Journal of Hydroinformatics

Journal of Hydroinformatics Vol 00 No 0, 1 doi: 10.2166/hydro.2026.187

A trunk-preserving, demand-weighted edge betweenness centrality framework for partitioning water distribution networks

Rodger Millar Munthali [a,b], Jinliang Gao [a,b,*], Huizhe Cao[c], Wenyan Wu[d], Jingyang Yu[e] and Shihua Qi[e]

[a] School of Environment, Harbin Institute of Technology, Harbin 150090, China
[b] State Key Laboratory of Urban-Rural Water Resource & Environment, Harbin Institute of Technology, Harbin 150090, China
[c] Key Laboratory of Cold Region Urban and Rural Human Settlement Environment Science and Technology, School of Architecture and Design, Ministry of Industry and Information Technology, Harbin Institute of Technology, Harbin 150090, China
[d] School of Engineering and the Built Environment, Birmingham City University, Birmingham B47XG, UK
[e] Department of Municipal and Environmental Engineering, Heilongjiang Institute of Construction Technology, Harbin 150025, China
*Corresponding author. E-mail: gjl@hit.edu.cn

RMM, 0000-0002-7274-6245; JG, 0000-0002-6662-0187

ABSTRACT

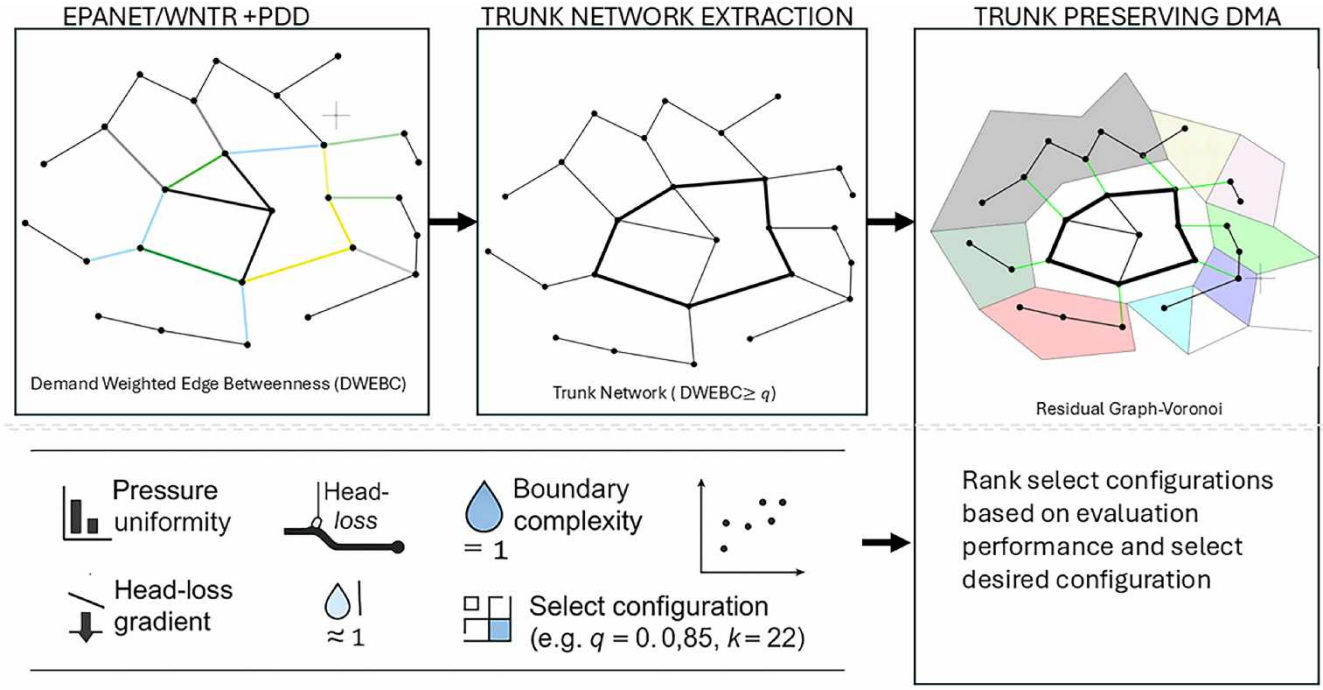
District metered areas (DMAs) play a crucial role in managing leakages and pressure, and monitoring operations within water distribution networks (WDNs). However, purely topological design of DMAs overlooks hydraulics and the operational role of highly critical pipes, undermining redundancy and resilience. This study proposes a trunk-preserving hydraulically informed partitioning framework that (i) utilises demand weighted edge betweenness centrality (DWEBC) to delineate a trunk network (TN) and (ii) grows capacitated, size-balanced DMAs on the residual network using an inverse of the DWEBC as edge weight. The approach employs farthest-first *k*-centre seeding, graph-Voronoi growth, Kernighan–Lin/Fiduccia–Mattheyses connectivity repair, uniformity balancing, and minimum DMA size enforcement. The DMAs are then evaluated using a pressure uniformity index (PUI), head-loss gradient (HLG), demand satisfaction ratio (DSR), and boundary control complexity (BCC) metrics, after which collector edges interfacing TN and DMAs are flagged. The approach is deterministic, bounded in runtime, and implemented in Python using EPANET/WNTR. Results on the benchmark networks demonstrate easy identification of TN, formation of balanced DMA in terms of size and pressure, and interpretable interfaces for metering and control. The framework generalises to real-world WDNs and can seed multi-objective refinement if Pareto exploration of PUI–HLG–DSR–BCC is desired.

