

A Graph-Theoretic Framework for Assessing the Resilience of Sectorised Water Distribution Networks

Manuel Herrera¹ · Edo Abraham² · Ivan Stoianov²

Received: 3 July 2015 / Accepted: 24 January 2016 /

Published online: 13 February 2016

© The Author(s) 2016. This article is published with open access at Springerlink.com

Abstract Water utilities face a challenge in maintaining a good quality of service under a wide range of operational management and failure conditions. Tools for assessing the resilience of water distribution networks are therefore essential for both operational and maintenance optimization. In this paper, a novel graph-theoretic approach for the assessment of resilience for large scale water distribution networks is presented. This is of great importance for the management of large scale water distribution systems, most models containing up to hundreds of thousands of pipes and nodes. The proposed framework is mainly based on quantifying the redundancy and capacity of all possible routes from demand nodes to their supply sources. This approach works well with large network sizes since it does not rely on precise hydraulic simulations, which require complex calibration processes and computation, while remaining meaningful from a physical and a topological point of view. The proposal is also tailored for the analysis of sectorised networks through a novel multi-scale method for analysing connectivity, which is successfully tested in operational utility network models made of more than 100,000 nodes and 110,000 pipes.

Keywords Water distribution networks · Resilience · Complex systems · Graph theory · Multiscale graph decomposition

✉ Ivan Stoianov
ivan.stoianov@imperial.ac.uk

Manuel Herrera
a.m.herrera.fernandez@bath.ac.uk

Edo Abraham
edo.abraham04@imperial.ac.uk

¹ EDEn - Department of Architecture and Civil Engineering, University of Bath, Bath, BA2 7AY, UK

² InfraSense Labs - Department of Civil and Environmental Engineering, Imperial College London, London, SW7 2AZ, UK

1 Introduction

Resilience can be defined as the ability of a system to maintain and adapt its operational performance in the face of failures and other adverse conditions (Laprie 2005; Strigini 2012). Depending on the application domain, different approaches are applied to assess the resilience of man-made systems. There is no single widely used definition for quantifying the resilience of water distribution networks (WDNs), a common method is to formulate the hydraulic resilience as a measure of the ability of a network to maintain supply under failure conditions. Todini (2000) proposed a resilience index based on the steady state flow analysis of WDNs and dissipated energy; consequently, the resilience of a water network was defined using a measure of the available surplus energy. Recent work extends this approach; for example, Raad et al. (2010) modelled the outcomes of this index proposing to work with alternative approaches to the original in Todini (2000); this was done to overcome computational difficulties in optimising a network design problem using Todini's resilience index as an objective function. Therefore, by using parameters of physical significance in the supply as surrogates and not the real hydraulic details, di Nardo et al. (2013a) included these surrogate indices for optimizing the resilience of sectorised water distribution networks. Baños et al. (2011) proposed an extension to Todini (2000) by considering failures under various water demand scenarios. Other approaches based on a steady state hydraulic analysis include Prasad and Park (2004) who adapted Todini's index by incorporating the effects of both surplus pressure and reliability of supply loops assessed by the variability in the diameter of pipes connected to the same node. They also applied the method to a multi-objective problem for the optimal design and rehabilitation of a water distribution network. Jayaram and Srinivasan (2008) used the same method from Prasad and Park (2004) for the analysis of networks with multiple sources. An alternative approach was proposed by Wright et al. (2015) who assess the resilience of networks with dynamically configurable topologies by estimating what percentage of operational demand can be supplied when a disruptive event occurs.

A common constraint of the above approaches is that the combination of possible failure scenarios grows exponentially as the network becomes bigger (Berardi et al. 2014) together with possible inconsistencies and uncertainties associated with hydraulic simulations (Gupta and Bhawe 1994). Trifunović (2012) explores hydraulic properties of the network based on the statistical analysis of common parameters under normal operation and proposed them as indices to assess the resilience of a water network. However, these statistical parameters may not be valid under failure conditions (Giustolisi et al. 2008). Another inherent drawback is the dependence on well calibrated (accurate) hydraulic models for large networks. Beyond the implicit non-linearity, the number of parameters involved in the hydraulic equations (Abraham and Stoianov 2015) and their large number of possible combinations introduce high complexity to the WDN calibration problem (an NP-Hard problem, which cannot be solved in polynomial time Takahashi et al. (2010).) For large WDNs, these have an expensive computational cost since, even for moderate size networks, the number of possible failure scenarios grows exponentially and are approached using either some linearising assumptions or the use of heuristics for failure simulations (Berardi et al. 2014).

This paper assesses the resilience of WDNs from a topological perspective where properties such as network configuration and redundancy in connectivity are taken into account together with physical-based flow properties. This assessment is based on a computationally efficient implementation of the K -shortest paths algorithm (Eppstein 1998). It measures in a statistically robust way how every WDN demand node would be affected by disruptions in supply. This is done by analysing alternative paths between demand nodes and

water sources (i.e. tanks and reservoirs). In addition, the paths are weighted by the hydraulic attributes of the supply routes to make the measures physically meaningful. Graph-theoretic indices have been applied to assess the resilience of water networks and used also as a surrogate measure in design problems (Yazdani et al. 2011). These surrogate indices are measures that correlate to physical/hydraulic based indices of resilience but are not necessarily based on physical indices whose parameters are either unavailable or are difficult to compute. These indices can be broadly classified into “statistical” and “spectral” measures (Gutiérrez-Pérez et al. 2013). Among the statistical indices, the meshedness coefficient (Buhl et al. 2006) provides an estimation of the topological redundancy of a network. Central-point dominance (Freeman 1977), gives a measurement of the network vulnerability to failures corresponding to central locations. Flow Entropy (Raad et al. 2010) measures the strength of supply to a node both in terms of the number of connections and their similarity. The so-called demand-adjusted entropic degree in Yazdani and Jeffrey (2012) is another alternative that uses demand on nodes and volume capacity on edges to compute a weighted entropic degree. More recently, (Liu et al. 2014) have extended the flow entropy measure to also consider the impact of pipe diameter on reliability.

