Evaluating Segment and Valve Importance and Vulnerability

Noha Abdel-Mottaleb[1] and Tom Walski F.ASCE*[2]

ABSTRACT

Because consideration of segments and valves is essential for evaluating the reliability and resilience of water distribution networks (WDNs) when shutdowns are required, a quick method of identifying critical and vulnerable segments and valves would benefit utilities. While the importance and vulnerability of segments can best be evaluated by extensive hydraulic analysis, hydraulic analyses can be time consuming. It can also be challenging to visualize the segments of a water distribution network and their associated valves. To address these limitations, this study develops a method based on graph theory to identify important and vulnerable segments without hydraulic calculations. The method generates a matrix that represents how reachable water sources are from segments when a given segment must be isolated, while distinguishing between continuous water sources and ephemeral storage. This study also applies measures from graph theory to determine the number of valves to operate to isolate a segment and provides a rigorous proof to support the intuitive equation. A method to visualize the connectivity of segments with the graph theory measures is demonstrated. The developed methods are applied to multiple valving scenarios of a case study and two real water distribution networks. Correlations between graph-theory based measures derived from the segment-valve topology and hydraulic simulation-based criticality are higher than in previous studies that apply graph theory to the pipe-junction topology of WDNs ($r \geq 0.6$). Results indicate that the developed methods can be used by utilities as a preliminary screening to eliminate the need for some hydraulic simulations. These findings are expected to provide decision-support for utilities.

[1]Ph.D. Candidate, Department of Civil and Environmental Engineering, University of South Florida 4202 E. Fowler Ave., ENB 118 Tampa, Florida, 33620. Email: nohaa@usf.edu

[2]Bentley Fellow, Bentley Systems, Nanticoke, PA. E-mail: tom.walski@bentley.com

Keywords: water distribution networks, complex network analysis, graph theory, segments, isolation valves, vulnerability, criticality

INTRODUCTION

Resilience of Water Distribution Networks (WDN) is the ability of the WDN to maintain adequate function in face of diverse and unanticipated failures and to bounce back (Vugrin et al., 2010; Diao et al., 2016; Klise et al., 2017; Meng et al., 2018; He and Yuan, 2019). As such, enhancing the resilience of WDNs continues to be an important, open research topic (Diao et al., 2016; Butler et al., 2017; Abdel-Mottaleb et al., 2019; Pagani et al., 2020). Many municipalities struggle with pipe breaks as systems age or morph in unexpected ways. Because there may be insufficient isolation valves in place to isolate each pipe in the system, when a pipe breaks, needing repair or replacement, it is often more than one pipe that must be isolated for repairs. This exacerbates the effects failures have on customers and is a situation that utilities seek to avoid (Walski, 1993; Liu et al., 2017; Zischg et al., 2019; Giustolisi, 2020).

Much of the literature focuses on identifying critical pipes in WDN or evaluating resilience by assuming individual pipes can always be isolated in the system (Shuang et al., 2014; Abdel-Mottaleb et al., 2019; He and Yuan, 2019; Balekelayi and Tesfamariam, 2019). However, it is the valves located along pipes that isolate failures, limiting widespread impacts due to the failure (Giustolisi, 2020). In evaluating WDN reliability, a common but faulty assumption has been that a single pipe can be isolated in the network. More often than not, there are not valves at each end of every pipe in the network. This means that a shutdown cannot always be isolated to a single pipe. A WDN segment is a portion of the system that can be isolated by valves; segments often contain more than one pipe (Walski, 1993; Giustolisi and Savic, 2010). Thus, a crucial step in enhancing reliability and resilience of WDNs is to evaluate the vulnerabilities and criticality of isolation valve systems in place. Isolation valves are critical when their operability is important for maintaining the performance of a given WDN.

Figure 1 shows a typical pipe, P-1, in a hydraulic model connecting nodes J-1 and J-2. Accounting for isolation valves, the pipe is part of three distribution segments S-101, S-102 and S-103.

Segments are bounded by isolation valves or other control devices. Segments may be made up of node elements and parts of pipe elements from a hydraulic model. Since there is no way to close a node, there is not a one-to-one relationship between pipe elements and segments.

Both graph theory and the emerging field of complex network analysis (CNA) are powerful tools to analyze networks and have been applied to WDN over the years for varying purposes, including evaluation of various aspects of reliability and resilience (Agathokleous et al., 2017; Hwang and Lansey, 2017; Ulusoy et al., 2018; Abdel-Mottaleb and Zhang, 2019; Yazdani and Jeffrey, 2012a,b; Torres et al., 2017; Giustolisi et al., 2017, 2019; Zeng et al., 2017; Balekelayi and Tesfamariam, 2019; Pagano et al., 2019). CNA builds on graph theory to focus more on the relationship between structure and function of a graph and the complex system (e.g., a WDN) it represents (Zweig et al., 2016). CNA provides tools to analyze more complex networks such as networks with multiple edges connecting nodes, weighted edges and nodes, among other things. However, both graph theory and CNA have often been applied to what is called the pipe-junction topology or the primal graph, where nodes represent demand junctions or other point features and edges represent pipes (Zischg et al., 2019). This representation of WDN is the most common, being necessary for hydraulic analyses and an accurate spatial representation of the system in maps (Walski, 1993). Further, this is most commonly visualized topology of WDNs, as there have been computer programs automating it and most discussions have been pipe-centered. Thus, some methods for identifying critical WDN components have been focused on this representation.

The number and location of valves has generally received very little attention in the research literature compared with other aspects of distribution design and analysis with some exceptions (Walski, 1994; Deb et al., 2006; Walski, 2011). This includes the application of CNA. While CNA has increasingly been applied to the pipe-junction representation of WDNs, limitations include that the pipe-junction topology is not representing the operational reality in event of pipe failures and their associated shutdowns because valves are not considered. For example, a cut set, which by definition has to do with isolation of portions of a network, is not accurate if applied to the pipe-junction topology because an edge that represents a pipe may not actually have valves on either end. However, if applied to the segment-valve representation could in fact represent the collection of valves that can result in excessive unintended isolation. Graph theory and CNA measures have hardly been applied to the dual, segment-valve representation where nodes represent segments and edges represent valves connecting them (Walski, 1993; Zischg et al., 2019; Giustolisi, 2020).

Using a segment-valve representation facilitates inspection and evaluation of potential isolation scenarios that can face a system to allow repair when system elements fail (Jun and Loganathan, 2007). This representation was first explained and visualized in (Walski, 1993), where the author differentiated between the Bouchart and Goulter (1991) definition of segments. The first commercial software to utilize segment-valve topology was WaterGEMS (Bentley Systems, 2020) which introduced its segmentation and criticality tool in the early 2000s. The first application to a full-scale system was Walski et al. (2008). Since then, there has been considerable research on identifying WDN segments, analyzing shutdowns with different valving scenarios and more recently, isolation valve system design.

There are two major research areas with respect to WDN segments and valving: analysis and planning (Jun and Loganathan, 2007). Planning refers to identifying optimal isolation valve systems (IVS) (e.g., Jun (2005); Alvisi et al. (2011); Choi et al. (2018); Giustolisi (2020)), DMA configuration (e.g., Santonastaso et al. (2019); Creaco and Haidar (2019)). Analysis involves identifying and evaluating impact of unintended isolations (e.g., Deb et al. (2006); Jun and Loganathan (2007); Creaco et al. (2012)), comparing different valving scenarios (e.g., Liu et al. (2017); Zischg et al. (2019); Atashi et al. (2020)), identifying optimal near-real time response strategies (e.g., Mahmoud et al. (2018)), and improving hydraulic modeling (e.g., Vasilic et al. (2018)). Zischg et al. (2019) and Jun and Loganathan (2007) focused on obtaining insight from the structure, or physical configuration of segments and valves, in lieu of hydraulic simulations.

However, there remain gaps to be addressed regarding connectivity analyses of WDNs using graph theory. As Meng et al. (2018) and Giustolisi et al. (2019) note, water sources are often not accounted for in studies that use connectivity analysis to evaluate reliability and resilience of WDNs (e.g., Yazdani and Jeffrey (2012a)). The concept of reachability (emphasizing the connectivity to water sources) has been introduced in previous research, but previously has referred to the pipe-junction representation. While Dziedzic and Karney (2014) used a connectivity matrix to evaluate WDN component connectivity to water sources, the matrix did not account for segments or different types of water sources, such as distribution tanks and reservoirs. Similarly, to evaluate WDN reliability, Wagner et al. (1988) quantified the reachability of water sources to demand junctions rather than to segments and did not distinguish between different water sources. Though the matrices developed by Jun and Loganathan (2007) are based on the segment-valve representation, they also do not account for the differences between remaining connected to water sources, such as reservoirs and more ephemeral storage, such as tanks. While the matrices allow tracing unintended isolations, they do not provide a prioritization scheme, or ranking, of segments and they do not measure the aggregate vulnerability of segments due to potential unintended isolations.

Additionally, the developed segment-valve diagrams would be useful if automated such that it is feasible for utilities to readily use them for mapping large system valving. While Zischg et al. (2019) apply CNA to the segment-valve topology, their methodology provides more global network-wide insight than it does local, component-level (see Pagano et al. (2019) for distinction between local and global graph theory or CNA-based measures). Similarly, Atashi et al. (2020) apply global resilience analysis to the segment-valve topology that provides global network-wide insight. Such a methodology, while practical to compare among isolation valves in different systems, does not provide details on which segments are problematic.

Further, traditional techniques of evaluating topological reliability and quantifying connectivity that have been applied to the pipe-junction representation have not been applied to the segment-valve representation (e.g., minimum edge cut sets by Su et al. (1987) and Yang et al. (1996), articulation points by Jacobs and Goulter (1989); Jacobs and Coulter (1991)). However, there are direct operational insights that might be gained by using these metrics from graph theory on the segment-valve representation because of how the graph theory measures physically correspond to this representation of WDNs.

The most accurate way to determine the criticality of each segment in a WDN is to remove each segment from a water distribution system model and conduct a pressure dependent demand, extended period simulation to identify impacted customers and calculate the demand shortfall. This type of calculation is performed in software such as WaterGEMS (Bentley Systems, 2020). Utilities attempting to analyze this problem manually would have to run many simulations because there no quick preliminary screening tool to hone in on important segments. Moreover, there is no tool to quickly identify vulnerable segments to help utilities identify problematic locations of a given WDN.

To address these gaps, the authors expand on previous research (specifically, connectivity analyses) that applies graph theory to WDNs. The major contribution of this study is in applying graph theory to the segment-valve topology for preliminary screening to identify problem areas. The study provides methods for displaying results in a useful way to utilities by visualizing various measures to help quickly identify important (i.e., critical) and vulnerable segments and also to help map segments that do not necessitate closing all downstream valves. Additionally, the paper presents an importance and vulnerability index that account for the difference between continuous water sources and ephemeral water storage when evaluating the impact of segment isolation. The paper first presents an improved approach to visualizing the segment-valve topology, followed by the use of a reachability matrix to determine important and vulnerable segments. Then, a method to determine the number of valves that must be operated to isolate segments in different cases is presented. Several case study scenarios are studied to illustrate the use of the topology display and reachability-based indices for a variety of systems, including some real systems.