Key words: demand weighted edge betweenness centrality (DWEBC), district metered areas (DMAs), graph partitioning, network resilience, trunk network

HIGHLIGHTS

- A trunk network preserving partitioning framework that keeps critical pipes outside any district metered area (DMA) to ensure resilience.
- Establishes graph-Voronoi DMAs on a residual graph using demand weighted edge betweenness centrality and Kernighan–Lin/Fiduccia–Mattheyses.
- Evaluates DMAs and identifies collector edges for metering and control.
- The method yields DMAs with lower pressure variance and boundary complexity.

GRAPHICAL ABSTRACT



1. INTRODUCTION

The water supply infrastructure, which is designed and operated to supply potable water to communities with acceptable levels of satisfaction (Pagano *et al.* 2019; Yu *et al.* 2024), relies on effective partitioning of the water distribution network (WDN) into district metered areas (DMAs) to reduce real losses, improve pressure management, and enable targeted metering (Morrison *et al.* 2007; Giustolisi & Ridolfi 2014; Savić & Ferrari 2014; Campbell *et al.* 2015). However, in the process of partitioning the WDN, the community detection and modularity-based methods can include hydraulically critical pipelines inside a DMA, which reduces network redundancy and risks service disruption during operation or failures. DMAs initially were rigid structures with a fixed topology; however, this has changed as now we have dynamic DMAs that are design(ed) for control (dfc) and adjust according to the prevailing conditions in the WDN (Sharma *et al.* 2022; Ulusoy *et al.* 2022). As such, an accurate and practical partitioning of a WDN must be hydraulically informed, operationally conventional, and computationally tractable.