The work also extends the proposed adaptation of K -shortest paths algorithm for the analysis of WDNs divided into sectors or District Metered Areas (DMAs) by utilizing multiscale network decomposition (Albert and Barabási 2002; Lee and Maggioni 2011). Demand nodes of a given DMA are aggregated into a sector-node, where a new graph (DMA-graph) is defined with vertices representing DMA areas and edges that abstract the sector-to-sector connectivity (Izquierdo et al. 2011). A mapping function is used to establish a relationship between the network graph of demand nodes and the DMA-graph, where every DMA-vertex (or sector-node) has new features such as number of customers, total demand, and an average pressure aggregated from information of individual nodes of the DMA. The physical connections between sectors, which consist of pipes and boundary valves, form the edges of this DMA-graph and define the level of its connectivity. Resilience assessments carried out on individual demand nodes in the DMAs are also linked to the corresponding sector-nodes, which are then used for a system-level resilience analysis. Considering both the network graph of demand nodes and a DMA-graph, a multiscale analysis of resilience is proposed.

2 A Graph Theory Perspective on the Definition of WDN Resilience

The network connectivity of a WDN can be modelled as a nearly-planar mathematical graph¹, $G = (V, E)$, where V (vertices) corresponds to n nodes and E (edges) corresponds to m pipes of the water system. This graph also has the particular characteristics that every node or edge should normally have at least one path of edges connecting it to a source node (tanks and/or reservoirs).

For a given WDN, our approach starts with computing a measure of resilience for each node by identifying the number of water sources a node is connected to and estimating how well that node is supplied by the available water sources. To accomplish this, all routes connecting a node i to all tanks and reservoirs are computed, and used to quantify the capacity or resistance of the corresponding routes. Since computing all routes which link

¹A planar graph is one that can be drawn on a plane without having edges intersecting other than at a node mutually incident with them. WDN graphs are not strictly planar due to the existence of some possible crossovers but are almost planar.

nodes to sources is prohibitively expensive, our approach becomes feasible by restricting the computations to the K ‘shortest’ routes (Eppstein 1998) for the flow between a node i and its j -th source $s_j(i)$. In calculating this index, every route should be weighted by a surrogate measure of flow resistance. For example, the proposed surrogate flow resistance measure depends on the physical characteristics of connecting pipes as water travels from j -th source, $s_j(i)$, to i and it can be estimated by a measure proportional to the energy loss associated with a route of supply. The estimation of energy loss in water transportation is commonly done using the Darcy-Weisbach and Hazen-Williams formulae. Although both formulae depend on hydraulic measures like velocity and roughness, only the proportional dependence on topological constants of the network has been considered as no detailed hydraulic modelling is used. Energy loss across a link is proportional to $f \times L/D$; where f is the friction factor, L is the length of the pipe and D represents its diameter. This resistance measure is used to weight every path in the calculus of shortest paths, providing a surrogate measure for potential energy loss associated with each path connecting a node and its water source. The proposed resilience index for a node is

$$I_{GT}(i) = \sum_{s=1}^S \left(\frac{1}{K} \sum_{k=1}^K \frac{1}{r(k, s)} \right), \quad (1)$$

where S is the total number of sources in the network and $r(k, s)$ is a surrogate measure of the energy loss (the resistance to water transport) associated with the k^{th} path to source s . One such measure of energy loss could be approximated as:

$$r(k) = \sum_{m=1}^M f(m) \frac{L_m}{D_m}, \quad (2)$$

where M is the number of pipes in path k and $f(\cdot)$ estimates the friction factor by pipe age and material (Christensen 2009).

The weighted KSP algorithm returns infinity for Eq. 2 if there is no path of water distribution connecting the given node with a supply source. The contribution of such a source to the resilience index for a node in Eq. 1 will be zero. Consequently, both the resistance of paths to sources and redundancy in supply are explicitly taken into account by the resilience index.

In Eq. 2, the flow resistance over K paths is averaged since using the shortest path only may not be sufficient to represent the connectivity and indirectly resilience. By averaging over the complete set of routes (i.e. $K \rightarrow 2^n$, where n is a very big number for operational networks (Eppstein 1998)), the true connectivity of a node to a source is measured. However, the computational complexity of this measure grows exponentially ($\mathcal{O}(2^n n)$ for the proposed index) and it quickly becomes infeasible for medium to large networks. Averaging over the K -shortest paths is a computationally feasible approximation that takes into account a sufficiently large number of paths in a robust statistical method.

2.1 Validation of the Resilience Assessment Method

A previously published network C-Town (Ostfeld et al. 2012) is used as a validation case-study for the graph-theoretic method proposed. This is a medium size network that consists of 333 nodes, 429 pipes, 4 valves, 5 pump stations, 7 elevated tanks, and 1 reservoir. A graph-theoretical analysis of resilience in WDNs (Yazdani et al. 2011) identifies a node of

special importance as those with high betweenness centrality index, which measures the relative number of shortest paths from all vertices to all others that pass through a node. Nodes with betweenness centrality values larger than 2 times the standard deviation (SD) from the network average computed over all nodes will be denominated as high betweenness nodes. This work considers “critical transfer nodes” to those of topological importance with respect to the flow distribution with a WDN. Thus, critical transfer nodes are, a priori, those belonging to the main trunk and those with high betweenness.

