in pipe size distribution between the original and modeled network. However, the lack of available data for a real network restricts us from expanding our validation and, thus, calibrating the model further. Although while generating a synthetic WDN, we have to assume some parameters (e.g., flow velocity, pipe roughness coefficient) in the process as well as do not consider the history of remodeling and renovation because of the lack of available information, which could be improved with access to additional data. Moreover, we only validate our model against a small network, whereas a comparison against a real and sizeable WDN can provide more insight to improve the accuracy of the algorithm. Furthermore, although almost all nodes in the synthetic network can get water from at least two different directions (except for maybe cul-de-sacs), the resilience of a synthetic WDN can be enhanced by adding more loops (Mair et al., 2017). Moreover, the long-term goal of SyNF is not only to synthesize a water distribution network (the scope of this article), but also incorporate synthesized power distribution, building, and transportation systems to better understand urban resilience by facilitating interdependency analysis. Details on SyNF and how to obtain the model are available on http://synthetic.resilientinfrastructure.org/. Overall, the model can generate a reasonable synthetic WDN and provides important information on water infrastructure, which can be significantly useful to advance our understanding of the criticality, vulnerability, and resilience of a WDN.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the National Science Foundation [GCR-1934933]; National Science Foundation [SRN-1444755]; National Science Foundation [SCC-1831475]; National Science Foundation [CSSI-1931324]; Office of Naval Research [N00014-18-1-2393].

Notes on contributors

Nasir Ahmad is a post-doctoral scholar in the School of Sustainable Engineering and the Built Environment at Arizona State University.

Mikhail Chester is an associate professor in the School of Sustainable Engineering and the Built Environment at Arizona State University.

Emily Bondank is a post-doctoral researcher in the School of Sustainable Engineering and the Built Environment at Arizona State University.

Mazdak Arabi is a professor of Civil and Environmental Engineering at Colorado State University.

Nathan Johnson is an associate professor in the Polytechnic School at Arizona State University.

Benjamin L. Ruddell is an associate professor in the School of Informatics, Computing, and Cyber Systems at Northern Arizona University.

ORCID

Nasir Ahmad http://orcid.org/0000-0001-5067-7368
Mikhail Chester http://orcid.org/0000-0002-9354-2102
Emily Bondank http://orcid.org/0000-0002-9577-8637
Nathan Johnson http://orcid.org/0000-0002-1456-2384
Benjamin L. Ruddell http://orcid.org/0000-0003-2967-9339

References

ADEQ Arizona Department of Environmental Quality. (1978). *Guidelines for the construction of water systems*. https://azdeq.gov/node/3366

Allenby, B., & Chester, M. Reconceptualizing Infrastructure in the Anthropocene. *Issues in Science and Technology*. 2018, May 25. https://issues.org/reconceptualizing-infrastructure-in-the-anthropocene/

American Water Works Association. Retrieved August 14, 2019, from https://www.awwa.org/

AWWA. (2012). *Buried no longer: Confronting America's Water Infrastructure Challenge*. AWWA Denver.

Bartos, M. D., & Chester, M. V. (2014). The conservation nexus: Valuing interdependent water and energy savings in Arizona. *Environmental Science & Technology*, *48*(4), 2139–2149. doi:10.1021/es4033343

Boeing, G. (2017). OSMnx: New methods for acquiring, constructing, analyzing, and visualizing complex street networks. *Computers, Environment and Urban Systems*, *65*, 126–139. doi:10.1016/j.compenvurbsys.2017.05.004

Bondank, E. N., Chester, M. V., & Ruddell, B. L. (2018). Water distribution system failure risks with increasing temperatures. *Environmental Science & Technology*, *52* (17), 9605–9614. doi:10.1021/acs.est.7b01591

Brière, F. G. (2007). *Drinking-water distribution, sewage, and rainfall collection*. Montréal: Presses Internationales Polytechnique.

Cascos, G., Ciaccia, J., Tripp, J., & Bella, J. (2004). Emergency Backup Power Supply. *American Water Works Association. Journal; Denver*, *96*(1), 41.