WDN partitioning falls into two main phases: clustering and sectorisation. The clustering algorithms identify a number of feasible DMAs by identifying clusters. This is done through the aggregation of nodes and pipes in the WDN based on metrics that reflect the level of connectivity and relevant hydraulic or operational characteristics. On the other hand, sectorisation procedures physically separate the network by installing valves and/or flow metres at strategic positions identified through graph theory (GT) (Giudicianni *et al.* 2020; Sharma *et al.* 2022; Creaco *et al.* 2023). Apparently, most of the methods used for WDN analysis and partitioning use GT (Khoa Bui *et al.* 2020). A graph is a pair of sets $G = (V,E)$ where V represents the vertices (nodes or junctions or tanks or reservoirs) and E represents edges (links or pipes or pumps or valves) (Schaeffer 2007; Shekofteh *et al.* 2023). The use of GT for partitioning involves establishing sub-graphs within the network. This ranges from sectorisation approaches that use transversal and shortest path methods to modern clustering modularity frameworks. These include the classical graph methods like Girvan–Newman betweenness edge removal and modularity maximisation or simply tailoring modularity to WDNs (Giustolisi & Ridolfi 2014). Modularity-based community detection, such as Fast-Greedy/Louvain and random-walk/Infomap variants, has been transferred from complex networks to WDNs. Comparative studies on EXNET show trade-offs, as Fast-Greedy often attains higher modularity. Multilevel recursive partitioning (METIS) yields well-balanced cluster sizes, useful for uniform DMAs, whereas random-walk methods can minimise cuts but at the expense of balance (Liu *et al.* 2018). Weighted spectral methods, graph Laplacians, integrate hydraulic

parameters such as diameter, length, conductance, and flow via edge weights, which improves partition layouts relative to purely topological partitions, and offers a mathematical route to DMA formation (Di Nardo *et al.* 2017). Spectral normalised cuts, Kernighan–Lin/Fiduccia–Mattheyses (KL/FM) refinement, and multilevel partitioning methods, such as METIS/Karlsruhe High Quality Partitioning, which pursue balanced and well-cut partitions, are also used for WDN partitioning. However, they are not tailored to hydraulic weights nor to operational trunk constraints.

There have been many studies on WDN partitioning. The depth-first search (DFS) and breadth-first search (BFS) use an exploration approach that tracks nodes to establish connectivity between nodes that have been visited. DFS and BFS were used on a network in Mexico to identify nodes belonging to sectors and examine configurations of sectors and their size and node linkages (Tzatchkov *et al.* 2012). Hence, stronger connections could be identified between nodes, as exemplified by the independent works of Ostfeld and Lifshitz (Perelman & Ostfeld 2011; Khoa Bui *et al.* 2020). A DFS has also been used to partition a single-source and multiple-source WDN (Tzatchkov & Alcocer-Yamanaka 2019). Others have used these algorithms combined with other methods to identify sectors. The DFS, BFS, and the algorithm for finding the shortest path (Dijkstra algorithm) were used to automatically create DMAs in a WDN (Alvisi & Franchini 2014). There has also been work on using the Floyd–Warshall algorithm to partition a WDN into DMAs (Gomes *et al.* 2012). Giustolisi & Ridolfi (2014) developed a modularity index tailored to WDNs and applied it to an example network.

High connectivity indicates a robust network that can withstand disruptions in specific locations without compromising water delivery to other areas (Ulusoy *et al.* 2022). Conversely, low connectivity suggests a more fragile network where a single pipe failure could isolate entire sections. That is the graph-theoretic centrality measures are used as rapid, surrogate indicators of critical elements in network analysis. The measures assign values to nodes and edges that represent their importance in the network (Brandes 2008; Ugurlu 2022). Betweenness centrality (BC), which is one of the most popular measures that is based on the shortest path principle (Brandes 2008; Ugurlu 2022), defines a node/edge as being central if it has many shortest paths between pairs of nodes passing through it (Zarghami 2018).

While BC is widely used in identifying and assessing central nodes (junctions) and edges (pipes) (Brandes 2008; Bergamini & Meyerhenke 2016; Zarghami 2018; Ugurlu 2022) and defining the partitions, it has several limitations when applied to WDNs owing to its focus on topology and the shortest path principle. By definition, it calculates the number of shortest paths passing through a node. However, water flow paths, just like electric current and vehicles in traffic, are not equal in terms of service consequence (Newman 2005; Ulusoy *et al.* 2018), for hydraulic principles and parameters such as demand, pressure, and pipe diameters affect the flow paths and operation of the WDN. The traditional BC is applied to a static WDN, while WDNs are dynamic in nature as flow changes based on demand, head loss, operational controls, and pressure, making the purely topological BC metric less representative of the WDN. Although previous studies have incorporated some hydraulic parameters, a clear gap remains. Existing metrics may include flow in the edge weights, but they generally neglect spatial heterogeneity in demand; specifically, no current formulation of edge betweenness simultaneously captures the demand contributions of both the origin and destination along a path. A pipe that sits on many shortest paths between two low-demand nodes can score as highly as one that connects major supply and demand centres, even though the latter's failure would inflict far greater service loss. This work aims to ensure that the BC reflects the physical and operational conditions and is hydraulically relevant, and explores extending Brandes's algorithm by using the demand at the target nodes in the edge weights calculations for the edge BC. The metric can highlight pipes whose failure would sever high-demand corridors, even if they lie on less-travelled topological backbones. The work uses a demand-weighted edge betweenness centrality (DWEBC) metric to delineate a WDN trunk network (TN).