In the case of the proposed K -shortest path index, for a demand node i and source s , let

$$g(K; i, s) = \frac{1}{K} \sum_{k=1}^K \frac{1}{r(k)}, \tag{3}$$

represent the inverse of the resistance to supply in Eq. 1 averaged over K paths. Figure 1a shows the difference $g(K + 1) - g(K)$ for two nodes of varying centrality, A and B, where the latter belongs to the set of central nodes. In the two cases, the average of Eq. 3 taking K shortest paths converges when the number of paths is increased. That is, the difference on computing $g(K + 1)$ or $g(K)$ changes significantly as one moves away from using only the shortest path (i.e. $K = 1$) but not as much when one goes from K to $K + 1$ paths for larger values of K (over 30 paths in the case of Fig. 1a). Then, this measure is proposed as a decision criterion on the number of paths to be used in Eq. 1 to compute $I_{GT}(\cdot)$.

Using K paths has two main advantages. Firstly, it emphasises the need to consider multiple paths of transport that contribute to the resilient operation of a network and it guarantees a robust statistical estimation since multiple paths are used to capture the average connectivity and capacity instead of the unique shortest path connecting two nodes. Secondly, the sum in Eq. 1 converges for relatively low values of K to obtain a good approximation of the analysis over all possible routes.

Once the tuning process provides an estimate of the K value, the proposal proceeds to compute the inverse of the resistance to supply, $g(K)$, for the chosen K . In Fig. 1b $g(30)$ is plotted for all individual nodes. The same plot also represents the values of $g(30)$ for central nodes (main-trunk nodes and those with a betweenness centrality of at least 2 times the SD of the network average; they are plotted with red triangles). Most of the critical nodes have high $g(K)$ compared to the average node $g(K)$ value. In general, nodes of higher betweenness centrality are those that will appear more frequently in any shortest path, including paths to sources, and so will tend to have higher resilience as defined by Eq. 1.

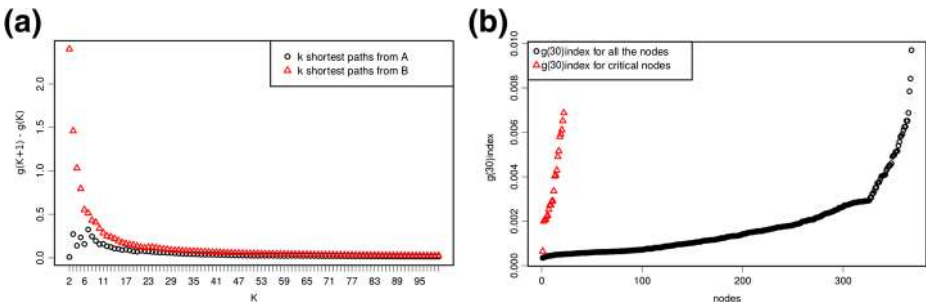


Fig. 1 Validation analysis on the K routes to the water sources

Table 1 Resilience indices computed when a percentage of pipes are removed

Index	0 % removed	2.5 % removed	5 % removed
I_r	0.4476	0.3697	0.2561
I_{GT}	0.6306	0.5711	0.4166
I_{CB}	0.0715	0.0620	0.0550
I_m	0.0597	0.0508	0.0419
I_S	0.0369	0.0355	0.0351

2.2 Comparison of K-Shortest Paths Index (I_{GT}) with Alternative Approaches

The next stage of the analysis is a comparison of the various indices of resilience under failure conditions in a supply network. Three scenarios generated from C-Town are considered, (Table 1, Fig. 1). The first scenario represents normal operational conditions without failures. The second scenario is created by the failure (removal) of 2.5 % of randomly selected pipes. The third scenario assumes the failure of 5 % of pipes. The resilience indices are expected to degrade as more pipes are taken off. The aim of this example is to evaluate the consistency of the new index, I_{GT} , by validating that it positively correlates with results obtained from hydraulic simulations. The comparison is done with Todini’s hydraulic resilience index, I_r (Todini 2000). Other graph-theoretic indices, such as central-dominance betweenness (I_{CB}), meshedness (I_m), and a demand-adjusted entropic index (I_S) (Yazdani and Jeffrey 2012), are also included in this analysis. Table 1 summarises the average of the computed indices under the repetition of 20 random experiments for each scenario.

To compare the performance of the various indices, their relative change under failure conditions are investigated. The process starts with the computation of a nominal resilience value; these are the computed under normal operation of a WDN. Then, there are computed the indices under failure configurations and are plotted the ratio of the new resilience values to their corresponding nominal index. Figure 2a shows that the degradation in resilience is captured by the considered indices to a different degree. The simulation of failure indices in Fig. 2a is repeated 20 times (by randomly sampling the corresponding disrupted pipes at each repeated simulation; in this case, 20 samples is enough to arrange statistically meaningful comparisons between groups/indices). At each iteration, different sets of pipes (2.5

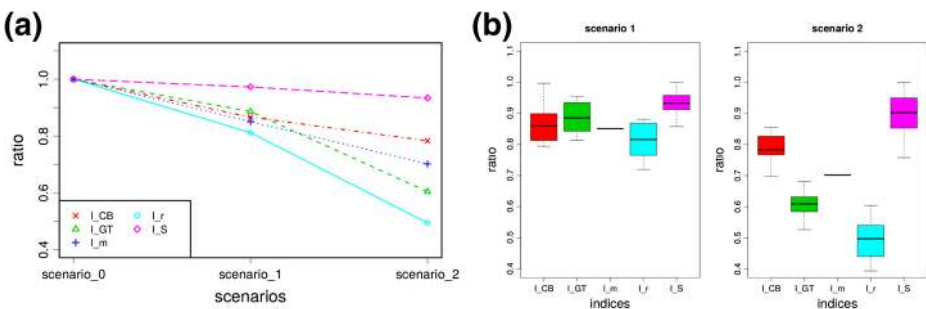


Fig. 2 Graph-theoretic indices vs. hydraulic simulation based index

% and 5 %) are randomly sampled and the indices are averaged for each proportion of pipe failures in the WDN. Figure 2b shows the distribution of the indices in Fig. 2a under these multiple experiments.