METHODOLOGY

Development of Segment-Valve Topology Display

Using the graph representation of segment connectivity to visualize the interconnections between segments is valuable and advantageous to communicate the relationship of valves and segments (Jun and Loganathan, 2007). Such a segment-valve graph can be constructed by treating each segment as a node and each valve as an edge connecting the segments. A segment-valve graph can be considered as a dual topology compared with the usual pipe-junction topology.

The information necessary to construct a segment-valve display includes: the list of segments and their associated pipes, locations of pipes (x,y) coordinates used to place the segment spatially, and valves connecting the segments. Segments containing water sources are also identified, and their IDs stored. For this study, WaterGEMS is used to identify segments and their elements. The segment-valve graph can be constructed using only open source software packages within python. An empty *networkx* (Hagberg et al., 2008) graph is initialized, and all segments are stored as nodes.

In WDNs where there are segments connected to each other by more than one valve, a *multigraph* object is initialized instead of a graph object to allow for multiple edges between nodes. Similarly, if a WDN has directional elements (e.g., pumps, PRVs), a *digraph* object (i.e., directed graph) is initialized to account for direction. When studying a given pressure zone with directional elements, segments flowing into the zone should be treated as open links while those flowing out of the zone should be treated as closed to prevent water from moving backward through a directional element. For each node, (x,y) coordinates are assigned as the centroid of the pipes belonging to the given segment. The centroid is calculated using the *shapely* package with the *multipoint* and *centroid* functions (Gillies, 2013). If two segments are connected by a valve, then an edge or link is assigned between the two nodes in the graph. The exact location of the valve is not needed, and it is assumed that all valves are operable although the graph can be reconstructed to eliminate inoperable valves. A schematic of a segment-valve diagram is shown in Figure 2a.

To display the topology, the *matplotlib* and *networkx* packages are used. The graph object is used as input into the *networkx draw* function along with settings for node and edge (i.e. segment and valve, respectively) sizes and displaying labels and coordinates. The size and color of nodes and edges can be used to represent different attributes by passing the dictionary object of the attribute values as input to the *networkx* draw function. For the case study presented in this paper, sample code is provided in a repository (see Abdel-Mottaleb and Walski (2020b)).

Reachability Matrix

Determination of Reachability Matrix

Constructing the reachability matrix, $\boldsymbol{R}$, is an integral step of assessing importance and vulnerability of segments and was introduced in Abdel-Mottaleb and Walski (2020a). Rows, m, of the matrix represent segment closure and columns, n, represent segment impact. This section illustrates matrix construction with a small example. First, every segment closure is considered. For each segment closure, the impact on other segments connectivity to water sources is evaluated. Water sources are distinguished as continuous or ephemeral. Continuous sources such as large reservoirs, wells and treatment plants are long lasting and ephemeral sources, such as distribution tanks, are temporary because they can drain in a few hours if they are not connected to a continuous source.

If the closure of a given segment (m) causes another segment (n) to lose connection with all water sources, a value of 2 is assigned to R_{mn} for all isolated segments. If the closure of a given segment (m) causes another segment (n) to lose connection with all continuous water sources yet maintain connection to an ephemeral water source, a value of 1 is assigned to R_{mn} the closure of a given segment (m) does not cause another segment (n) to lose connection with continuous water sources, a value of 0 is assigned to R_{mn} is repeated for each segment in the network until the cells of matrix $\boldsymbol{R}$ are populated, as shown in Figure 2 for a small example. In Figure 2, the reservoir is the continuous source while the Tank is an ephemeral source.

Summing the values of row m, corresponding to a given segment, gives an indication of how important a segment is because its closure results in many unintended isolations, and this sum is called the importance index. Summing the values of column n, corresponding to a given segment, gives an indication of how vulnerable that segment is to unintended isolation. A high value of the importance index indicates shutting down this segment will affect many other segments, while a value of 2 indicates that segment will only impact itself. A high value for the vulnerability index indicates that a shutdown of several other segments can isolate this segment, while a value of 2 indicates that failures in other segments will not affect it.

Calculations based on the reachability matrix can be represented on the segment-valve topology.

To compare the indexes of different valving scenarios, the importance and vulnerability indexes are min-max normalized between 0 and 1. Meaning, for a given network scenario, the minimum index is subtracted from all the values and then divided by the difference between the maximum and minimum values of the segments in the network scenario.

Different valving scenarios of a given network will also differ in their segments. Because each scenario would also have a different number of segments, the data (i.e., importance and vulnerability indexes) measured for each scenario is on a different scale. When the data for each scenario is re-scaled or min-max normalized between 0 and 1, the different valving scenarios can be compared with each other. For example, if the spread of the vulnerability index of a scenario is mostly above 0.5 and another scenarios vulnerability index is mostly below 0.5, this provides information relating how vulnerable a valving scenario is compared with another. The normalization procedure also allows location-specific comparisons between valving scenarios. Because the scale is normalized, comparing segments within a given system is also easier as the maximum and minimum values are known to be 1 and 0, respectively.

Validation by Comparison with WaterGEMS

The reachability matrix importance index is validated with WaterGEMS criticality as follows. First the reachability matrix is constructed, and the importance index is calculated for each given segment isolation. The WaterGEMS (Bentley Systems, 2020) criticality tool is also run for the three valving scenarios of the case study using a 24-hour extended period simulation (pressure driven) to quantify the system demand shortfall for each segment isolation starting at $t = 0$ and lasting through the 24-hour duration.

For the case study presented in the following section, the average supplied flow in the network is $44 L/s$. The tank has a base elevation of $49m$, initial elevation of $55m$ and maximum elevation of $58m$. The reservoir and pumps are designed to fill the tank during the simulations. The analysis mode was set to hydraulics only, the reference pressure is $50m$, and the threshold pressure is set to the reference pressure.

The Pearson correlation coefficient and linear regression statistics are calculated in python using scipy (Jones et al., 2016), and are used to evaluate the correlations between the reachability matrix value and the hydraulic simulation-based result.

Identifying Valves that must be Operated for Segment Shutdowns

The location of a segment in the network has operational implications for valve closure when segments must be isolated. Depending on the structure that a given segment is part of (e.g. loop, tree, or both) the number of valves that must be closed to isolate the segment differ. Trees refer to the portions of the network where two or more segments are connected by only one path (i.e., free of loops). Additionally, a failed valve may mean that more valves must be closed to isolate a segment, or it might not have an impact– depending on its location.

Impact of Failed (OPEN) valves

Failed valves Require Extensive Shutdowns In this work, a valve is referred to as "failed" when it cannot be closed when required, hence the term failed (OPEN). Isolation valves can be considered to have failed for various reasons including being paved over, cannot be operated for a variety of reasons, or cannot be located. The more failed valves there are, the more unintended isolations there will be, and the more valves need to be operated to isolate a given segment. This leads to the observation that the number of valves that must be operated when the sole valve between two segments fails can be given by

$$N = N_1 + N_2 - 2 \tag{1}$$

for looped areas, where N is the number of valves to be closed off to isolate a segment in the event of a broken valve between two adjacent segments, 1 and 2. N_1 is the number of valves to close segment 1 and N_2 is the number of valves to close segment 2. Graphically, an inoperable valve between two segments means that the two segments become merged into one. The proof of this equation is as follows.

Proof Let a connected graph, G, with a number of nodes greater than two, represent a water network, where the nodes represent segments and the edges represent valves. Let G contain two adjacent nodes, 1 and 2, representing two adjacent segments, 1 and 2, in a water network. The degree, k, is the number of edges that are incident on a node (i.e., the number of connections). Let the degree of node 1 be denoted by k_1 and the degree of node 2 be denoted by k_2. Because node 1 and node 2 are connected to each other, k_1 and k_2 each account for the edge between nodes 1 and 2. Let a modified graph from G, denoted by G', contain nodes 1 and 2 merged as a super-node, $1 + 2$ (see Figure 3). The degree of node $1 + 2$ in G', k_{1+2}, is equivalent to $k_1 + k_2 - 2$, where k_1 and k_2 are the degrees of node 1 and node 2 from the original graph, G. Two is subtracted from the total degree because when node 1 and node 2 are merged, they are considered as a single node without a connection between them (i.e., there is no longer node 1 or node 2, but only node $1 + 2$). Therefore, the number of edges that must be removed from G to disconnect a subgraph consisting of only nodes 1 and 2 is also equal to $k_1 + k_2 - 2$. Meaning that to isolate segment 1 or segment 2 in a water network where segments 1 and 2 are connected by a broken valve, $N_1 + N_2 - 2$ valves must be operated, where N_x is the number of valves connected to a given segment x. This illustrates how a failure of the valve between segments 1 and 2 can greatly impact the extent of the outage caused by a failure in either segment 1 or 2.

There are two exceptions to Equation 1. The first is when either segment 1, segment 2, or both are connected to a tree structure of segments that are fed by the loop containing segments 1 and 2. The second is when there is more than one valve between the two segments that will be merged as a single node in their corresponding graph representation due to a at least one valve failure. In this case, the number of valves that must be operated can be generalized to Equation 2:

$$N = N_1 + N_2 - 2N_{1,2} \tag{2}$$

where N is the number of valves that must be operated, N_1 is the number of valves connected to segment 1, N_2 is the number of valves connected to segment 2, and $N_{1,2}$ is the number of valves connecting segments 1 and 2.

Thus far, the segment-valve graph has been referred to as undirected. Considering the segment-valve graph as directed, with link direction indicating flow through, observations are made in the following section related to the operation of valves that isolate segments that bridge loops and the tree structures they feed.

Identifying of Valves Needed to close Tree Segments

A segment can be part of a loop on the upstream side but on the downstream side (the side that is isolated from any source), it can be the beginning of a tree. The standard practice for such segments is only closing valves upstream of the segment to isolate it. It is not necessary to operate downstream valves to isolate a segment. Possible reasons include time constraints, and the location of valves and traffic disruption. Though uncommon, utilities may opt to close downstream valves.

Identifying segments that are part of a loop and have a downstream tree structure can be achieved by identifying and mapping which valves are connected to loops (called cycles in graph theory) and which are the beginning of a tree and do not need to be closed. To identify those valves that are located on the downstream side of a segment, first, articulation points are identified. An articulation point (also called a cut vertex) is any node (i.e., segment) whose removal disconnects the graph. Physically, an articulation point is a segment, which when shut down, also isolates downstream segments beyond itself, if there is not a source downstream. Any segment in a tree or serving as the root of a tree is an articulation point. In a completely looped system, there are no articulation points. The $articulation\_points$ function within the *networkx* package used to identify articulation points.