The metric is formulated and implemented in Python using the Water Network Tool for Resilience (WNTR). We validate the framework on D-town, Battle of the Water Sensor Network 2 (BWSN2), and Exnet benchmark networks. While other studies have tried to apply the TN isolation duration partition (Campbell *et al.* 2015), this study takes a different approach through DWEBC instead of just the accumulated shortest path value as in Campbell *et al.* (2015). The metric offers a computationally efficient, globally informed ranking that is responsive to both network structure and demand distribution, and it fills the identified gap.

WDN partitioning studies have used modularity, spectral cuts, and BC to guide DMA design, but a time-robust, hydraulically weighted BC that identifies a TN that is to be preserved in partitioning the WDN remains under-explored.

Building on these insights, this study proposes a framework that identifies the TN in a WDN using a time-robust DWEBC. The DWEBC is computed on directed hourly graphs using weighted edges that incorporate nodal demand and head loss and then averaged across the timesteps to obtain a robust score. The TN is identified through quantile thresholding of the

DWEBC scores. This TN, with its edges and their start and end nodes, remains unassigned to any DMA. On the residual graph, partitions are realised using a capacitated graph-Voronoi growth obtained from k-centre seeds at hydraulic distance (inverse of demand-weighted) from the TN. This is followed by KL/FM partition refinement, connectivity repair, and uniformity balancing, with the user able to specify a minimum size per DMA. The resulting DMAs are then assessed via the proposed pressure uniformity index (PUI), head-loss gradient (HLG), demand satisfaction ratio (DSR), and boundary control complexity (BCC) metrics, and we flag collector edges, edges connecting TN to the DMAs, for instrumentation.

This paper contributes (1) a demand, direction, and length aware edge betweenness metric (DWEBC) that identifies hydraulically significant pipes and stabilises rankings over diurnal operation, (2) a trunk preserving partitioning workflow that ensures critical pipes are outside any DMA and grows zones on the residual network, with size and connectivity safeguards, and (3) an interpretable evaluation set of PUI, HLG, DSR, and BCC metrics, with explicit identification of DMA linking/collector edges for valve placement and control.

Section 2 describes the DWEBC and partitioning methodology, Section 3 presents the D-town, BWSN2, and Exnet application and results, and Section 4 summarises the conclusions.

2. METHODS

We propose a GT partitioning method for WDN that leverages DWEBC computed under a pressure-dependent demand (PDD) hydraulic engine to identify a TN and develop partitions on the residual graph. We initially represent the WDN as an undirected graph $G = (V, E)$ with nodes (V) as junctions, reservoirs, and tanks and edges (E) as pipes. Using EPANET/WNTR, we simulate the network for 24 hourly timesteps that result in outputs as flowrate (Q_e^t), head loss (h_e^t), and nodal demand (d_i^t) for every edge (e) and node(i) and timestep (t). For each t, we orient the pipe in the direction of the simulated flow, that is, reversing if ($Q_e^t < 0$) to obtain a directed graph, G_t.