As shown in Fig. 2b, graph-theoretic indices show consistency with the hydraulic index². The index I_{GT} changes in a proportional way to variations in the condition of the water distribution network. The entropic degree, I_S , is the measure which provides the poorest estimation of the deterioration in resilience as more pipes are removed, similarly to the results presented by (Greco et al. 2012). Figure 2b also shows that the meshedness index, I_m , has a narrow distribution. This is because meshedness is a measure that depends only on the relative number of nodes and pipes, without taking into account their physical or hydraulic properties.

3 Assessing Resilience for Large Scale Water Networks

The network topology of a WDN includes both tree-like (usually hierarchical and lower redundant networks, i.e. having an equivalent proportion of nodes and links connecting them) and well meshed structures, where there usually exists several paths connecting two nodes and so a greater number of links compared to the number of nodes. These two type of structures can be interconnected by links forming graph-bridges, which transport water from the transmission mains to these distribution areas. One way to analyse large-scale WDNs is by partitioning the network into different components (assets: valves, pumps, pipes, and meters, among others). Alternatively, a WDN can be divided into sub-networks in order to make the analysis simpler than working with the whole system model. This is of interest especially in the case of large-scale WDNs, which can contain hundreds of thousands of nodes and pipes, where consequently the classical methods of analysis are not efficient because of the large dimensions.

Many WDNs have been divided into sectors (DMAs) mainly for leakage management (Tabesh et al. 2009; Gomes et al. 2013; Xin et al. 2014). For this and a variety of other reasons (Herrera 2011), WDNs are divided into sectors nowadays and researches have focused their attention on different sectorisation approaches. The most studied works on water network division are based on variations of graph clustering (di Nardo and di Natale 2011; Perelman and Ostfeld 2011), spectral clustering (Herrera et al. 2010; Candelieri et al. 2014; Herrera et al. 2012), community detection (Diao et al. 2012), multi-agent systems (Izquierdo et al. 2011; Hajebi et al. 2013), breadth and depth first search (Deuerlein 2008; di Nardo et al. 2014) or multilevel partitioning (di Nardo et al. 2013b). The work of Mandel et al. (2015) is the first to propose a sectorisation approach for large-scale water networks. For these previously sectorised networks, the resilience index in Eq. 1 measures the level of connectivity of each demand node through various supply roots. A DMA division of the network automatically allows the use of two different scales of this resilience measure, having approaches of resilience for every individual node or measures summarised by sector. This provides new ways for analysing and managing the system and allows network configurations that have a major impact on resilience to be easily identified; for example multi-feed, single-feed, cascading DMAs.

²Please note that the proposed index does not replace/estimate hydraulically simulated indices. The presented analysis validates the consistency of results as failure conditions change.

3.1 Multiscale Resilience Analysis

The proposed graph-theoretic index is extended to water networks that are divided into sectors thus enabling resilience analysis. A DMA-based graph is a graph of lower level of detail (or coarser granularity) that captures the topology of existing DMAs. The DMA-scale resilience is derived from a user-defined function that transforms the resilience of the demand nodes in a DMA to a scalar representing the DMA resilience. For example, by transforming the index of Eq. 1 to the coarser sector level graph to analyse the relative connectivity of DMAs. Here, it is used the trimmed mean (García-Escudero et al. 2003) to transform the resilience of n_i^* demand nodes of DMA_i to the sector graph as:

$$I_{GT}^*(DMA_i) = \sum_{j=1}^{n_i^*} \frac{I_{GT}(j)}{n_i^*}. \tag{4}$$

In calculating the mean, the trimmed approach discards points of very high or very low value before computing the mean for the remaining nodes. Then, the aggregated measure is the summary of the single values for n_i^* nodes of the DMA, with $n_i^* \leq n_i$. As such, a trimmed mean is a more robust centrality measure as the average value is not skewed by extreme low/high value data. For instance, if a DMA with predominantly high resilient nodes has a sub-area that has critically low resilience, the arithmetic average would not identify the presence of these in the DMA resilience index, whereas the trimmed method does. Furthermore, the presence of individual critical customers such as hospitals could significantly affect the system level resilience assessment. The set of excluded nodes in the computation of the trimmed mean, \mathcal{H} , are identified to validate the existence of critical nodes among them, and the subsequent impact on the DMA resilience assessment. As a result, Eq. 5 formulates a constitutive rule which is used to rank DMAs as low resilient if they have low values of I_{GT} for critical users (hospitals, schools, or industrial users) regardless of the fact that a DMA as a whole or a majority of its demand nodes might have high resilience.

$$I_{GT}(DMA_i) = \begin{cases} \min\{I_{GT}(H_i)\} & \text{if at least one critical user } \in \mathcal{H} \\ I_{GT}^*(DMA_i) & \text{otherwise,} \end{cases} \tag{5}$$

where H_i is the set of critical users in DMA_i .

For a DMA with a large variation in the I_{GT} value, it is possible to trim a significant number of nodes (i.e. a number of demand nodes representing a significant number of customers) in computing Eq. 4. To avoid this, the variability of the resilience is used, which is defined here as standard deviation normalized by the mean resilience value:

$$S_{GT}^2(DMA_i) = \sum_{j=1}^{n_i^*} \frac{1}{n_i^* - 1} \left(\frac{I_{GT}(j) - I_{GT}^*}{I_{GT}^*} \right)^2. \tag{6}$$

Based on the variability of $I_{GT}(\cdot)$ in each DMA_i (Eq. 6), a second constitutive rule is formulated and applied:

$$I_{GT}(DMA_i) = \begin{cases} \min_{DMA_i}\{I_{GT}(\cdot)\} & \text{if } S_{GT}^2(DMA_i) > t \\ I_{GT}^*(DMA_i) & \text{otherwise} \end{cases} \tag{7}$$

where t is a user-defined threshold.