A cycle is a path in the network that begins and ends at the same node (in this case, the same segment) without repeating segments. The cycle basis of a graph is a minimal collection of cycles such that any cycle in the network can be written as a sum of cycles in the basis (a basis is a set of elements from which all other elements can be derived as a combination of that set of elements). In other words, the cycle basis of a given WDN, contains all the segments that are part of loops. The cycle basis of the graph is identified using the $cycle\_basis$ function within *networkx*.

Then, to exclude any segments from the articulation points that are not also part of a cycle, the intersection between the sets of segments in the articulation points and cycle basis is identified and stored. Thus far, the method does not account for the location of the water sources in the network.

With an additional step, we can exclusively identify which valves are connected to loops and which are the beginning of a tree and do not need to be closed. The subgraphs containing segments with water sources must be identified and excluded from the intersection of the articulation points and cycle basis. The *biconnected_component_subgraphs* function in networkx is used to identify subgraphs that contain segments with water sources; that way, only segments with downstream trees are included and not upstream trees.

CASE STUDY

The WDN used as a case study has previously been studied by Liu et al. (2017). The network has 279 pipes and 188 junctions. The segment-valve topology is obtained for three valving scenarios for this network (N valves at each intersection, N-1 valves at each intersection and a scenario with fewer than $N - 1$ which is referred to as scarce valving). The N valve rule refers to the most complete allocation of valves, with valves located at the end of each pipe (i.e., at the junction) and the $N - 1$ valve rule refers to one less valve than the number of pipes at each junction located at the given junction. The scarce valving scenario refers to the WDN having fewer valves installed than according to the $N - 1$ rule (but more than $N - 2$). N valves corresponds to 183 valves and 157 segments, $N - 1$ valves corresponds to 130 valves and 104 segments, and scarce valves corresponds to 91 valves and 65 segments. The pipe-junction topology of the N-valve scenario is shown in Figure 4. Figure 5 is a display of the segments highlighted in the WaterGEMS software. Figure 6 is the corresponding segment-valve display in python from which the reachability matrix is constructed.

RESULTS AND DISCUSSION

Reachability Matrix

The importance and vulnerability indexes for three different valving scenarios of the case study are displayed in Figures 7, 8, and 9, where larger and darker nodes indicate higher importance (left) and vulnerability (right). The axes represent the x and y coordinates of segment centroids and the legend represents a normalized index for each scenario, where 0 is the minimum value and 1 is the maximum value. The segment identifiers are labeled in each node. For high resolution maps, refer to the supplementary information. Such figures make it easy to identify important (i.e., critical) and vulnerable segments. It is also possible to compare how important and vulnerable segments are in a given location for different scenarios. For example, the location with the three circled segments (11, 13, 24 in N and $N-1$ and 10, 12, 22 in scarce) increases in vulnerability as the number of valves decreases. On the other hand, the importance of a given segment decreases as the number of valves increase. Due to the reduction of valves and increased size of segments, there are more critical segments. In another example, the location of Segment 23 in Figure 8 becomes more critical than it was in Figure 7 (i.e., has a higher importance index as indicated by the darker color and larger node size). These results are consistent with the findings from Liu et al. (2017) and Atashi et al. (2020) that demonstrated N valving was more resilient than $N-1$ valving for a given network. This method adds to these previous works by allowing spatial comparison of network locations in addition to the comparison of overall global network-wide performance.

Validation by Comparison with WaterGEMS

The importance index values significantly correlated with demand shortfall values from WaterGEMS. Figure 10 shows that there is a strong correlation between the importance index and the model-based system demand shortfall for the system, for the three valving scenarios ($p \leq 0.001$). The standard error for N and $N-1$ valving scenarios is 2 percent, and for the scarce valving scenario is 3 percent.

Properties of the reachability matrix provide a rough method to quickly evaluate the impacts segment closures or failed valves may have, given a valving scenario of a network. This can help identify areas needing additional study. Alternatively, it can narrow down the number of scenarios to simulate for failure analysis.

Impact of Failed (OPEN) valves

Failed valves cause extensive segment shutdowns. The more failed valves there are, the more unintended isolations there will be, and the more valves need to be operated to isolate a given segment. For example, in Figure 11, consider that segment 9 must be isolated but the valve between segment 9 and segment 121 is not operable. Instead of simply closing two valves, between 125 and 9 and 121 and 9, now the valves between 90 and 121 and 6 and 121 must be closed. This also means that water cannot pass quickly from segment 90 to segment 6. This provides impetus for identifying valves needed to close tree segments and identifying loops that feed trees.

Identifying Valves Needed to Close Tree Segments

As an example, in Figure 12, if segment 138 must be isolated, the valves between segment 138 and 112 and segment 138 and 31 must be isolated because the water source is to the north of 138. However, valves in the tree downstream of segment 138 do not need to be closed off because there is not a water source between the segments downstream and 138 itself. This observation indicates that, in contrast to segments attached only to loops where the sum of the number of valves determines how many valves must be closed to shut down that segment, when some segments are the root of a tree, there is no need to shut down valves in the downstream direction. A segment can be part of a loop on the upstream side but on the downstream side (the side that is isolated from any source), it can be the beginning of a tree.

Interestingly, from the reachability matrix, for segments that are part of loops without being the beginning of a tree, the importance index values would always be 2. Whereas for segments in loops that are also the beginning of trees, the importance index values would always be greater than 2. Segments with importance values of 2 would automatically be eliminated from the set of segments that are both part of a loop and the beginning of a tree.

Identifying Loops that Feed Trees

For the scarce valving scenario of the case study network shown in Figure 13a, the articulation points are shown in Figure 13b. While many of the segments in the set of articulation points are actually important as indicated by their importance index, there are some segments that can disconnect the graph by only isolating a single segment making them less important (e.g., segment 12 circled in the bottom right of Figure 9b in comparison to segment 60, upstream of it). All of the segments with valves that do not necessarily need to be shut off are included in the set of articulation points. But the set of articulation points can also include segments that are not also part of a loop, such as segments 12, 43, 15 and 16 (in Figure 13b). After obtaining the intersection of the segments in the sets of articulation points and cycle basis, shown in Figures 13c and 13d, segments that are not part of loops are no longer highlighted.

This provides different information than the reachability matrix. Though the location of segment 23 has a high importance index value as shown in Figure 9a and is an articulation point that can disconnect the WDN, if it must be isolated, all valves (upstream and downstream) must be closed because it does not have tree segments downstream of itonly loops. That is why the location of segment 23 is not highlighted in Figure 13e, after the location of water sources is accounted for. This result can help utilities focus on valves that must actually be shut off. As this process is conducted in python, the programming language used in ArcGISit is easy to integrate in the GIS systems utilities already use.

APPLICATIONS TO REAL SYSTEMS

The proposed methodology has been applied to real WDNs, System 1 and System 2 to maintain confidentiality. The primary limitation of the reachability matrix methodology occurs when there are many directional components (e.g., PRVs) because the model assumes flow occurs in both directions, which occurred with one of the real models tested. However, even without accounting for directionality, the correlation coefficient between the matrix importance index and system demand shortfall is comparable to the values obtained by Balekelayi and Tesfamariam (2019) and Meng et al. (2018). System 1 has two reservoirs, two tanks, six pumps, seven PRVs, three TCVs, and 1236 isolation valves. The segment-valve topology is displayed in Figure 14a. When directionality of PRVs was accounted for (i.e., a directed graph) in generating the reachability matrix, the correlation coefficient between the importance index and WaterGEMS criticality of segments is higher than that obtained by Balekelayi and Tesfamariams methodology ($r = 0.6$ as opposed to 0.37) while using a larger network (System 1). Similarly, the correlation coefficient is higher than most of the coefficients reported by Meng et al. (2018).

When the data from the boundary elements (e.g., directional elements, reservoirs, tanks) is not included in the analysis, the correlation coefficient becomes more than double ($r = 0.86$). However, as can be seen from the scatterplot in Figure 14b, the matching between the importance index and system demand shortfall is more accurate for higher values of importance. Utilities already know to pay attention to the segments with boundary elements and they do not need a preliminary screening to indicate that these elements are important.

System 2 has two reservoirs, four tanks, six pumps, 1730 isolation valves, a PRV, and a PSV. For this WDN, a correlation coefficient of 0.89 is obtained between the importance index and system demand shortfall after removing the segments with boundary elements from the analysis. This suggests that proximity to the boundary elements may also impact how accurately the importance index of a segment represents its actual importance. More experiments with water distribution networks containing directional elements must be conducted to discern this.

For different valving scenarios of the real networks, the indexes from the matrix were able to capture that the more valving there is in the network (applied according to rules such as N and $N-1$), the less vulnerable and important the most vulnerable and important segments become for a given network which has been shown before in Liu et al. (2017). See the supplementary information for box plots from the analysis of variance between valving scenarios of a given network.

These results indicate that graph theory measures can in fact be used for preliminary screening for failure scenario analysis or initial design. As it pertains to the articulation points, and identification of segments that have valves downstream that do not necessarily need to be operated, the physical correspondence between the WDNs and the graph theory measures indicates that the measures can still be used for real systems, when quick screening is needed to facilitate decision-making.

SUMMARY

Using a segment-valve topology is more powerful than a pipe-junction topology when analyzing the reliability of water distributions systems. This study advanced the state-of-the-art in reliability analysis based on segments and valves in several areas by developing:

A method to quickly display segment and valve topology in a way that makes it easy to graphically identify and display important and vulnerable segments, A method to quickly screen a network to identify important and vulnerable segments that not only quantifies these properties but also accounts for whether water sources are continuous sources with relatively continuous supply vs. those (e.g. tanks) that are ephemeral in nature and can only serve as a supply for a limited time,

A method to identify the number of valves to be operated to isolate a segment depending on the topology of the system, and A method to identify segments that will be isolated when other segments are closed.

The methods are both theoretically justified and useful for practicing engineers and operators. Applying CNA to the segment-valve topology demonstrates the impact certain valve failures may have on segment isolation and helps identify valves that do not need to be closed when segments must be isolated. While the pipe-junction topology of WDNs is necessary for design and hydraulic analysis, the segment-valve topology provides insight that may provide decision-support for utilities attempting to ensure their system is resilient. Future research can expand on the current study by including population-based metrics and likelihood of failure in the analysis. More WDN configurations under a wide array of loading conditions can also be tested to more clearly establish limitations of the methods.

DATA AVAILABILITY STATEMENT

- All data, models, or code generated or used during the study are available from the corresponding author by request.
- Some or all data, models, or code generated or used during the study are available in a repository or online in accordance with funder data retention policies (https://github.com/N-abdel/segment-valve-visualization).

ACKNOWLEDGEMENTS

N. Abdel-Mottaleb is supported by the National Science Foundation under Grant Number 1638301. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

REFERENCES

Abdel-Mottaleb, N., Ghasemi Saghand, P., Charkhgard, H., and Zhang, Q. (2019). "An exact multiobjective optimization approach for evaluating water distribution infrastructure criticality and geospatial interdependence." *Water Resources Research*, 55(7), 5255–5276.