The edge costs for the DWEBC calculation are defined for the direction-aware graphs, termed demand weighted costs for each edge e = (s,t), using the edge weighting function as follows:

$$w_{(s,t)} = \alpha_{(s,t)} l_{(s,t)} \times \frac{1}{d(t) + \epsilon} \tag{1}$$

where $l_{(s,t)}$ is the pipe length between demand nodes s and t, $d(t)$ is the demand at the target node t (downstream demand), ϵ is a small constant to avoid division by zero, and α is a scaling factor calibrated using observed data: typically (h_e^t), $l_{(s,t)}$, and $d(t)$. The scaling factor α controls the relative importance of the demand compared with the pipe length. This has been modelled to be in the order of observed (computed head loss). The scaling factor calibrates the weighting function to a physically meaningful range, as revealed through dimension analysis, since the weighting function has dimensions comparable to head loss, which facilitates interpretation of DWEBC as a surrogate for energy-weighted flow importance.

By inversely weighting demand in Equation (1), high-demand nodes incur a lower traversal cost; shortest-path algorithms will thus preferentially route through corridors serving great demands, aligning path enumeration with service-loss priorities. BC interprets weight as an edge cost; hence, shorter pipe length will be preferred. In the same understanding, a path with a lower target node demand, with a lower cost, will be regarded as a high-cost path, which is against the goal of what we are trying to achieve, for in WDN, higher demands should generally have higher centrality and are critical in the network. This way, low-demand nodes (with high inverse demand) will appear more central in the network, while nodes with high demand (low inverse demand) will appear less central. Intuitively, hydraulically favourable, high-demand corridors are cheaper and more likely to carry the shortest-path flow.

We consider a case prevalent in networks with heterogeneous materials or conditions, in which the WDN is overly sensitive to nodal demand and pipe material, as when we have widely varying demand patterns and profiles, and the network has widely varying pipe materials, as opposed to the lengths between nodes. In this scenario, the variable ($\alpha_{(s,t)}$) is defined as:

$$\alpha_{(s,t)} = h_{f(s,t)} \times \frac{d_{(s,t)}}{l_{(s,t)}} \tag{2}$$

In Equation (2) for the variable scaling factor, the terms $h_{f(s,t)}$, $l_{(s,t)}$, and $d_{(s,t)}$ are computed for the pair of nodes concerned, in which $d_{(s,t)}$ is the average of the demand at nodes s and t. However, $h_{f(s,t)}$ and $d_{(s,t)}$ are not averaged over the timesteps for EPS.

For each G_t, we compute the weighted edge BC using $w^t_{(s,t)}$ as the path cost. We then aggregate over time (average across 24 h) onto the undirected physical graph H, yielding a robust DWEBC score DWEBC_e for every pipe.

A TN (backbone) $B \subseteq E(H)$ is defined as a set of edges with DWEBC_e above a quantile threshold q_b. This set represents pipes with high, system-wide conveyance significance that are critical to maintaining supply in the network. As such, these edges must not become partition boundaries or be included in a partition. We then build a graph on the residual distribution network and establish connectors C, which are edges crossing the interface between the TN and residual nodes, such that all valves and metres, hence control and measurements, are meant to be on C:

$$C = \{(s,t) \in E : s \in T,\ t \notin T \text{ or } t \in T,\ s \notin T \tag{3}$$

We build super nodes on the residual graph and grow the partitions based on the desired number of DMAs, where each component (including isolated nodes) is a must-link group we call a super node. A weighted and undirected graph is established using the initial graph and a hydraulic weight per edge, $w^-_{(s,t)}$, defined as:

$$w^-_{(s,t)} = \left(\frac{1}{\frac{1}{T}\sum_{t=1}^{T} w_{(s,t)} + \varepsilon} \right) \gamma_{(s,t)} + \eta L_{(s,t)} \tag{4}$$

In the previous equation, we define the edge cost $w^-_{(s,t)}$ in terms of the inverse of the mean of $w_{(s,t)}$ over time T, $\gamma_{(s,t)}$ = trunk penalty > 1 for $(s,t) \in B$ and equal to 1 otherwise, and $\eta \geq 0$ is a length weight factor to gently penalise elongated DMAs. We do not remove trunk edges, penalising them simply discourages DMA growth across the trunk, with a larger $\gamma_{(s,t)}$ representing a stronger discouragement of cross TN growth.

