Clustering analysis of water distribution systems: identifying critical components and community impacts

K. Diao, R. Farmani, G. Fu, M. Astaraie-Imani, S. Ward and D. Butler

ABSTRACT

Large water distribution systems (WDSs) are networks with both topological and behavioural complexity. Thereby, it is usually difficult to identify the key features of the properties of the system, and subsequently all the critical components within the system for a given purpose of design or control. One way is, however, to more explicitly visualize the network structure and interactions between components by dividing a WDS into a number of clusters (subsystems). Accordingly, this paper introduces a clustering strategy that decomposes WDSs into clusters with stronger internal connections than external connections. The detected cluster layout is very similar to the community structure of the served urban area. As WDSs may expand along with urban development in a community-by-community manner, the correspondingly formed distribution clusters may reveal some crucial configurations of WDSs. For verification, the method is applied to identify all the critical links during firefighting for the vulnerability analysis of a real-world WDS. Moreover, both the most critical pipes and clusters are addressed, given the consequences of pipe failure. Compared with the enumeration method, the method used in this study identifies the same group of the most critical components, and provides similar criticality prioritizations of them in a more computationally efficient time.

K. Diao (corresponding author)  
R. Farmani  
G. Fu  
M. Astaraie-Imani  
S. Ward  
D. Butler  
Centre for Water Systems, University of Exeter, Exeter EX6 7HS, UK  
E-mail: K.Diao@exeter.ac.uk

Key words | clustering analysis, critical component, urban development, water distribution system

INTRODUCTION

A water distribution system (WDS) is a network configured with hydraulic components that provide a water supply. Specifically, the vertices in the network could be the consumers, sources, and tanks; while the edges could be the connecting pipes, pumps, and valves (Perelman & Ostfeld 2014). Along with the urbanization process, WDSs, acting as urban lifelines, have become enormous both in size and complexity, which may create tremendous economic loss and adverse social impacts (Farmani & Butler 2014). For instance, during the summer of 2009, 101 pipe breaks occurred in Los Angeles; these breaks were followed by an investigation as a consequence of the public request for an explanation on the cause of these blowouts (Bardet et al. 2010). Again, the water loss in a WDS could exceed 50% either in developing or developed countries (Society for International Development 2008; Giugni et al. 2008). There are both topological and behavioural complexities in WDSs. For instance, real-world WDSs commonly have a vast number of nodes and links organized in intricate topologies (e.g., a combination of loops and branches). Again, the WDSs are nonlinear dynamic systems. There are considerable uncertainties in system behaviours, such as the stochasticity of flow rates, pressures, and water demands, etc. Admittedly, nowadays, most of the details of a WDS could be mimicked by a hydraulic model due to developments in both modelling techniques and data availability. Nevertheless, setting up such a model is just the prerequisite for understanding a WDS. As there are a large number of components with a variety of properties, it is usually rather difficult to identify the system’s properties (e.g., the network structure, the operational scheme, and interactions between its components, etc.). In search of solutions to this problem, one way is to simplify the network layout for visualization and data processing through clustering analysis. Clustering of networks consists of dividing a network into a number of sub-networks (i.e., clusters) with vertices and edges (Schaeffer 2007). The resulting layout of clusters reveals the network structure and interactions between components more explicitly.

In the context of WDS analysis, there are a few studies exploring the application of clustering approaches. Bartolin
et al. (2005) applied the minimum spanning tree method to assist several model-based analyses of the Valencia WDS. The analyses involved network segmentation, detection of isolated parts of the network, isolation of broken pipes, and identification of the conveyance capacity of critical flow paths (removal of which divides the network into two separate parts). Similarly, Tzatchkov (2006) applied depth-first and breadth-first based graph algorithms to the segmentation and water quality analysis of a large WDS. Specifically, as for the segmentation, the method proposed is able to identify independent networks and to detect errors in network data. As for water quality analysis, it speeds up the process of obtaining the contribution of multiple water sources to the consumption at nodes by mapping WDSs into directed graphs and making use of stacks for computation. Deuerlein (2008a) developed a generalized graph decomposition model that simplifies a network into a graph consisting of two main elements, called forests and cores, respectively. The model was subsequently applied to facilitate WDS analysis including, reliability analysis (Deuerlein et al. 2009), model calibration (Deuerlein 2008b), and the risk analysis of sensor placement (Deuerlein et al. 2011). Perelman & Ostfeld (2011) developed topological clustering tools for analysis of flow patterns in WDSs. As clusters result from the flow directions in pipes, the placement of sensors can be proposed. Yazdani & Jeffrey (2011) discussed the suitability of the clustering coefficient in graph theory as an indicator of path redundancy in WDSs. Sempewo et al. (2008) and Di Nardo & Natale (2010) presented spatial analysis tools for metered areas design in WDSs. Similarly, Izquierdo et al. (2009) and Herrera et al. (2010) designed two partitioning methods based on machine learning, with both graphical and vector information considered. More recently, Scibetta et al. (2013) and Diao et al. (2015) proposed complex network clustering techniques for the same task.