When high variability exists, this constitutive rule computes the DMA resilience index over the whole set of demand nodes in the DMA and identifies the existence of low resilient nodes. The two Eqs. 5 and 7 assist operators to identify areas for a further investigation where a threshold of demand nodes or critical customers might not have the required level of resilience.

3.2 Multiscale Representation of a WDN

An essential part of the sector-level resilience analysis is the pre-processing of network data of varying types and complexity. For example, geographic information system (GIS) data, telemetry measurements, and hydraulic information is aggregated into an unified system that provides access to heterogeneous data sets. A novel augmented graph is proposed which transforms DMAs into sector-nodes. Each sector-node has aggregated data about the assets it contains, abstracting the high-granularity graph of hydraulic nodes to a low-granularity graph. The pipes and valves that connect different DMAs represent the edges between sector-nodes and inherit their corresponding properties. A schematic view of this multiscale decomposition process is represented in Fig. 3.

By considering all links hydraulically connecting DMAs, the edges of the new graph are formed by taking into account the open/closed status of valves, and consequently links which can provide additional operational resilience are identified. The weights of the new graph edges are proportional to the aggregated volumetric properties of individual pipes and valves that connect sector-nodes. Sector-nodes can also be used to visualize average levels of demand, pressure, elevation, etc. and the sector-node size can inherit topological characteristics of DMAs (eg. number of nodes or customers).

A multiscale network decomposition to assess the resilience of water systems is approached by studying their performance and network structure at various granularity levels and by analysing topological and physical information inherited from single demand nodes to sectors. Other important details that are obtained in the analysis include the number and location of multi-feed DMAs, the number of valves within each DMA, and the pressure differential between neighbouring DMAs. In addition, the sector-node graph is used to identify which part of the network topology should require further investigation to improve resilience.

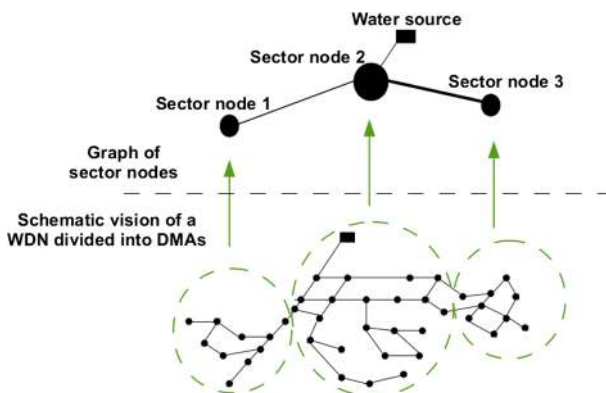


Fig. 3 Mapping high-granularity demand-node into low-granularity DMA-based graphs

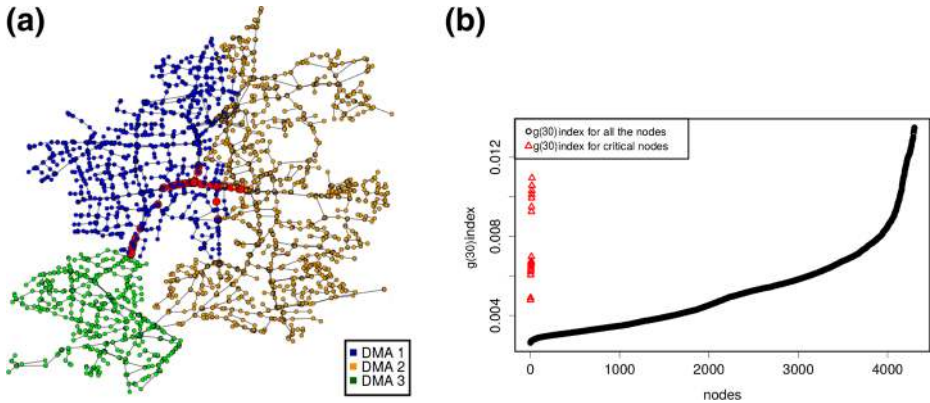


Fig. 4 Layout and $g(K)$ index for the analysis of WDN-A

4 Experimental Study

Two operational WDNs were used to analyse the proposed methodology for assessing the resilience of sectorised water distribution networks. Firstly, a medium size WDN divided into 3 DMAs is investigated. A resilience assessment is carried out for each DMA. The second case-study corresponds to a large water supply zone with high population density in England, UK. In both cases, certain sensitive information about the networks is withheld in order to maintain confidentiality.

4.1 Case-Study A

WDN-A (Fig. 4a) has 4820 nodes, 1 tank, and 5234 pipes and valves sectorised into 3 DMAs. Detailed information for each DMA is summarised in Table 2. Figure 4a also shows the DMAs layout and the location of critical transfer nodes (i.e. nodes with high values of central-betweenness).

In order to compute the proposed graph-theoretic index, I_{GT} of Eq. 1, a sufficient number of shortest paths K need to be used. The analysis outlined in Fig. 1 was carried out on different nodes and it showed that $K = 30$ as sufficient for WDN-A. In Fig. 4b, we plot the values of $g(30)$ for each node of the WDN and for demand nodes with high betweenness centrality (labelled as critical nodes and also marked in red in Fig. 4a). For this network, analysis in Fig. 4b shows that the measure $g(K)$ of the most critical nodes exceeds the average for the network.

From Table 2, it can be concluded that DMA-3 is the sector with lowest resilience given the index $I_{GT}(DMA_3)$, while $I_{GT}(DMA_2)$ index is higher than the other two. DMA-2 is

Table 2 Summary of the DMA configuration and resilience indices for WDN-A

Sector ID	# nodes	# pipes	$I_{GT}(DMA)$
DMA-1	1626	1770	0.0654
DMA-2	2416	2638	0.0869
DMA-3	778	826	0.0465
TOTAL	4820	4032	0.0663

directly connected with the only source that supplies water to the whole network. Thus, topologically and hydraulically, DMA-2 is the most resilient sector of this network.