Abdel-Mottaleb, N. and Walski, T. (2020a). "Identifying vulnerable and critical water distribution segments." *World Environmental and Water Resources Congress 2020: Hydraulics, Waterways, and Water Distribution Systems Analysis*, American Society of Civil Engineers Reston, VA, 329–339.

Abdel-Mottaleb, N. and Walski, T. (2020b). "Segment valve visualization, <https://github.com/N-468abdel/segment-valve-visualization>.

Abdel-Mottaleb, N. and Zhang, Q. (2019). "Logical architecture of water distribution networks." *World Environmental and Water Resources Congress 2019: Hydraulics, Waterways, and Water Distribution Systems Analysis*, American Society of Civil Engineers Reston, VA, 506–519.

Agathokleous, A., Christodoulou, C., and Christodoulou, S. E. (2017). "Topological robustness and vulnerability assessment of water distribution networks." *Water Resources Management*, 31(12), 4007–4021.

Alvisi, S., Creaco, E., and Franchini, M. (2011). "Segment identification in water distribution systems." *Urban Water Journal*, 8(4), 203–217.

Atashi, M., Ziaei, A. N., Khodashenas, S. R., and Farmani, R. (2020). "Impact of isolation valves location on resilience of water distribution systems." *Urban Water Journal*, 17(6), 560–567.

Balekelayi, N. and Tesfamariam, S. (2019). "Graph-theoretic surrogate measure to analyze reliability of water distribution system using bayesian belief network–based data fusion technique." *Journal of Water Resources Planning and Management*, 145(8), 04019028.

Bentley Systems (2020). "Bentley WaterGEMS.

Bouchart, F. and Goulter, I. (1991). "Reliability improvements in design of water distribution networks recognizing valve location." *Water Resources Research*, 27(12), 3029–3040.

Butler, D., Ward, S., Sweetapple, C., Astaraie-Imani, M., Diao, K., Farmani, R., and Fu, G. (2017). "Reliable, resilient and sustainable water management: the safe & sure approach." *Global Challenges*, 1(1), 63–77.

Choi, Y. H., Jung, D., Jun, H., and Kim, J. H. (2018). "Improving water distribution systems robustness through optimal valve installation." *Water*, 10(9), 1223.

Creaco, E., Franchini, M., and Alvisi, S. (2012). "Evaluating water demand shortfalls in segment analysis." *Water resources management*, 26(8), 2301–2321.

Creaco, E. and Haidar, H. (2019). "Multiobjective optimization of control valve installation and dma creation for reducing leakage in water distribution networks." *Journal of Water Resources Planning and Management*, 145(10), 04019046.

Deb, A. K., Snyder, J., Hammell, J. O., Jun, H., and McCammon, S. B. (2006). *Criteria for valve location and system reliability*. American Water Works Association.

Diao, K., Sweetapple, C., Farmani, R., Fu, G., Ward, S., and Butler, D. (2016). "Global resilience analysis of water distribution systems." *Water research*, 106, 383–393.

Dziedzic, R. and Karney, B. (2014). "Water distribution system performance metrics." *Procedia Engineering*, 89, 363–369.

Gillies, S. (2013). "The shapely user manual, <http://tolerity.org/shapely/manual.html>. Accessed: 2019-12-06.

Giustolisi, O. (2020). "Water distribution network reliability assessment and isolation valve system." *Journal of Water Resources Planning and Management*, 146(1), 04019064.

Giustolisi, O., Ridolfi, L., and Simone, A. (2019). "Tailoring centrality metrics for water distribution networks." *Water Resources Research*, 55(3), 2348–2369.

Giustolisi, O. and Savic, D. (2010). "Identification of segments and optimal isolation valve system design in water distribution networks." *Urban Water Journal*, 7(1), 1–15.

Giustolisi, O., Simone, A., and Ridolfi, L. (2017). "Network structure classification and features of water distribution systems." *Water Resources Research*, 53(4), 3407–3423.

Hagberg, A., Swart, P., and S Chult, D. (2008). "Exploring network structure, dynamics, and function using networkx." *Report no.*, Los Alamos National Lab.(LANL), Los Alamos, NM (United States).

He, X. and Yuan, Y. (2019). "A framework of identifying critical water distribution pipelines from recovery resilience." *Water Resources Management*, 33(11), 3691–3706.

Hwang, H. and Lansey, K. (2017). "Water distribution system classification using system characteristics and graph-theory metrics." *Journal of Water Resources Planning and Management*, 143(12), 04017071.

Jacobs, P. and Coulter, I. (1991). "Estimation of maximum cut-set size for water network failure." *Journal of Water Resources Planning and Management*, 117(5), 588–605.

Jacobs, P. and Goulter, I. C. (1989). "Optimization of redundancy in water distribution networks using graph theoretic principles." *Engineering optimization*, 15(1), 71–82.

Jones, E., Oliphant, T., Peterson, P., et al. (2016). "Scipy: Open source scientific tools for python, 2001.

Jun, H. (2005). "Strategic valve locations in a water distribution system.

Jun, H. and Loganathan, G. (2007). "Valve-controlled segments in water distribution systems." *Journal of Water Resources Planning and Management*, 133(2), 145–155.

Klise, K. A., Bynum, M., Moriarty, D., and Murray, R. (2017). "A software framework for assessing the resilience of drinking water systems to disasters with an example earthquake case study." *Environmental modelling & software*, 95, 420–431.

Liu, H., Walski, T., Fu, G., and Zhang, C. (2017). "Failure impact analysis of isolation valves in a water distribution network." *Journal of Water Resources Planning and Management*, 143(7), 04017019.

Mahmoud, H. A., Kapelan, Z., and Savić, D. (2018). "Real-time operational response methodology for reducing failure impacts in water distribution systems." *Journal of Water Resources Planning and Management*, 144(7), 04018029.

Meng, F., Fu, G., Farmani, R., Sweetapple, C., and Butler, D. (2018). "Topological attributes of network resilience: A study in water distribution systems." *Water research*, 143, 376–386.

Pagani, A., Meng, F., Fu, G., Musolesi, M., and Guo, W. (2020). "Quantifying resilience via multiscale feedback loops in water distribution networks." *Journal of Water Resources Planning and Management*, 146(6), 04020039.

Pagano, A., Sweetapple, C., Farmani, R., Giordano, R., and Butler, D. (2019). "Water distribution networks resilience analysis: a comparison between graph theory-based approaches and global resilience analysis." *Water Resources Management*, 33(8), 2925–2940.

Santonastaso, G., Di Nardo, A., and Creaco, E. (2019). "Dual topology for partitioning of water distribution networks considering actual valve locations." *Urban Water Journal*, 16(7), 469–479.

Shuang, Q., Zhang, M., and Yuan, Y. (2014). "Performance and reliability analysis of water distribution systems under cascading failures and the identification of crucial pipes." *PloS one*, 9(2).

Su, Y.-C., Mays, L. W., Duan, N., and Lansey, K. E. (1987). "Reliability-based optimization model for water distribution systems." *Journal of Hydraulic Engineering*, 113(12), 1539–1556.

Torres, J. M., Duenas-Osorio, L., Li, Q., and Yazdani, A. (2017). "Exploring topological effects on water distribution system performance using graph theory and statistical models." *Journal of Water Resources Planning and Management*, 143(1), 04016068.

Ulusoy, A.-J., Stoianov, I., and Chazerain, A. (2018). "Hydraulically informed graph theoretic measure of link criticality for the resilience analysis of water distribution networks." *Applied network science*, 3(1), 31.

Vasilic, Z., Stanic, M., Kapelan, Z., Ivetic, D., and Prodanovic, D. (2018). "Improved loop-flow method for hydraulic analysis of water distribution systems." American Society of Civil Engineers.

Vugrin, E. D., Warren, D. E., Ehlen, M. A., and Camphouse, R. C. (2010). "A framework for assessing the resilience of infrastructure and economic systems." *Sustainable and resilient critical infrastructure systems*, Springer, 77–116.

Wagner, J. M., Shamir, U., and Marks, D. H. (1988). "Water distribution reliability: analytical methods." *Journal of Water Resources Planning and Management*, 114(3), 253–275.

Walski, T. M. (1993). "Water distribution valve topology for reliability analysis." *Reliability engineering & system safety*, 42(1), 21–27.

Walski, T. M. (1994). "Valves and water distribution system reliability." *AWWA National Convention, New York, NY*.

Walski, T. M. (2011). "How many isolation valves are needed in a water distribution system?." *Urban Water Management: Challenges and Oppurtunities-11th International Conference on Computing and Control for the Water Industry, CCWI 2011*.

Walski, T. M., Weiler, J. S., and Culver, T. (2008). "Using criticality analysis to identify impact of valve location." *Water Distribution Systems Analysis Symposium 2006*, 1–9.

Yang, S.-L., Hsu, N.-S., Louie, P. W., and Yeh, W. W. (1996). "Water distribution network reliability: Connectivity analysis." *Journal of Infrastructure Systems*, 2(2), 54–64.

Yazdani, A. and Jeffrey, P. (2012a). "Applying network theory to quantify the redundancy and structural robustness of water distribution systems." *Journal of Water Resources Planning and Management*, 138(2), 153–161.

Yazdani, A. and Jeffrey, P. (2012b). "Water distribution system vulnerability analysis using weighted and directed network models." *Water Resources Research*, 48(6).

Zeng, F., Li, X., and Li, K. (2017). "Modeling complexity in engineered infrastructure system: Water distribution network as an example." *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 27(2), 023105.

Zischg, J., Reyes-Silva, J. D., Klinkhamer, C., Krueger, E., Krebs, P., Rao, P. S. C., and Sitzenfrei, R. (2019). "Complex network analysis of water distribution systems in their dual representation using isolation valve information." *World Environmental and Water Resources Congress*, ASCE Pittsburgh, Pennsylvania, 484–497.

Zweig, K. A. et al. (2016). *Network analysis literacy*. Springer.