The research introduced above has indeed provided substantially useful insights regarding WDS clustering. However, there is still no definition describing what could be a cluster in a WDS. Since urban development is one origin of the complexities of WDSs, there may exist cluster structures in WDSs formed along with urban expansion. As WDSs are expanding along with urban development and might therefore be developed community-by-community, clusters in a WDS could be sub-networks serving corresponding communities (e.g., residential zones, industrial sites, business centres, etc.) in the covered urban area. Detection of such a cluster structure in the WDS can therefore provide crucial clues for understanding the properties of the system, and for the evaluation of the social-economic impacts of water supply from a cluster-based point of view.

Accordingly, this paper proposes a clustering strategy that intends to decompose a WDS into the cluster structure resulting from urban development. Being identical to the idea of decomposition theorems (Simon & Ando 1961; Simon 1962; Courtois 1985; Salingaros 2001), the method, termed modularity-based clustering (Clauset et al. 2004; Newman 2004; Newman & Girvan 2004), divides a complex system into a number of clusters with stronger internal connections than external connections (Salingaros 2001), as shown in Figure 1. The quality of the division could be measured using ‘modularity’ ($0 < Q < 1$) as a metric. A high modularity ($Q > 0.3$) indicates a significant cluster structure.

The suitability of the method confirms the fact that the topological features of WDSs fit into the decomposition theorems very well. WDSs are expanding along with urban evolution and may therefore be developed community by community. The urbanization starts from initial building blocks. These entities subsequently expand or are combined with each other to form larger blocks step by step. During this process, water distribution pipes are organized following the development of such city blocks (communities), and therefore the distribution systems are also formed block-wise. As pipe redundancy is required for water security considerations (Awumah et al. 1991; Walski et al. 2003), the number of intra-block connections is usually higher than that of inter-block connections, since the former is to ensure service while the latter is mainly to deliver water among regions. Furthermore, at geographical boundaries (e.g., rivers and elevation boundaries), the density of pipes crossing those boundaries is commonly low for the sake of

![Figure 1](https://iwaponline.com/wst/article-pdf/70/11/1764/175008/1764.pdf)
cost saving. Thus, the blocks could be regarded as clusters, given the higher fraction of intra-block connections. Each cluster is a subnetwork mainly serving a city block. Hence, the layout of those clusters is a good indicator of the WDS’s structure. For instance, an aggregated visualization of a WDS could be generated. On the basis of the visualization, it is should be possible to identify all the crucial characteristics of a WDS, particularly all the most critical clusters, nodes, and links for a given concern. In this study, such a practical application is discussed in detail based on a real-world case study.

## METHODS AND MATERIALS

### Modularity-based clustering

There are two steps for modularity-based clustering of WDSs.

#### Step 1: WDS mapping

The distribution system is mapped into an undirected graph $G = (V, E)$ in which the vertices $V$ represent the consumers, sources, and tanks; the edges $E$ the connecting pipes, pumps, and valves (Perelman & Ostfeld 2011).

#### Step 2: clustering

The cluster layout of $G$ is identified by using an algorithm proposed by Clauset et al. (2004). The method uses ‘modularity’ ($Q$) as an indicator to quantify the quality of the graph division into clusters (Novák et al. 2010). By mapping a variety of networks onto undirected graphs and applying clustering analysis, researchers in complex systems science have found that networks with a strong cluster structure typically have a value of $Q \geq 0.3$ (Clauset et al. 2004; Newman & Girvan 2004). The strong cluster structure refers to a structure consisting of clusters with dense edges within them, but with sparse connections between them. The modularity is defined as:

$$Q = \frac{1}{2m} \sum_{(i,j)} \left( A_{ij} - \frac{k_i k_j}{2m} \right) \delta(c_i, c_j) \quad (1)$$

where $A_{ij}$ is an element of the adjacency matrix of the network ($A_{ij} = 1$ if vertices $v_i$ and $v_j$ are connected, and $A_{ij} = 0$ otherwise). $m = \frac{1}{2} \sum_{(i,j)} A_{ij}$ is the total number of edges; $k_i = \sum_j A_{ij}$ is the degree of vertex $v_i$, defined as the number of edges connected to that vertex; $\delta(c_i, c_j)$ is 1 if $c_i = c_j$, and 0 otherwise. $c_i$ and $c_j$ represent two different clusters; $v_i$ and $v_j$ represent vertices in $c_i$ and $c_j$, respectively.