4.2 Case-Study B

WDN-B (shown in Fig. 5) is comprised of 106,115 nodes, 4 constant head sources, 28 reservoirs, 111,096 pipes, 102 pumps, and 1,900 boundary valves (Fig. 5a). GIS models, telemetry data, and physical and hydraulic information are combined in an integral database framework. Figure 5 shows the process of multiscale decomposition from the network of demand nodes to a coarser representation by sector-nodes. Through the graph transformation process, every DMA (sector-node) inherits the information related to the nodes within the area; such as, demand, elevation, pressure, and geographic coordinates. Sector-nodes also gather the information of the pipes and valve: diameter and length, and their own characteristics and the open/closed status for valves.

The proposed multiscale decomposition of the network topology by sector-nodes reduces the dimensionality of complex networks. Figure 5b shows the spatial distribution of these sector-nodes together with their connectivity in terms pipes and valves. The volumetric capacity of aggregated links between DMAs are represented by the thickness of the connecting links. The colour code of the sector-nodes identifies whether a DMA is multi-feed (in colour red) or single feed (grey colour). The analysis shows that almost 40 % of the DMAs (i.e. 90 of the 228 DMAs) in WDN-B are fed by more than one supply connection, Fig. 5b. All this information together with the topology of the assets in the DMAs facilitate the estimation of the resilience indices at a DMA level, which is a new property of individual sector-nodes. Any measure based on individual assets within each DMA can be represented in the multiscale DMA-graph. Figure 6a shows the average resilience for each DMA, $I_{GT}(DMA_i)$, using a traffic light system. Red represents low resilience with an I_{GT} value 2 SDs below the network average for all DMAs; in this example, the network average is 0.0395 and 2 SDs below that is 0.02. Similarly, high resilience sectors (with values 2 SDs above average DMA resilience) are represented using green colour, and DMAs in

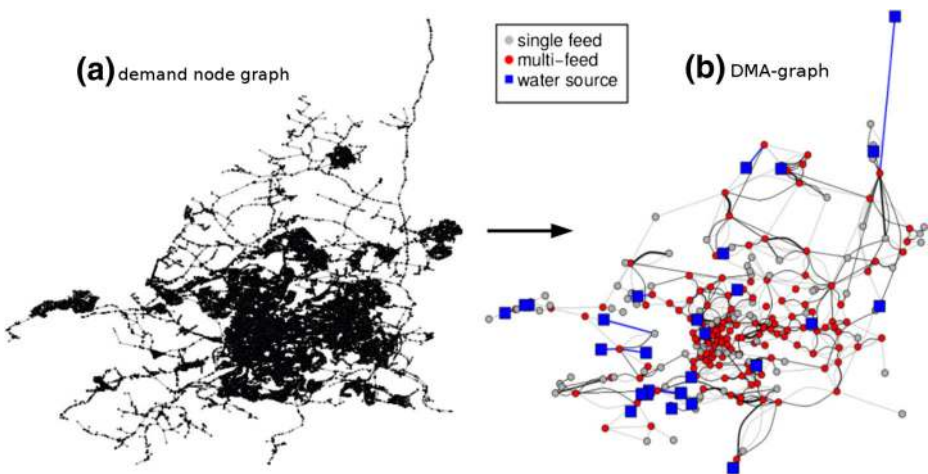


Fig. 5 Layout of WDN-B and its representation by a graph of sector-nodes. Single-feed and multi-feed DMAs are differentiated visually.

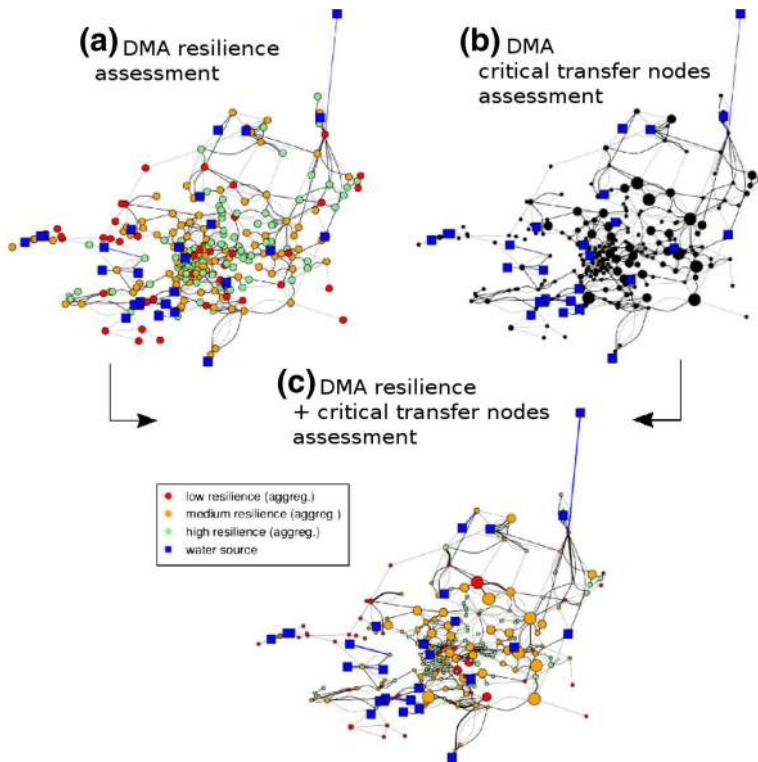


Fig. 6 Schematic layout of WDN-B network resilience: $I_{GT}(DMA_i)$ given in Eq. (5) is represented by colour code (low=red, high=green) and the level of critical transfer by node size.

between the two extremes with orange/amber colour. The augmented DMA-graph can also provide further information which assists operators in understating and analysing topological and operational constraints that affect resilience. Figure 6b highlights the critical transfer nodes of every DMA in distribution using the size of the circle plot associated with sector-nodes. In this figure, criticality is determined based on the number of high betweenness nodes.