The boundary edges are selected only from C or between DMAs, never on TN, and if a proposed boundary would touch B, it must be redirected to the nearest feasible C.

The algorithm aims to partition the residual WDN/graph into k DMAs that (i) respect a TN which is essential for system-wide water supply; (ii) are hydraulically coherent; and (iii) are operationally practical with a bounded size and manageable boundary nodes.

We hence form the residual node set $V \backslash T$ and its connected components $\{C_i\}$ in H. To avoid insignificant DMAs and respect a minimum DMA size $S_{\min}$ (e.g., 15 nodes), we compute a per-component seed quota as follows:

$$k_i \approx \min\left(\frac{|C_i|}{S_{\min}},\ \mathrm{round}\left(K \cdot \frac{|C_i|}{|V \backslash T|} \right) \right) \tag{5}$$

with a global budget $\sum_i k_i \approx K$ where K is the target number of DMAs/partitions. We resolve rounding by distributing remaining seeds to components with the largest fractional share until the budget is met. Within each component C_i, we pick k_i seed nodes by farthest-first (k-centre) on the weight metric, that is, starting from the node farthest from the TN in $w^-_{(s,t)}$, and then iteratively add the node maximising its minimum weight-distance to existing seeds. This is the classic Gonzalez's algorithm (Gonzalez 1985), adapted to hydraulic distance and residual components.

For speed and determinism, single-source Dijkstra distances $d_s(\cdot)$ are precomputed on H for each seed s using edge weights $w^-_{(s,t)}$, which are reused every subsequent step of assignment and balancing.

We let the number of elements in the residual node set be represented by N_{res}, such that $N_{\mathrm{res}} = |V \backslash T|$, and K_s be the number of seeds. We target an average DMA size $\bar{S} = N_{\mathrm{res}}/K_s$ and define a uniformity band:

$$S_{\mathrm{lo}} = \max(S_{\min}, \lfloor \beta_{\mathrm{lo}} \bar{S} \rfloor),\ S_{\mathrm{hi}} = \max(S_{\mathrm{lo}}, \beta_{\mathrm{hi}} \bar{S}) \tag{6}$$

with $\beta_{\mathrm{lo}} \approx 0.9$, $\beta_{\mathrm{hi}} \approx 1.1$. The lower bound β_{lo} prevents the formation of undersized DMAs, while the upper bound β_{hi} limits excessive aggregation. We sort non-trunk nodes by their nearest-seed distance and assign each to its closest seed, capping each seed's size at S_{hi}. If the closest seed is full, it is assigned to the nearest non-full seed, or the currently smallest seed if all are full. This procedure balances initial DMAs without crossing the TN heavily, using the penalty in $w^-_{(s,t)}$.

Within each DMA, we identify connected components, and any non-main fragments, so-called orphans, are reattached to neighbouring DMAs. The target DMA is chosen by (i) maximum boundary strength across the cut (sum of $w^{-}_{(s,t)}$) on the cut edges, else (ii) nearest-seed distance. We bound this step to a few iterations to ensure predictable runtime.

To further reduce size variance, we iteratively grow undersized DMAs ($< S_{\text{lo}}$) by pulling boundary nodes from oversized neighbours, favouring moves with the smallest distance penalty $\Delta d = d_{\text{new seed}} - d_{\text{current seed}}$ to avoid disconnecting the donor DMA. We also shrink oversized DMAs ($> S_{\text{hi}}$) by donating low-penalty boundary nodes to undersized neighbours. Each pass is bounded; hence, moves are accepted only if they keep donors above $S_{\min}$ and below a set Δd threshold. This creates uniform-sized DMAs while preserving hydraulic plausibility. In concept, this is similar to balanced weighted label propagation (Pirouz 2021) and the KL/FM refinement algorithm (Dongarra *et al.* 2003), with balance constraints but using demand weighted edges as the objective proxy.