Chester, M. V., & Allenby, B. (2018). Toward adaptive infrastructure: Flexibility and agility in a non-stationarity age. *Sustainable and Resilient Infrastructure*, 1–19. doi:10.1080/23789689.2017.1416846

City of Phoenix. (2020a). *City of phoenix history*. Retrieved January 21, 2020a, from https://www.phoenix.gov/pio/city-publications/city-history

City of Phoenix. (2020b). *Water services design standards manual for water and wastewater systems*. Retrieved January 20, 2020b, from https://www.phoenix.gov:443/waterservices/publications/design-manuals/systems

Clark, S. S., Seager, T. P., & Chester, M. V. (2018). A capabilities approach to the prioritization of critical infrastructure. *Environment Systems and Decisions*, *38*(3), 339–352. doi:10.1007/s10669-018-9691-8

Department of Homeland Security. (2014). *Water and wastewater systems sector*. Department of Homeland Security. Retrieved March 26, 2019, from. https://www.dhs.gov/cisa/water-and-wastewater-systems-sector

Francis, R., & Bekera, B. (2014). A metric and frameworks for resilience analysis of engineered and infrastructure systems. *Reliability Engineering & System Safety*, *121*, 90–103. doi:10.1016/j.ress.2013.07.004

Fujiwara, O., & Ganesharajah, T. (1993). Reliability assessment of water supply systems with storage and distribution networks. *Water Resources Research*, *29*(8), 2917–2924. doi:10.1029/93WR00857

Gegner, K. M., Birchfield, A. B., Xu, T., Shetye, K. S., & Overbye, T. J. (2016). A methodology for the creation of geographically realistic synthetic power flow models. *2016 IEEE Power and Energy Conference at Illinois (PECI)*, Urbana, Illinois, USA, 1–6.

GIS - Parcel Database—Maricopa County Assessor's Office. Retrieved October 16, 2019, from https://mcassessor.maricopa.gov/data-sales/gis.php

Guinyard, T., Woo, O., & Arvin, C. (2015). *Water main break floods parking garage*. NBC Southern California. http://www.nbclosangeles.com/news/local/Water-Main-Break-Floods-Street-in-Newport-Beach-322258912.html

Homer, C., Huang, C., Yang, L., Wylie, B., & Coan, M. (2004). Development of a 2001 National Land-Cover Database for the United States. *Photogrammetric Engineering and Remote Sensing*, *70*(7), 829–840. doi:10.14358/PERS.70.7.829

Idaho National Laboratory (INL). (2006). *Critical Infrastructure Interdependency Modeling: A Survey of U.S. and International Research (INL/EXT-06-11464, 911792; p. INL/EXT-06-11464, 911792)*. doi:10.2172/911792

Király, Z., & Kovács, P. (2012). Efficient implementations of minimum-cost flow algorithms. ArXiv:1207.6381 [Cs]. http://arxiv.org/abs/1207.6381

Loganathan, G., Greene, J., & Ahn, T. (1995). Design heuristic for globally minimum cost water-distribution systems. *Journal of Water Resources Planning and Management*, *121*(2), 182–192. doi:10.1061/(ASCE)0733-9496(1995)121:2(182)

Mair, M., Rauch, W., & Sitzenfrei, R. (2014). Spanning tree-based algorithm for generating water distribution network sets by using street network data sets. *World Environmental and Water Resources Congress, 2014*, 465–474. doi:10.1061/9780784413548.050

Mair, M., Zischg, J., Rauch, W., & Sitzenfrei, R. (2017). Where to find water pipes and sewers?—On the correlation of infrastructure networks in the urban environment. *Water*, *9*(2), 146. doi:10.3390/w9020146

Mays, L. W. (Ed.). (2000). *Water distribution systems handbook*. New York, NY: McGraw-Hill.

Mays, L. W. (Ed.). (2004). *Water supply systems security*. New York, NY: McGraw-Hill.

Mesinger, F., DiMego, G., Kalnay, E., Mitchell, K., Shafran, P. C., Ebisuzaki, W., Jović, D., Woollen, J., Rogers, E., Berbery, E. H., Ek, M. B., Fan, Y., Grumbine, R., Higgins, W., Li, H., Lin, Y., Manikin, G., Parrish, D., & Shi, W. (2006). North American regional reanalysis. *Bulletin of the American Meteorological Society*, *87*(3), 343–360. doi:10.1175/BAMS-87-3-343

Möderl, M., Sitzenfrei, R., Fetz, T., Fleischhacker, E., & Rauch, W. (2011). Systematic generation of virtual networks for water supply: SYSTEMATIC GENERATION. *Water Resources Research*, *47*(2), 2. doi:10.1029/2009WR008951

NIPP. 2013: Partnering for critical infrastructure security and resilience (2014). Washington DC. Retrieved from https://www.cisa.gov/publication/nipp-2013-partnering-critical-infrastructure-security-and-resilience

North Marin Water District. (2019). Retrieved June 24, 2019, from https://www.nmwd.com/about.php

Orlin, J. B. (1997). A polynomial time primal network simplex algorithm for minimum cost flows. *Mathematical Programming*, *78*(2), 109–129. doi:10.1007/BF02614365

Park, H., & Liebman, J. C. (1993). Redundancy-constrained minimum-cost design of water-distribution nets. *Journal of Water Resources Planning and Management*, *119*(1), 83–98. doi:10.1061/(ASCE)0733-9496(1993)119:1(83)

Park, J., Seager, T. P., Rao, P. S. C., Convertino, M., & Linkov, I. (2013). Integrating risk and resilience approaches to catastrophe management in engineering systems: Perspective. *Risk Analysis*, *33*(3), 356–367. doi:10.1111/j.1539-6924.2012.01885.x

Pederson, P., Dudenhoeffer, D., Hartley, S., & Permann, M. (2006). Critical infrastructure interdependency modeling: A survey of US and international research. *Idaho National Laboratory*, *25*, 27.