Resilience and critical transfer measures are simultaneously represented in Fig. 6(c), which combines a traffic light system and the sector-node size. As a result, the centre and centre-North of WDN-B include DMAs that have medium and low resilience measures and medium and high levels of flow transferability for central nodes (see Table 3). All these areas also supply critical customers and could be prioritised for further analysis and maintenance to improve network resilience. The boundaries of the WDN layout have low resilience but also low level of criticality. Table 3 presents a summary of attributes for the top 5 sectors in terms of their resilience and flow transfer critical nodes.

As shown in Table 3, where the DMAs are ranked by resilience, 3 of the 5 DMAs contain more than 2,000 nodes, while only one of them has a size less than 1,000 nodes. Taking into account the number of critical transfer nodes within each sector (column #CNs of Table 3), it is shown that DMA-29, DMA-101, and DMA-208 are sectors with larger number of transfer

Table 3 Summary of the DMA configuration of the top 5 WDN-B sectors

Sector ID	# nodes	# pipes	Multi-feed	$I_{GT}(DMA)$	#CNs
DMA-29	2003	2081	Yes	0.0124	118
DMA-184	1119	1242	No	0.0179	32
DMA-85	784	834	Yes	0.0194	25
DMA-101	2106	2181	No	0.0212	60
DMA-208	2081	2249	Yes	0.0268	114

The resilience index, I_{GT} , is computed for each DMA. The number of critical transfer nodes (i.e. the number of nodes with betweenness centrality values larger than 2 times the SD from the average computed over all the DMA nodes) is given by #CNs

nodes. These DMAs also have a varying level of resilience; for example DMA-29 has the smallest resilience index and DMA-208 has the highest.

5 Conclusions

A novel resilience index based on weighted K -level shortest paths from consumption nodes to water sources has been derived to estimate WDN resilience together with information about critical customers. Computing these K different routes for every node provides numerical robustness to our analysis compared to analysing only the shortest path. The index is shown to be consistent with hydraulically-defined resilience measures. This is because the K -shortest paths are weighted by quantities associated with energy losses of water transport routes. The current proposal has the following advantages:

- An inclusion of physical attributes of the water distribution network within graph theoretic indices gives the resilience analysis a hydraulically relevant and easy to quantify measure of how well connected a node is to the available water sources.
- Computations for the proposed index scale quasi-linearly as a function of the number of nodes ($\sim \mathcal{O}(m + n \log(n) + kn)$). This makes the measure feasible for application to large scale water networks, where the classical critical link hydraulic based analysis of network resilience would scale combinatorially.
- Constitutive rules and functions facilitate the inclusion of empirical knowledge.

The proposed graph-theoretic framework is complimented by a novel multiscale decomposition method that converts the original WDN layout to DMA-based graphs. The developed method allows us to work with a WDN at two different levels of abstraction: single demand node and DMA. This enhances the results obtained for assessing the resilience of a WDN, especially in the case of large scale networks.

Acknowledgments This work has been performed with the support of the NEC - Imperial research project on Big Data Technologies for Smart Water Networks.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Abraham E, Stoianov I (2015) Sparse null space algorithms for hydraulic analysis of large scale water supply networks. *J Hydraul Eng*:04015058
- Albert R, Barabási AL (2002) Statistical mechanics of complex networks. *Rev Mod Phys* 74(1):47
- Baños R, Reca J, Martínez J, Gil C, Márquez AL (2011) Resilience indexes for water distribution network design: a performance analysis under demand uncertainty. *Water Resour Manag* 25(10):2351–2366
- Berardi L, Ugarelli R, Røstum J, Giustolisi O (2014) Assessing mechanical vulnerability in water distribution networks under multiple failures. *Water Resour Res* 50(3):2586–2599
- Buhl J, Gautrais J, Reeves N, Solé RV, Valverde S, Kuntz P, Theraulaz G (2006) Topological patterns in street networks of self-organized urban settlements. *The European Physical Journal B - Condensed Matter and Complex Systems* 49(4):513–522
- Candelieri A, Conti D, Archetti F (2014) Improving analytics in urban water management: a spectral clustering-based approach for leakage localization. *Procedia-Social and Behavioral Sciences* 108:235–248
- Christensen RT (2009) Age effects on iron-based pipes in water distribution systems. Utha State University, PhD thesis
- Deuerlein JW (2008) Decomposition model of a general water supply network graph. *J Hydraul Eng* 134(6):822–832
- di Nardo A, di Natale M (2011) A heuristic design support methodology based on graph theory for district metering of water supply networks. *Eng Optim* 43(2):193–211
- di Nardo A, di Natale M, Santonastaso GF, Tzatchkov VG, Alcocer-Yamanaka VH (2013a) Water network sectorization based on a genetic algorithm and minimum dissipated power paths. *Water Science & Technology: Water Supply* 13(4):951–957
- di Nardo A, di Natale M, Santonastaso GF, Venticinque S (2013b) An automated tool for smart water network partitioning. *Water Resour Manag* 27(13):4493–4508
- di Nardo A, di Natale M, Santonastaso G, Tzatchkov V, Alcocer-Yamanaka V (2014) Water network sectorization based on graph theory and energy performance indices. *J Water Resour Plan Manag* 140(5):620–629
- Diao K, Zhou Y, Rauch W (2012) Automated creation of district metered area boundaries in water distribution systems. *J Water Resour Plan Manag* 139(2):184–190
- Eppstein D (1998) Finding the k shortest paths. *SIAM J Comput* 28(2):652–673
- Freeman LC (1977) A set of measures of centrality based on betweenness. *Sociometry* 40(1):35–41
- García-Escudero LA, Gordaliza A, Matrán C (2003) Trimming tools in exploratory data analysis. *J Comput Graph Stat* 12(2)
- Giustolisi O, Kapelan Z, Savic D (2008) Algorithm for automatic detection of topological changes in water distribution networks. *J Hydraul Eng* 134(4):435–446
- Gomes R, Marques AS, Sousa J (2013) District metered areas design under different decision makers options: Cost analysis. *Water Resour Manag* 27(13):4527–4543
- Greco R, di Nardo A, Santonastaso G (2012) Resilience and entropy as indices of robustness of water distribution networks. *J Hydroinformatics* 14(3):761–771
- Gupta R, Bhave PR (1994) Reliability analysis of water-distribution systems. *J Environ Eng* 120(2):447–461
- Gutiérrez-Pérez J, Herrera M, Pérez-García R, Ramos-Martínez E (2013) Application of graph-spectral methods in the vulnerability assessment of water supply networks. *Math Comput Model* 57(7–8):1853–1859
- Hajebi S, Barrett S, Clarke A, Clarke S (2013) Multi-agent simulation to support water distribution network partitioning. In: 27th European simulation and modelling conference - ESM'2013. Lancaster University, UK, pp 1–6
- Herrera M (2011) Improving water network management by efficient division into supply clusters. PhD thesis, Hydraulic Eng. and Environmental Studies, Universitat Politècnica de Valencia
- Herrera M, Canu S, Karatzoglou A, Pérez-García R, Izquierdo J (2010). In: Swayne DA, Yang W, Voinov A, Rizzoli A, Filatova T (eds) An approach to water supply clusters by semi-supervised learning
- Herrera M, Izquierdo J, Pérez-García R, Montalvo I (2012) Multi-agent adaptive boosting on semi-supervised water supply clusters. *Adv Eng Softw* 50:131–136
- Izquierdo J, Herrera M, Montalvo I, Pérez-García R, Cordeiro J, Ranchordas A, Shishkov B (2011) Division of Water Supply Systems into District Metered Areas Using a Multi-agent Based Approach. Software and data technologies, communications in computer and information science, vol 50, Springer Berlin Heidelberg, pp 167–180