FIG1

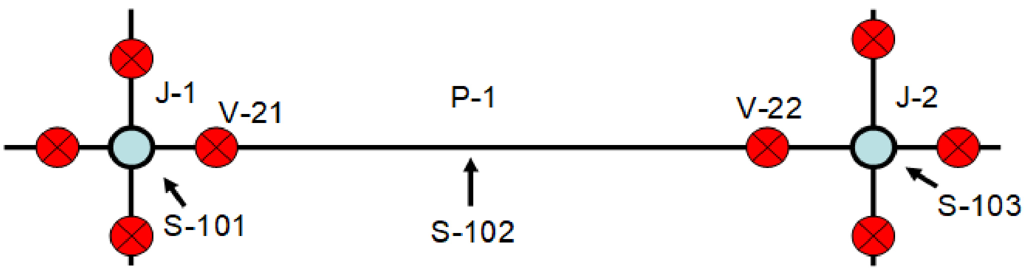
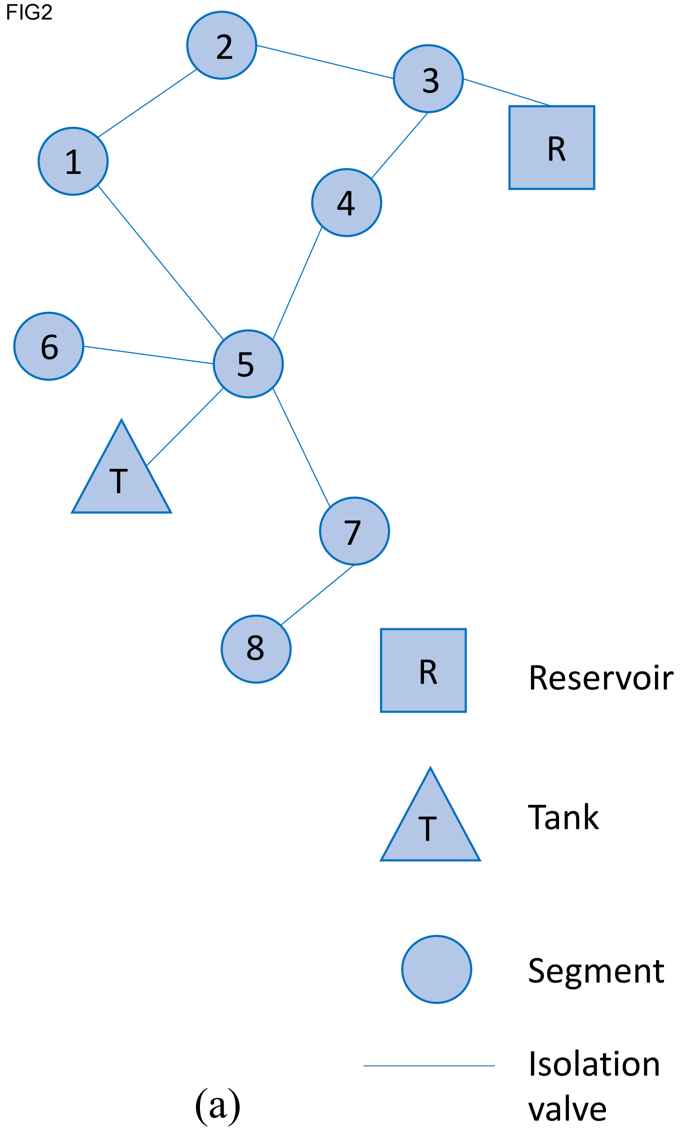


FIG2



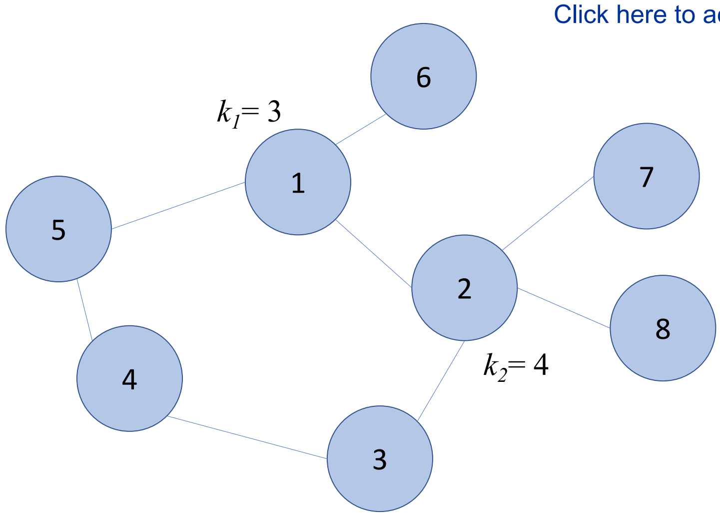
(a)

Affected

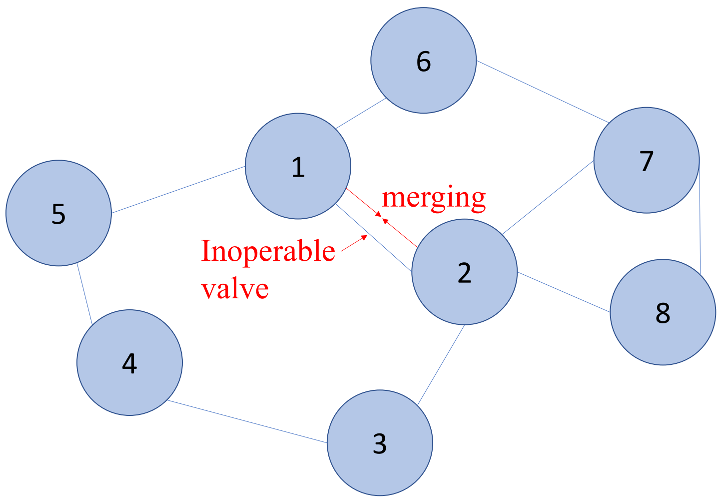
Isolated

Segment	1	2	3	4	5	6	7	8	Importance
1	2	0	0	0	0	0	0	0	2
2	0	2	0	0	0	0	0	0	2
3	1	1	2	1	1	1	1	1	9
4	0	0	0	2	0	0	0	0	2
5	0	0	0	0	2	2	2	2	8
6	0	0	0	0	0	2	0	0	2
7	0	0	0	0	0	0	2	2	4
8	0	0	0	0	0	0	0	2	2
Vulnerability	3	3	2	3	3	5	5	7	

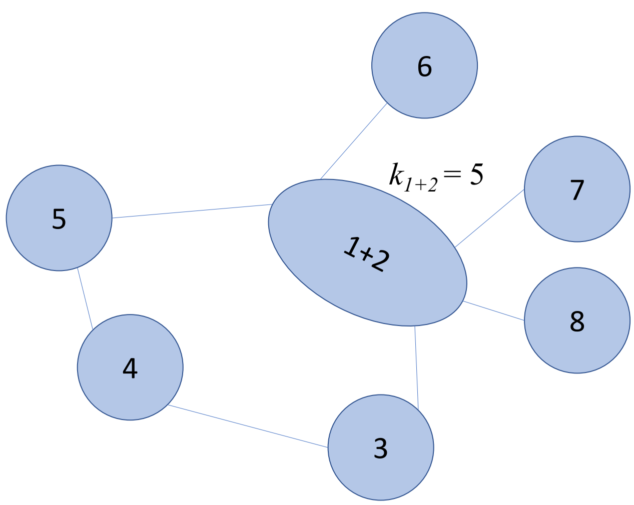
(b)


6
$k_1 = 3$
1
7
5
2
8
4
$k_2 = 4$
3

(a)


6
1
7
5
merging
Inoperable
valve
2
8
4
3

(b)


6
$k_{1+2} = 5$
7
5
1+2
8
4
3

(c)