Petit, F., Verner, D., Brannegan, D., Buehring, W., Dickinson, D., Guziel, K., Haffenden, R., Phillips, J., & Peerenboom, J. (2015). *Analysis of critical infrastructure dependencies and interdependencies (ANL/GSS–15/4, 1184636)*. Henry Stewart Publications. doi:10.2172/1184636

Petit, F., Wallace, K., & Phillip, J. (2014). Interactive dependency curves for resilience management. *Journal of Business Continuity & Emergency Planning*, *8*(2), 141–155.

Rinaldi, S. M., Peerenboom, J. P., & Kelly, T. K. (2001). Identifying, understanding, and analyzing critical infrastructure interdependencies. *IEEE Control Systems Magazine*, *21*(6), 11–25. doi:10.1109/37.969131

Savic, D. A., & Walters, G. A. (1997). Genetic algorithms for least-cost design of water distribution networks. *Journal of Water Resources Planning and Management*, *123*(2), 67–77. doi:10.1061/(ASCE)0733-9496(1997)123:2(67)

Schweitzer, E., Scaglione, A., Monti, A., & Pagani, G. A. (2017). Automated generation algorithm for synthetic medium voltage radial distribution systems. *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, *7*(2), 271–284. doi:10.1109/JETCAS.2017.2682934

Sharvelle, S., Dozier, A., Arabi, M., & Reichel, B. (2017). A geospatially-enabled web tool for urban water demand forecasting and assessment of alternative urban water management strategies. *Environmental Modelling and Software*, *97*, 213–228. doi:10.1016/j.envsoft.2017.08.009

Sitzenfrei, R., Fach, S., Kinzel, H., & Rauch, W. (2010a). A multi-layer cellular automata approach for algorithmic generation of virtual case studies: VIBe. *Water Science and Technology*, *61*(1), 37–45. doi:10.2166/wst.2010.782

Sitzenfrei, R., Möderl, M., Hellbach, C., & Rauch, W. (2011). Application of a stochastic test case generation for water distribution systems. *World Environmental and Water Resources Congress*, 113–120. doi:10.1061/41173(414)12

Sitzenfrei, R., Möderl, M., & Rauch, W. (2010b). WDS Designer—A tool algorithmic generation of water distribution systems based on GIS Data. In *World Environmental and water resources congress 2010: Challenges of Change* (pp. 4193–4204).

Urban Water Demand Forecasting. (2019). *One water solutions institute*. Retrieved November 19, 2019, from http://onewatersolutions.com/our-software/tools/urban-planning/urban-water-demand-forecasting/

US Census 2010. *Decennial Census by Decades*. Retrieved April 9, 2019, from https://www.census.gov/programs-surveys/decennial-census/decade.html

US EPA. *Resilient strategies guide for water utilities [Data and Tools]*. US EPA. 2019, February 25. https://www.epa.gov/crwu/resilient-strategies-guide-water-utilities

US EPA. (2014, June 24). *EPANET* [Data and Tools]. US EPA. https://www.epa.gov/water-research/epanet

Viessman, W., & Viessman, W. (Eds.). (2009). *Water supply and pollution control* (8th ed.). Upper Saddle River, NJ: Pearson Prentice Hall.

Wang, Z., Scaglione, A., & Thomas, R. J. (2010). Generating statistically correct random topologies for testing smart grid communication and control networks. *IEEE Transactions on Smart Grid*, *1*(1), 28–39. doi:10.1109/TSG.2010.2044814

Water treatment process. (2019). City of Mesa. Retrieved August 30, 2019, from. https://www.mesaaz.gov/residents/water-resources/services/water-treatment-process

Water. City of Tempe, AZ. (2019). Retrieved August 29, 2019, from. https://www.tempe.gov/government/municipal-utilities/water

Williams, G. S., & Hazen, A. (1908). *Hydraulic Tables: The elements of gagings and the friction of water flowing in pipes, aqueducts, sewers, Etc. As determined by the Hazen and Williams formula and the flow of water over sharp-edged and irregular weirs, and the quantity discharged, as determined by bazin's formula and experimental investigations upon large models*. New York: J. Wiley & sons.