- Jayaram N, Srinivasan K (2008) Performance-based optimal design and rehabilitation of water distribution networks using life cycle costing. *Water Resour Res* 44(1):1–15
- (2005) Resilience for the scalability of dependability. In: Laprie JC (ed) Fourth IEEE International Symposium on Network Computing and Applications, pp 5–6
- Lee JD, Maggioni M (2011) Multiscale analysis of time series of graphs. In: International Conference on Sampling Theory and Applications (SampTA)
- Liu H, Savić D, Kapelan Z, Zhao M, Yuan Y, Zhao H (2014) A diameter-sensitive flow entropy method for reliability consideration in water distribution system design. *Water Resour Res* 50(7):5597–5610
- Mandel P, Maurel M, Chenu D (2015) Better understanding of water quality evolution in water distribution networks using data clustering. *Water Res* 87(15):69–78
- Ostfeld A, Salomons E, Ormsbee L, Uber J, Bros C, Kalungi P, Burd R, Zazula-Coetzee B, Belrain T, Kang D, Lansley K, Shen H, McBean E, Yi Wu Z, Walski T, Alvisi S, Franchini M, Johnson J, Ghimire S, Barkdoll B, Koppel T, Vassiljev A, Kim J, Chung G, Yoo D, Diaio K, Zhou Y, Li J, Liu Z, Chang K, Gao J, Qu S, Yuan Y, Prasad T, Laucelli D, Vamvakeridou Lyroudia L, Kapelan Z, Savić D, Berardi L, Barbaro G, Giustolisi O, Asadzadeh M, Tolson B, McKillop R (2012) Battle of the water calibration networks. *J Water Resour Plan Manag* 138(5):523–532
- Perelman L, Ostfeld A (2011) Topological clustering for water distribution systems analysis. *Environ. Model Softw* 26(7):969–972
- Prasad T, Park N (2004) Multiobjective genetic algorithms for design of water distribution networks. *J Water Resour Plan Manag* 130(1):73–82
- Raad DN, Sinske AN, van Vuuren JH (2010) Comparison of four reliability surrogate measures for water distribution systems design. *Water Resour Res* 46(5):1–11
- Strigini L (2012) Fault tolerance and resilience: meanings, measures and assessment. In: Wolter K, Avritzer A, Vieira M, van Moorsel A (eds) *Resilience Assessment and Evaluation of Computing Systems*. Springer, Berlin, pp 3–24
- Tabesh M, Yekta AA, Burrows R (2009) An integrated model to evaluate losses in water distribution systems. *Water Resour Manag* 23(3):477–492
- Takahashi S, Saldarriaga J, Vega M, Hernández F (2010) Water distribution system model calibration under uncertainty environments. *Water Science & Technology: Water Supply*
- Todini E (2000) Looped water distribution networks design using a resilience index based heuristic approach. *Urban Water* 2(2):115–122
- Trifunović N (2012) Pattern recognition for reliability assessment of water distribution networks. PhD thesis, UNESCO-IHE Institute for Water Education, Delft University of Technology
- Wright R, Abraham E, Pappas P, Stoianov I (2015) Control of water distribution networks with dynamic DMA topology using strictly feasible sequential convex programming. *Water Resources Research* 51(12):9925–9941. doi:[10.1002/2015WR017466](https://doi.org/10.1002/2015WR017466)
- Xin K, Tao T, Lu Y, Xiong X, Li F (2014) Apparent losses analysis in district metered areas of water distribution systems. *Water Resour Manag* 28(3):683–696
- Yazdani A, Jeffrey P (2012) Water distribution system vulnerability analysis using weighted and directed network models. *Water Resour Res* 48(6):W06,517
- Yazdani A, Otoo RA, Jeffrey P (2011) Resilience enhancing expansion strategies for water distribution systems: a network theory approach. *Environ Model Softw* 26(12):1574–1582