$k_i k_j / 2m$ is the probability of an edge existing between vertices $v_i$ and $v_j$ if connections are randomly made but respecting vertex degrees. Hence, Equation (1) could be regarded as a scoring function that compares the actual fraction of intra-cluster edges with its expectation in the randomized network (Schuetz & Caflisch 2008). If the number of intra-cluster edges is no better than random then $Q = 0$, and conversely if there are no inter-cluster edges then $Q = 1$ (Zhu et al. 2008).

To facilitate the algorithm description, writing $\delta(c_i, c_j) = \sum_{(i,j)} \delta(c_i, i) \delta(c_j, j)$ and defining $e_{ij} = 1/2m \sum_{(i,j)} A_{ij} \delta(c_i, i) \delta(c_j, j)$ and $a_i = 1/2m \sum_k k_i \delta(c_i, k) = \sum_j e_{ij}$, Equation (1) is then written into:

$$Q = \sum_i \left[ e_{ii} - a_i \right]^2 \quad (2)$$

with $e_{ii}$ the fraction of the edges whose ends both belong to the cluster $i$, $e_{ij}$ the fraction of edges that link vertices in cluster $i$ to those in cluster $j$, and $a_i$ the fraction of edges that have at least one end in cluster $i$.

Merging two clusters $i$ and $j$ increases the modularity by (Newman 2004):

$$\Delta Q_{ij} = 2(e_{ij} - a_i a_j) \quad (3)$$

An example of how Equation (3) is deduced is included in Zhu et al. 2008.

Based on Equation (3), the modularity-based algorithm starts with all vertices being the sole member of a cluster of one, in which case $e_{ij} = 1/2m$ if $i$ and $j$ are connected and zero otherwise, and $a_i = k_i / 2m$. Thus, the initial values of $\Delta Q_{ij}$ and $a_i$ are:

$$\Delta Q_{ij} = \begin{cases} 1/2m - k_i k_j / (2m)^2 & \text{if } i, j \text{ are connected,} \\ 0 & \text{otherwise,} \end{cases} \quad (4)$$

and

$$a_i = k_i / 2m \quad (5)$$

for each $i$. Following a greedy strategy, the method then successively joins clusters with the greatest increase $\Delta Q$ in modularity; and eventually stops when the maximum
possible \( \Delta Q \leq 0 \) from joining any two is reached. A flowchart of the algorithm is provided below.

Flowchart of the community detection algorithm:
Each vertex is a community
Calculate the modularity change matrix \( \Delta Q \)
Determine \( a_i \)
while pair \((i, j)\) with \( \Delta Q_{ij} > 0 \) exist do
  Find the maximum element along its index \( (i', j') \) in each row of \( \Delta Q \);
  Search the maximum value \( \Delta Q_{i'j'} \) in all \( \{ (i, j, \Delta Q_{ij}) \}^n \) and its index \( (i', j') \);
  Eliminate the \( i' \)th row and column from \( \Delta Q \);
  Merging community \( i' \) into \( j' \) through updating the \( j' \)th row and column in \( \Delta Q \);
switch
case 1 \( k \) is connected to both \( i' \) and \( j' \)
  \( \Delta Q_{ij} = \Delta Q_{i'j'} + \Delta Q_{i'j'k} \)
case 2 \( k \) is connected to \( i' \) but not to \( j' \)
  \( \Delta Q_{ij} = \Delta Q_{i'j'k} - 2a_ia_k \)
case 3 \( k \) is connected to \( j' \) but not to \( i' \)
  \( \Delta Q_{ij} = \Delta Q_{i'j'k} - 2a_ia_k \)
end switch
end while

After decomposition, the WDS layout is simplified to a graph composed of clusters and interconnections between clusters. The properties of the WDS could then be explored by, for instance, addressing crucial interactions between clusters, identifying critical clusters, etc.