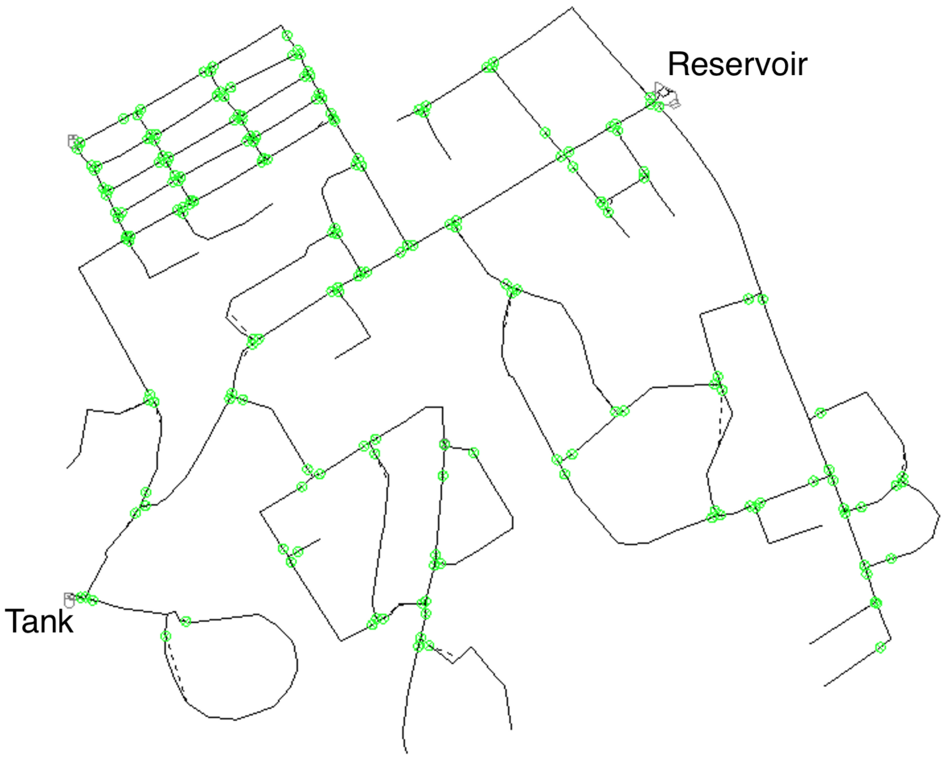

Reservoir
Tank

FIG5

Click here to access/download;Figure;FIG5_.pdf

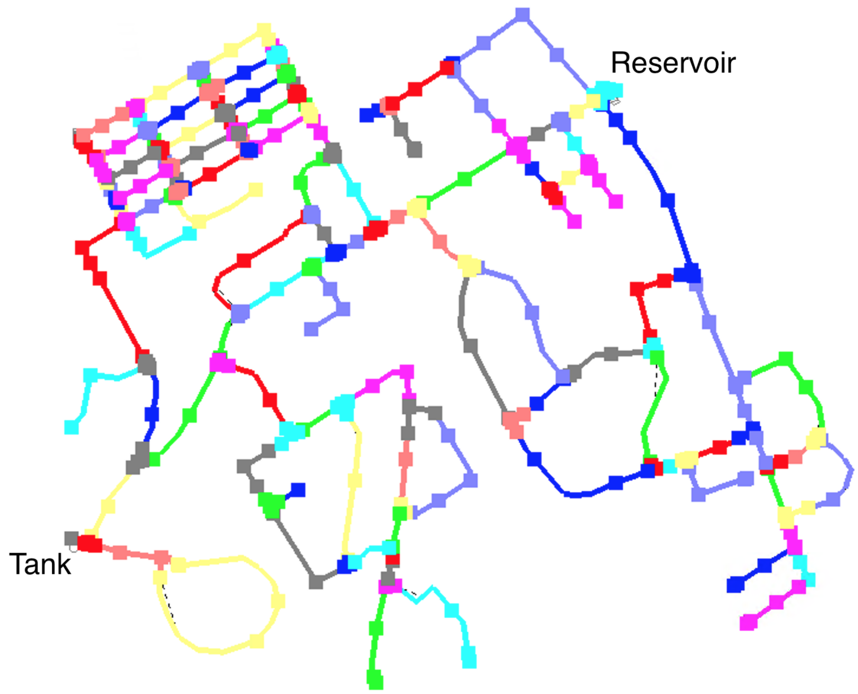


FIG6

Click here to access/download;Figure;FIG6.pdf

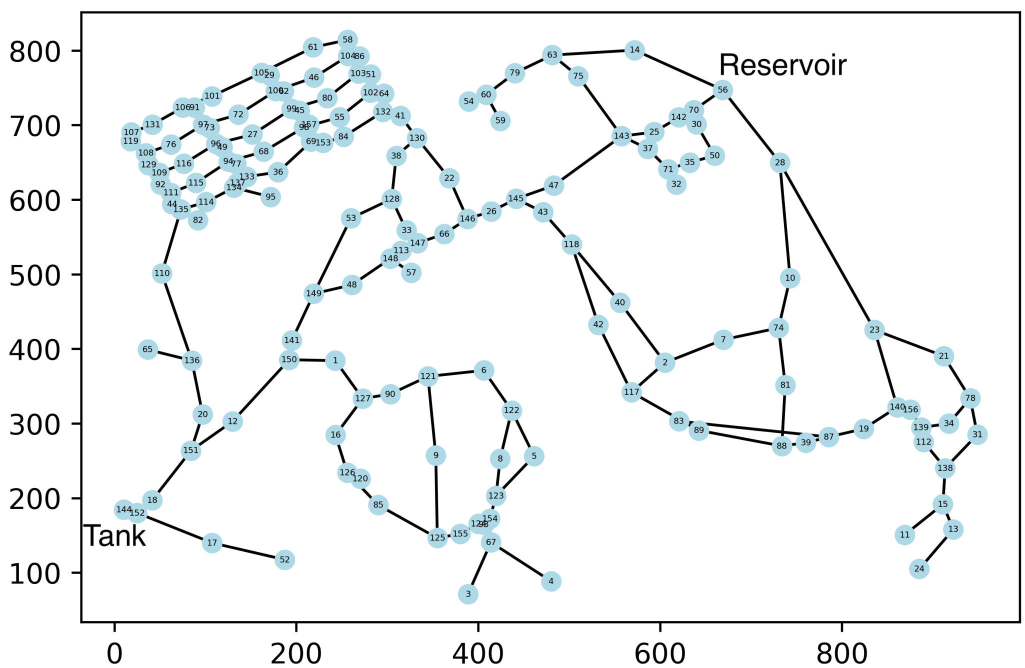


FIG7

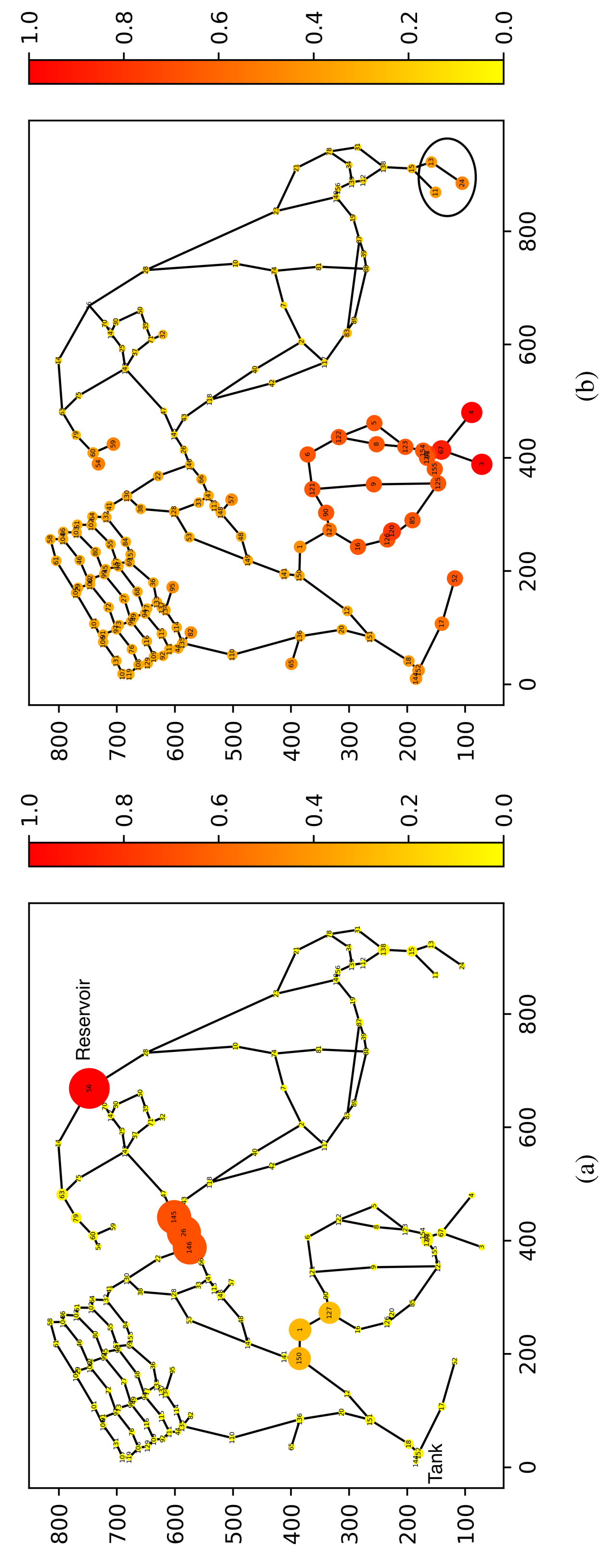


FIG8

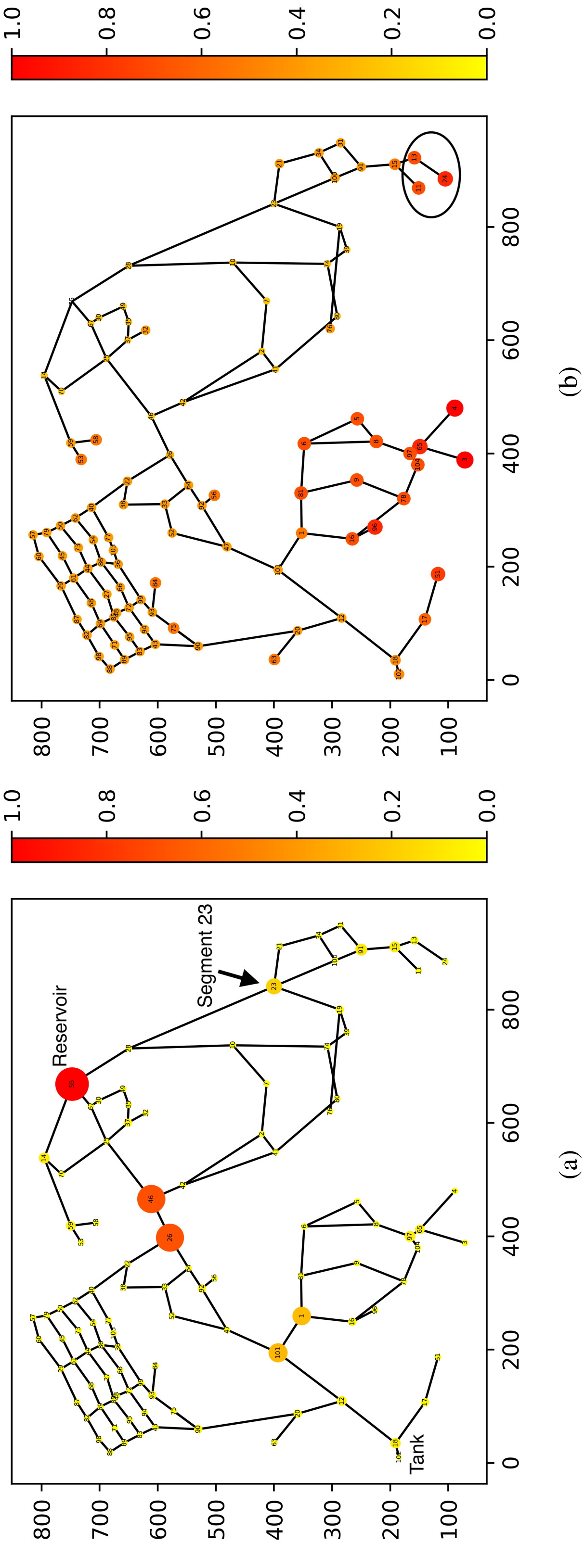
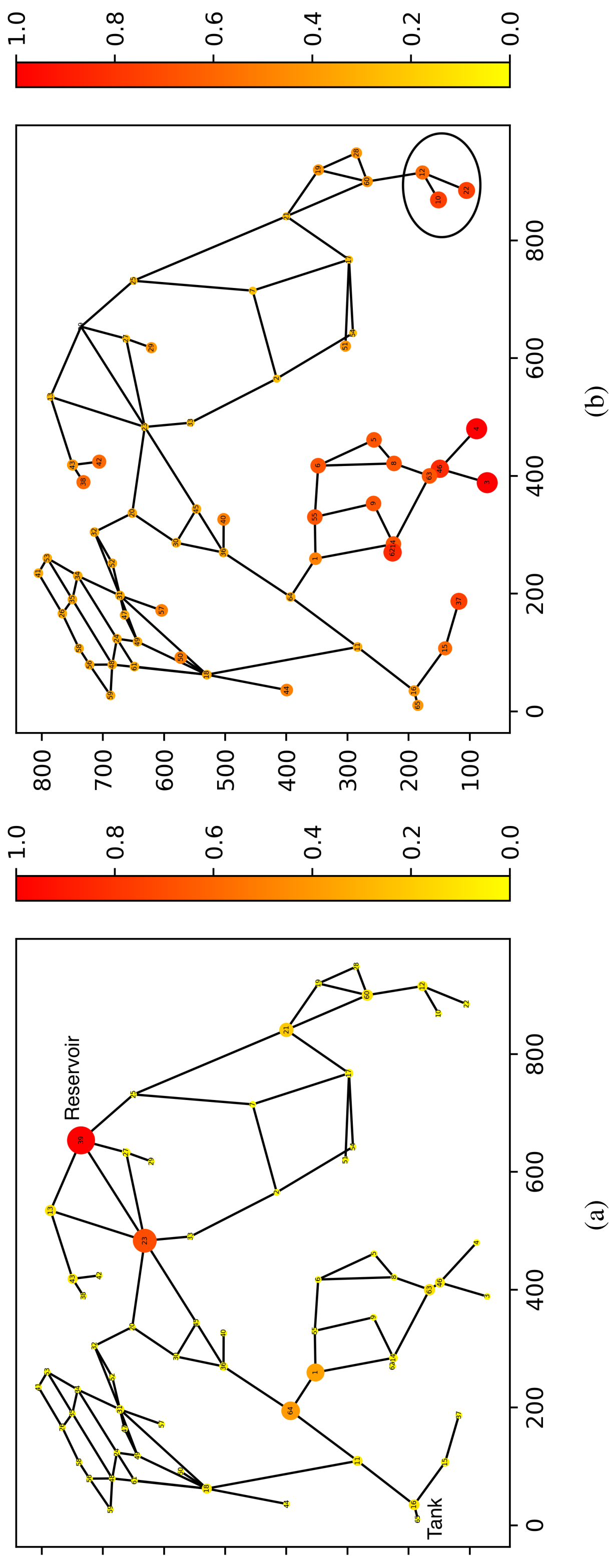
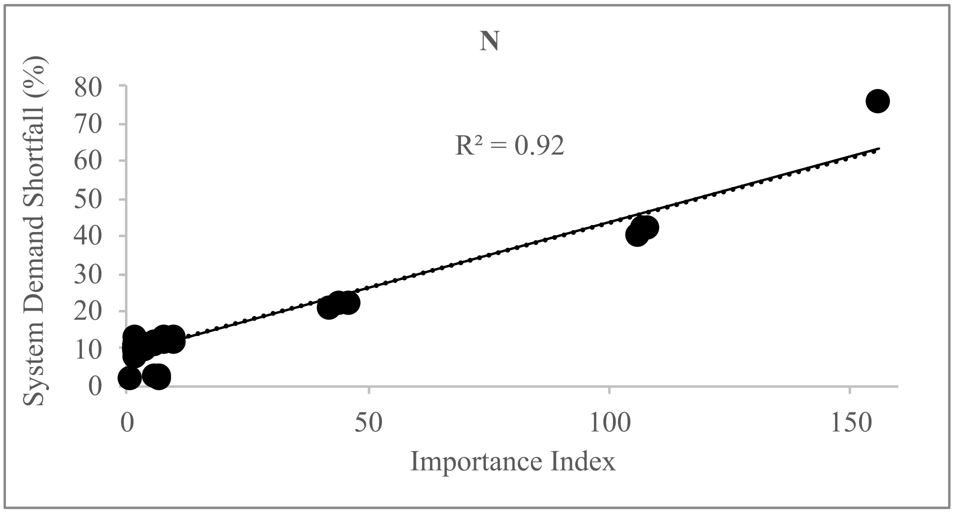
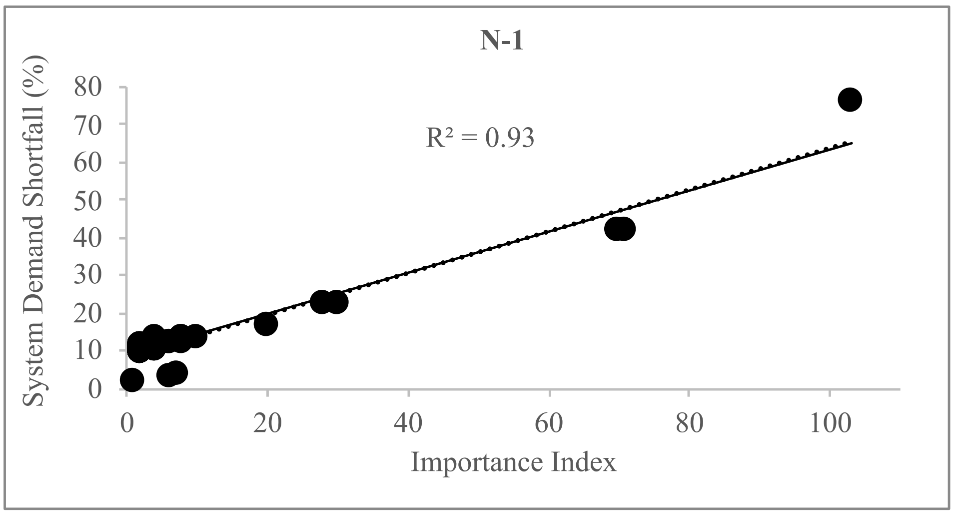