Application of the method

A dialectical example – the New York Tunnels

The New York Tunnels is a benchmark water supply system for an optimum design problem. Figure 4(a), provided by Savic (2002), illustrates the system layout and the territory where it is located. As can be seen, the city is comprised of five river-divided regions, which are Bronx, Manhattan, Queens, Brooklyn, and Richmond, respectively. Correspondingly, the water supply clusters constructed following urban development should fit into the terrain of the city; that is, the tunnel network could be regarded as a system naturally consisting of five interrelated clusters as a result of the terrain feature. Therefore, this typical case is an ideal example for testing the competence of the clustering method, since the approach is expected to detect the cluster structures of WDSs formed along with urban development.

Case study on an alpine WDS

To further verify the practical value of the proposed method, the clustering approach is applied to facilitate vulnerability analysis (VA) of an alpine WDS for 3,000 inhabitants with two pressure zones (Figure 3(a)). The hydraulic model of the system contains 127 junctions, 157 pipes, two reservoirs, four tanks, and six pumps (in three pump stations). The simulation period is 75 hours. Here, vulnerability is defined as the consequences to the recipient of system failure or a reduced level of service (Butler et al. 2014). In this study, the VA refers to identification of the most critical pipes in the system according to evaluation of failure impacts. Specifically, the events considered are firefighting and single pipe failure. The two events are assumed to appear simultaneously in all failure scenarios. For example, a failure of a pipe is combined with an additional firefighting water demand (26.67 l/s according to Austrian standards) at a junction. The consequence to the recipient (failure impact) is assessed by the hourly averaged change in the total number of population served due to a single pipe failure, i.e., \( \text{Pop} \). Calculation of \( \text{Pop} \) is described by Equations (6)–(11). Of course, reducing vulnerability to the relationship of affected population versus connecting pipes is a simplification. The method could be extended to achieve more in-depth risk analysis by using either a more comprehensive indicator taking additional information into account (e.g., failure probability of different infrastructures, degree of redundancy of supply), or a group of indicators emphasizing different aspects. In the latter case, the prioritization of crucial components based on each indicator could be addressed to identify components that are crucial for all indicators. To investigate all failure scenarios systematically (the total number of failure scenarios is \( 127 \times 157 = 19,939 \)), the node with fire-fighting demand is first located. Next, pipes are sequentially closed and the failure impact without each closed pipe is determined. Finally, prioritization of all pipes is made based on their impacts. A flowchart of the whole process is provided in Figure 2 for further illustration.

\[
q_{ij}(t) = \begin{cases} P_{n<} = 0.0 & 0 < P_n < P_{\text{min}}: d(t) \left( \frac{P_n}{P_{\text{min}}} \right)^{1/2} \\ P_{n>} = P_{\text{min}}: d(t) \end{cases}
\]

\[
PI_{ij}(t) = q_{ij}(t) - q_i(t)
\]
\[ PI_{ij} = \sum_{t=0}^{T} PI_{ij}(t) \tag{8} \]

\[ PI_j(FS) = \sum_{i=1}^{N} PI_{ij} \tag{9} \]

\[ PI_j = \sum_{FS} PI_{j}(FS) \tag{10} \]

\[ Pop_j = PI_j/(FS \cdot T \cdot c) \tag{11} \]

where \( d(t) \) — the expected nodal demand at junction \( i \); \( q_i(t) \) — the estimated actual nodal demand at junction \( i \) when there is no pipe failure; \( q_{ij}(t) \) — the estimated actual nodal demand at junction \( i \) when pipe \( j \) is closed; \( T \) — the total duration of model simulation; \( N \) — the total number of junctions; \( c \) — per capita water usage, 120 litre per person per day; \( Pop_j \) — the hourly averaged change in the total number of population served due to failure of pipe \( j \).

Undoubtedly, the VA for the case study network is solvable by an enumeration process that assesses the failure impacts of all pipes under all failure scenarios. Yet, this process might be inefficient and be unable to reveal the system's properties with a systematic view. For instance, the overall operational scheme of water supply to all regions in the service area may still be unclear, as well as the interactions of water transport among different regions. In this regard, the clustering analysis could identify the corresponding sub-distribution networks (clusters) of the served regions, and consequently help understanding of the overall operational scheme, i.e., how the water supply is distributed to customers in each cluster through inter- and intra-cluster connections. Accordingly, it identifies interactions between clusters, and also critical interconnections and clusters. For the reasons given above, the modularity-based clustering method is utilized to improve the VA. (1) The system is mapped into an undirected graph (see ‘Methods and materials: Step 1’). (2) The graph is decomposed into clusters (also see ‘Methods and materials: Step 2’); Figures 3(b) and 3(c) illustrate the decomposed network layout and the corresponding aggregated visualization, respectively. (3) Two groups of pipes, i.e., the interconnections between clusters and the connections to reservoirs and main storage tanks, are first selected for the failure impact (Pop) evaluation described above. (4) Based on the results from step 3, the most critical clusters are identified. (5) The failure impact (Pop) evaluation is carried out for all pipes in the most critical clusters. (6) Prioritization is made for all the pipes selected in steps 3 and 5 according to their impacts.

RESULTS AND DISCUSSION

A dialectical example – the New York Tunnels

The result demonstrated in Figure 4(b) indicates that the method detected clues implied by the network topology, and hence determined the natural cluster structure that matches the terrain partition in the real world very well, despite small discrepancies at certain boundaries. A modularity of 0.564 indicates a strong cluster structure.

Case study on an alpine WDS

Clustering

As Figure 3(b) shows, the cluster structure with the maximum modularity \( Q = 0.752 \) has 11 zones. As with the dialectical example, the method detected the natural division of the case study network, since the layout was divided into both an upper zone and a lower zone. Each zone is comprised of several clusters as subsystems (i.e., clusters 7–11 in the upper zone and cluster 1–6 in the lower zone). From the aggregated visualization (Figure 3(c)), the important role of the interconnection between cluster 4 and 7 is easily identifiable, as that pipe is the hydraulic hub between the upper and the lower zones.

Failure impact evaluation of pre-selected pipes

At this stage, only two groups of pipes are chosen for failure impact assessment (calculation of Pop), i.e., the
Figure 3  |  (a) The alpine water distribution network. (b) Cluster structure of the network. (c) Aggregated visualization of the network. Each coloured area represents a cluster. The meaning of all the symbols remains unchanged for all the remaining figures unless otherwise specified.

Figure 4  |  (a) New York Tunnel System. (b) Cluster structure of New York Tunnel System.
interconnections between clusters and the connections to reservoirs and main storage tanks. In this case, the total number of pre-selected pipes is 24, including 20 interconnections and four others. The results of Pop, ranging from 949 to 0, is shown in Figure 5.

Figure 5 | Failure impact evaluation of pre-selected pipes.

Prioritization of clusters

The clusters are ranked based on two principles, respectively. Next, the most critical clusters in both rankings are chosen for further analysis. The principles are as follows:

*Principle 1* – Cluster A is more critical than B if the averaged Pop of all pre-selected pipes related to it (including interconnections connected to it and the connections to reservoirs and main storage tanks within it, if any) is larger than that of B.

*Principle 2* – Cluster A is more critical than B if the pre-selected pipe with the highest Pop related to it has a larger Pop than that of B.

Using two principles ensures an overall evaluation of failure impacts of the clusters. Principle 1 is used to find clusters that may cause the most severe impacts on average if any pre-selected pipe within it fails. Principle 2 is checked to address clusters that may cause the most severe consequences due to one single pipe failure. The results of cluster prioritization are available in Figures 6(a)

Figure 6 | Prioritization of clusters. (a)–(b) Ranking based on principle 1. (c)–(d) Ranking based on principle 2.
and 6(b). In both cases, clusters 1, 2 and 11 are regarded as the most critical. The failure impacts ($Pop$) of the intra-cluster pipes (31 pipes in total) are thereby calculated following the same process illustrated in Figure 2.

**Identification of all the most critical pipes**

As the last step, both the newly added intra-cluster pipes and the pre-selected pipes are prioritized together to sort out all the most critical pipes of the case study network (Figure 7 and Table 1). For verification, an enumeration process is implemented using the same indicator as for failure impact assessment. The results are compared in Figure 7. As for accuracy, although only a limited number of components was analysed using the authors’ method, the same group of the most critical elements was identified. As for efficiency, both methods are executed on a normal desktop computer. The enumeration process took about 5,100 s to complete, while the clustering-based method speeded up the process by approximately 3.64 times (completed in 1,400 s). The increased speed mainly results from a reduction in the number of failure scenarios considered (from $127 \times 157 = 19,939$ to $127 \times (24 + 31) = 6,985$). More importantly, the method proposed in this study could be used to deepen insights on the properties of WDSs following a systematic downscaling of the network according to its cluster structure. For instance, critical clusters, and interactions between them, could be identified. Hence, the methodology provides an integrated view of a WDS’s characteristics for a given analysis.