FIG9

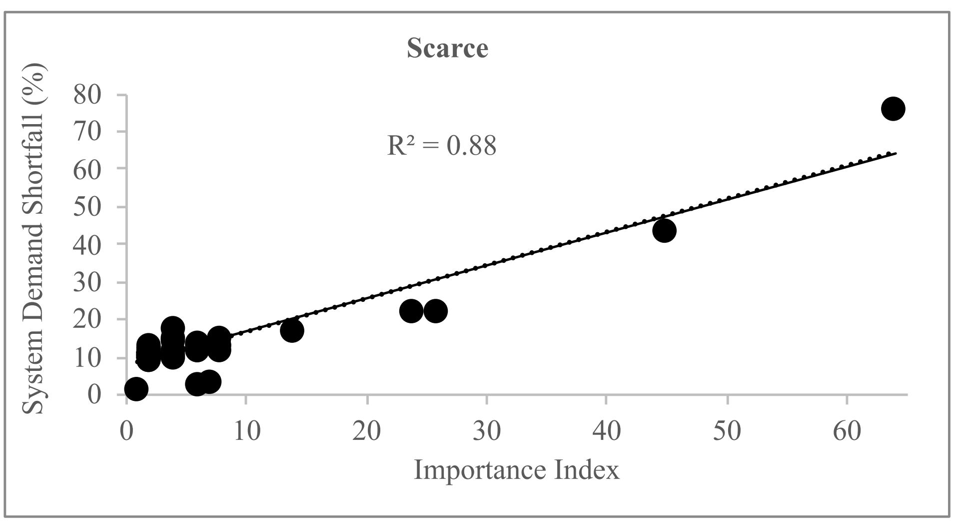

Reservoir
Tank
0 200 400 600 800
100 200 300 400 500 600 700 800
0.0 0.2 0.4 0.6 0.8 1.0
(a)
(b)


N
R² = 0.92
System Demand Shortfall (%)
80
70
60
50
40
30
20
10
0
0
50
100
150
Importance Index

(a)


N-1
R² = 0.93
System Demand Shortfall (%)
80
70
60
50
40
30
20
10
0
0
20
40
60
80
100
Importance Index

(b)


Scarce
R² = 0.88
System Demand Shortfall (%)
80
70
60
50
40
30
20
10
0
0
10
20
30
40
50
60
Importance Index

(c)

FIG11

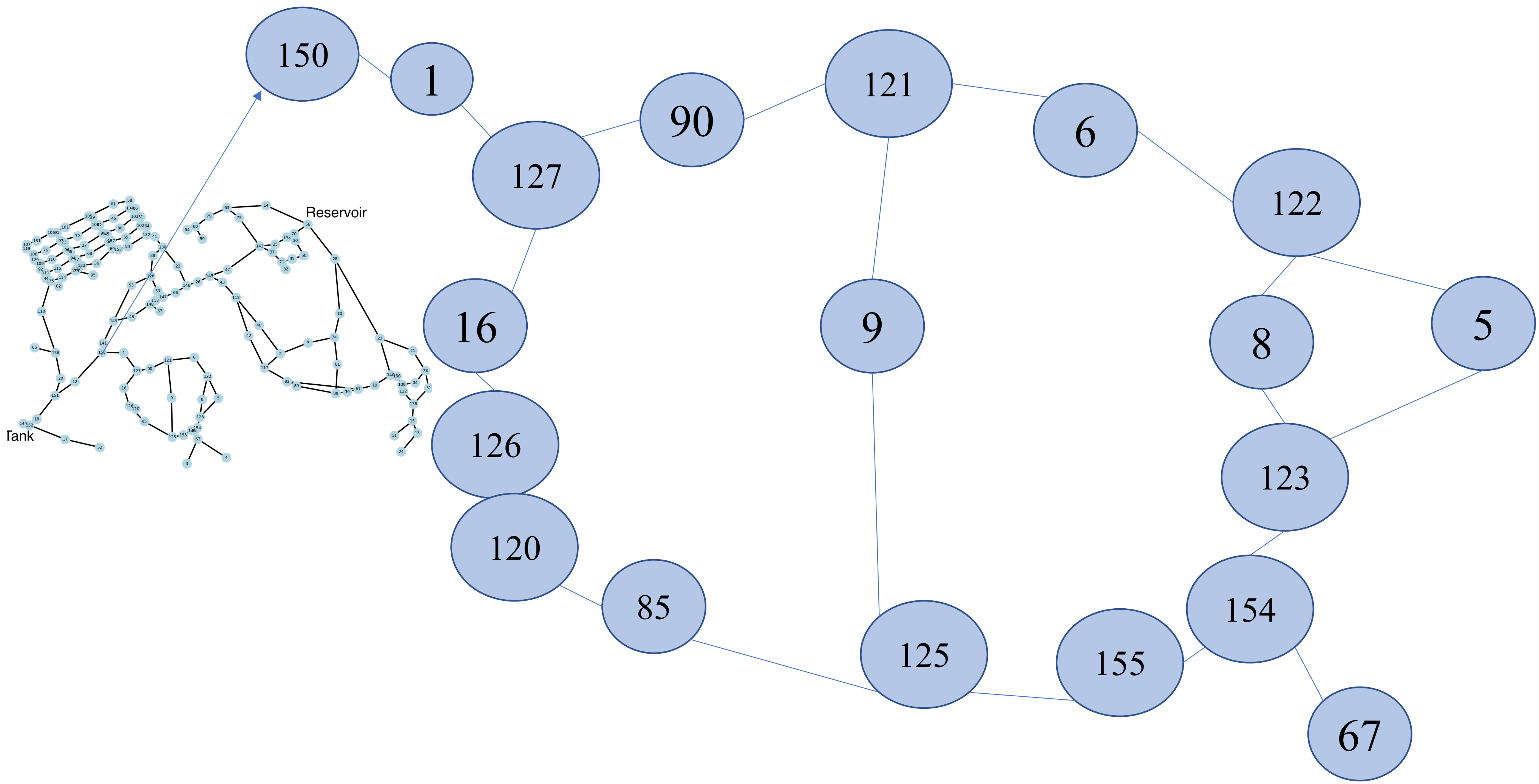
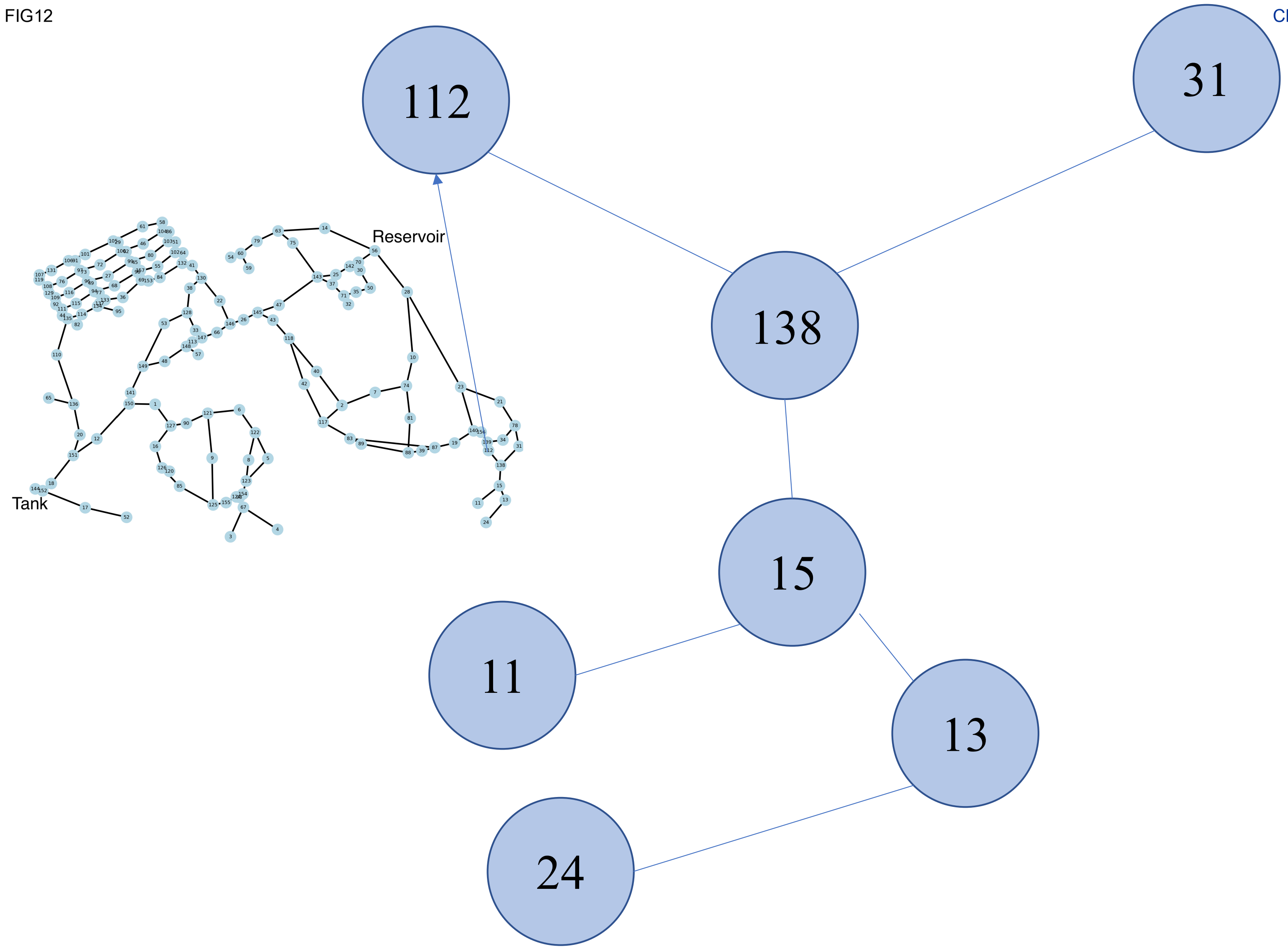
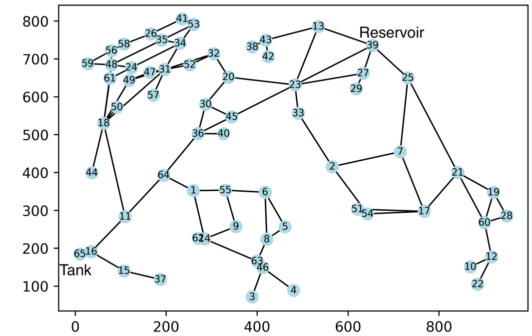
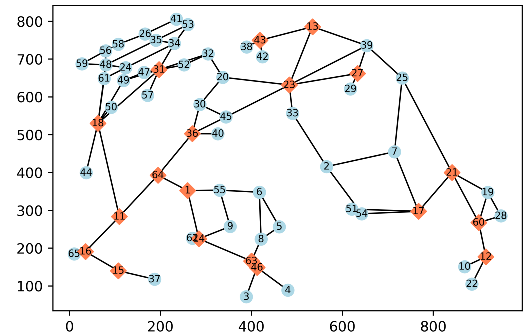