Further to the results, a long tail is seen in several cases throughout the case study (Figures 5, 6, and 7), especially in Figure 7. The long tail of pipe criticality may indicate that the vulnerability of WDSs to water supply is dominated by only a few components. From an integrated point of view, a few clusters and interconnections play key roles. From a component-by-component point of view, only a small group of pipes dominates (nine out of 157 in this case). This finding may justify the idea of locating critical components from only a few key clusters, since a vast number of components might be excludable due to comparative unimportance. Furthermore, the method may facilitate prioritisation of water network components for water system management and control, especially for rehabilitation.

**SUMMARY, CONCLUSIONS AND OUTLOOK**

This study explores possible solutions to reduce the layout complexity of WDSs. As this complexity, to a large extent, originates from urban development, a clustering method that may be able to detect the cluster structures of WDSs formed along with urban expansion is proposed. Modularity
Table 1: A summary of the most critical pipes

<table>
<thead>
<tr>
<th>Rank</th>
<th>Pipe ID</th>
<th>Population affected</th>
<th>Location (intra- or inter-clusters)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>60</td>
<td>1,669</td>
<td>Cluster 1</td>
</tr>
<tr>
<td>2</td>
<td>57</td>
<td>1,486</td>
<td>Cluster 2</td>
</tr>
<tr>
<td>3</td>
<td>59</td>
<td>961</td>
<td>Cluster 1</td>
</tr>
<tr>
<td>4</td>
<td>58</td>
<td>949</td>
<td>Interconnection</td>
</tr>
<tr>
<td>5</td>
<td>HZ7</td>
<td>502</td>
<td>Cluster 11</td>
</tr>
<tr>
<td>6</td>
<td>HZ42</td>
<td>499</td>
<td>Cluster 11</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>387</td>
<td>Interconnection</td>
</tr>
<tr>
<td>8</td>
<td>64</td>
<td>370</td>
<td>Cluster 2</td>
</tr>
<tr>
<td>9</td>
<td>3</td>
<td>344</td>
<td>Interconnection</td>
</tr>
</tbody>
</table>

(Q) is the indicator to measure the quality of the graph division into clusters. Q > 0.5 indicates a significant cluster structure.

For verification, the New York Tunnels case is analysed as a dialectical example, since there are obvious boundaries of the natural division of the distribution system. Again, the practical application of the clustering method is demonstrated by case studies on a more complicated distribution system. For the sake of VA, the method is applied to facilitate the identification of all the most critical pipes, given the pipe failure impacts on the number of served population. From the two case studies, the following is found:

(1) The method for clustering might be able to identify the natural division of WDSs. The clusters detected could be subnetworks of corresponding communities of the served urban area.

(2) Both of the WDSs have significant cluster structures (Q > 0.5) formed following urban development.

(3) Compared with enumeration-based identification of critical components, the clustering-based method identified the same group of the most critical components and the same prioritizations of those components in a more computationally efficient time (about 3.6 times faster).

(4) The cluster-based method identified all the same most critical components as the enumeration method.

(5) The cluster analysis provides an integrated view of a WDS’s properties. For instance, the criticality of clusters is measurable together with that of their interconnections, in addition to individual critical components.

For the future, more case studies are required to further confirm the reliability of the clustering method to a variety of WDSs. It would also be meaningful to work out more practical applications of the method. Regarding identification of the most critical components, the approach is still not a fully automated process. In this regard, automation of the whole process could be a next step.

ACKNOWLEDGEMENTS

The work reported was funded by the UK Engineering & Physical Sciences Research Council (EPSRC) project Safe & SuRe. The authors would like to thank Prof. Wolfgang Rauch, Dr Robert Sitzenfrei, and Associate Professor Manfred Kleidorfer at the Unit of Environmental Engineering, Faculty of Civil Engineering, Institute of Infrastructure, University of Innsbruck, for providing the hydraulic model of the alpine WDS.

REFERENCES


Scibetta, M., Boano, F., Revelli, R. & Ridolfi, L. 2013 Community detection as a tool for complex pipe network clustering. EPL 103 (48001), 1–6.


First received 31 January 2014; accepted in revised form 28 May 2014. Available online 13 June 2014