Reservoir
Tank
150
1
127
90
121
6
122
16
9
8
5
126
123
120
154
85
125
155
67

FIG12

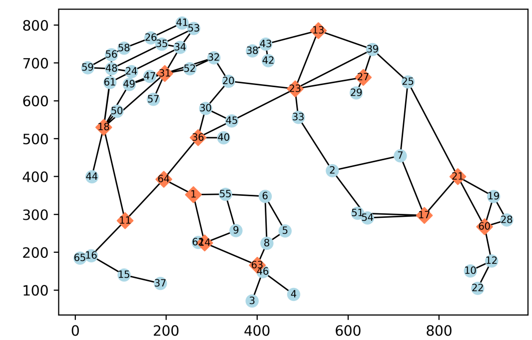




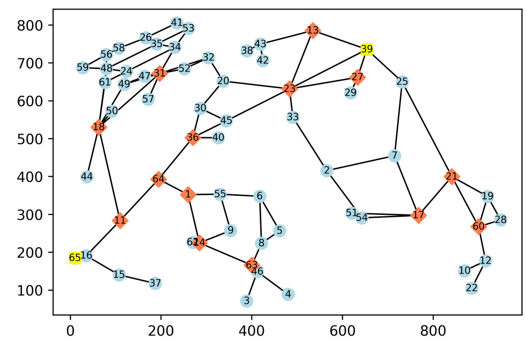
(a)



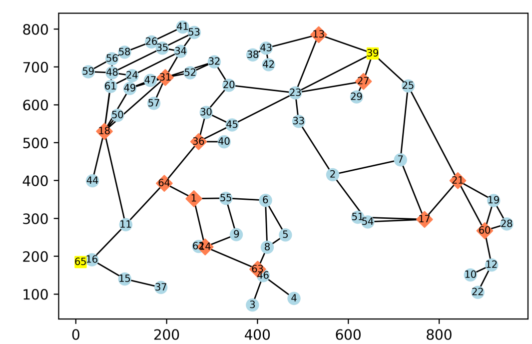
(b)



(c)

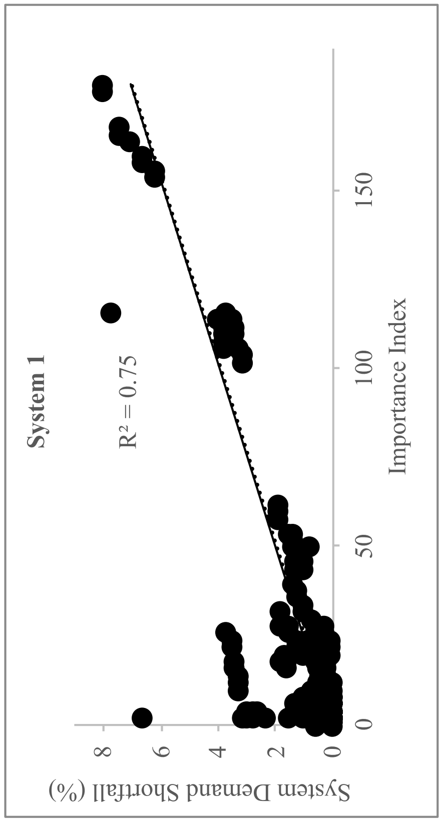


(d)

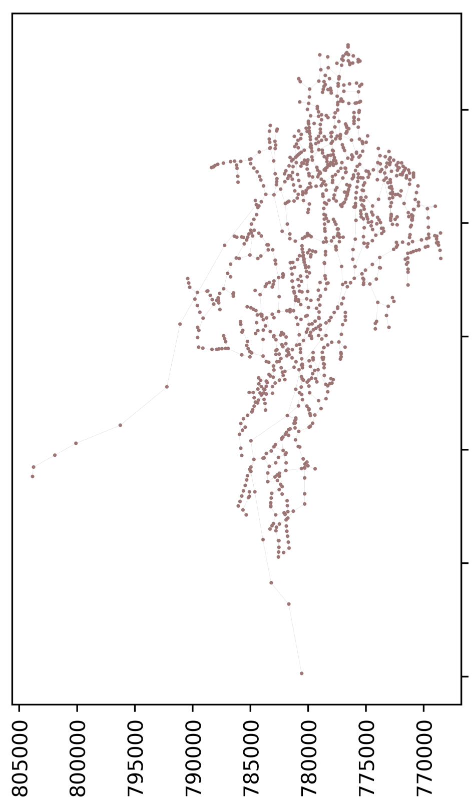


(e)

FIG14


System 1
R² = 0.75
Importance Index
System Demand Shortfall (%)
0
50
100
150
0
2
4
6
8

(b)


885000
890000
895000
900000
905000
910000
805000
800000
795000
790000
785000
780000
775000
770000

(a)

FIG. 1. Illustration of WDN segments, a single pipe can be part of several segments and segments can contain more than one pipe

FIG. 2. (a) Segment-valve topology of a small example (b) Reachability matrix

FIG. 3. Illustration of proof for a small example, showing segments 1 and 2 merging to form a single node in the segment-valve topology

FIG. 4. Pipe-junction topology of N-valve scenario highlighting valve locations

FIG. 5. Segments highlighted on the Pipe-junction topology of N-valve scenario in WaterGEMS

FIG. 6. Segment-valve topology of N valve scenario for the system shown in Figure 4, where segments are nodes and valves are edges

FIG. 7. N valve scenario, where normalized index values are represented in the legend on the right; smaller nodes correspond to lower values and larger nodes correspond to higher values; axes correspond to x,y coordinates of segment centroids (a) Importance (b) Vulnerability

FIG. 8. N-1 valve scenario, where normalized index values are represented in the legend on the right; smaller nodes correspond to lower values and larger nodes correspond to higher values; axes correspond to x,y coordinates of segment centroids (a) Importance, the location of segment 23 increases in criticality from the N valve scenario (b) Vulnerability, the bottom right tree segments increase in vulnerability from the N valve scenario

FIG. 9. Scarce valve scenario, where normalized index values are represented in the legend on the right; smaller nodes correspond to lower values and larger nodes correspond to higher values; axes correspond to x,y coordinates of segment centroids (a) Importance (b) Vulnerability

FIG. 10. Validation of importance index derived from the reachability matrix.

All three valving scenarios' segment importance values significantly correlated ($p<0.01$) with the system demand shortfalls obtained using WaterGEMS

FIG. 11. Zoomed in section of case study network, originating at source segment 150

FIG. 12. Zoomed in section of case study network, originating at source segments 112 and 31

FIG. 13. (a) Segment-valve topology of scarce valving scenario (b) Articulation points denoted with diamond shaped nodes of the scarce valving scenario (c) Intersection of articulation points and segments that are part of the cycle basis denoted with diamond shaped nodes for the scarce valving scenario (d) Intersection of articulation points and segments that are part of the cycle basis, with Reservoir (continuous source) and Tank (ephemeral source) highlighted (Segments 39 and 65, respectively) (e) Segments that are part of loops for which isolation does not necessitate downstream valve closures

FIG. 14. (a) Segment-valve topology of System 1 where the x-axis represents the x-coordinates and the y-axis represents the y-coordinates (b) x-axis represents the importance index for System 1 and the y-axis represents the system demand shortfall, where the correlation coefficient is 0.86 ($p<0.01$)