Any remaining DMA smaller than $S_{\min}$ is merged into the most compatible neighbour, chosen by (i) strongest boundary strength or (ii) nearest-seed distance. This guarantees no tiny DMAs survive.

After labelling, we designate collector edges as edges with one node in T (trunk) and the other in some DMA. These edges are the natural metering and isolation edges.

The quality of the resulting DMA k is evaluated using four main indicators: PUI, HLG, DSR, and BCC, as follows:

(a) PUI

It is desirable to have uniform pressure through the WDN; the partitioning process introduces unforeseen variations that may result in pressure deficit or increased pressure in some nodes. The PUI hence measures the extent to which the partitioning process has achieved uniformity in the DMAs formed, defined as:

$$\text{PUK}_k = \frac{\sigma(P_k)}{\mu(P_k)} \tag{7}$$

where P_k is the vector of nodal pressures in partition k, μ is the mean pressure, and σ is the standard deviation. A lower PUI indicates better pressure balance within the partition, assisting in leakage reduction methods.

(b) HLG

As a consequence of partitioning, water may take a longer path to a demand node because of valve closure or pipe removal. This increases or decreases energy loss in the partitions; as such, there is a need to quantify the effect using HLG:

$$\text{HLG}_\text{k} = \frac{1}{|E_k|} \sum_{(s,t)\in E_k} \left| \frac{H_s - H_t}{L_{(s,t)}} \right| \tag{8}$$

where H_s, H_t are the hydraulic heads at nodes s and t, $L_{(s,t)}$ are the lengths of pipes between node s and t, and E_k is a set of pipes (edges) in partition k. This parameter measures the friction loss frequency, in which lower values indicate hydraulically coherent zones.

(c) DSR

While Di Nardo *et al.* (2014) call the term flow deficit index (FDI), here, it is termed DSR, as there can be oversupply, not just deficit:

$$\text{DSR}_\text{k} = \frac{\sum_{i\in V_k} q_{i,\text{ supplied}}}{\sum_{i\in V_k} q_{i,\text{demanded}}} \tag{9}$$

where a DSR that approaches 1 signifies that the partition is meeting demand with minimal deficit or oversupply.

(d) BCC

Since DMA design has evolved from static entities to dynamically configurable partitions, it is imperative that boundaries are as simple as possible, such that any control settings for reopening and closing of valves take into account the network topology (Ulusoy *et al.* 2022) and the steady and unsteady hydraulic behaviour of the WDN (Wright *et al.* 2014; Wright 2016). This ensures the maximisation of resilience and minimisation of pressure deviations (Ulusoy *et al.* 2022). We take

a simpler approach that considers nodes in a partition as follows:

$$\mathrm{BCC}_k = \frac{|B_k|}{|V_k|} \tag{10}$$

where B_k is the number of boundary nodes interfacing with other partitions and V_k is the number of nodes in partition k. This BCC_k parameter reflects the operational simplicity of the partition, in which a lower BCC is preferred for control and monitoring.

All iterative phases are explicitly bounded to a small number of iterations, and the most expensive operations use precomputed single-source Dijkstra maps from each seed. For ≤ 30 seeds, precomputation dominates runtime, and subsequent steps are near-linear.

The algorithm is presented as follows.

ALGORITHM. Identify TN and C and Partition WDN (WN, *K*, *S*_min, *q*, gamma, beta_lo, beta_hi, eta); R1 and R2 are small user iteration limits

simulate 24h PDD → states[t], t=1…24
for t in 1…24:
    build directed G_t by flow sign
    compute $w(t, e)$ for all edges e in G_t
    EBC_t = edge_betweenness(G_t, weight = $w(t, \cdot)$)
DWEBC(e) = mean_t EBC_t(e)
B = (e: DWEBC(e) ≥ quantile_q(DWEBC)); T = endpoints(B)
build $H = (V, E)$ (undirected)
for e in E:
    cost(e) = (1/(mean_t $w(t,e)$ + eps)) * (gamma if $e \in B$ else 1) + eta * length(e)
V_res = $V \setminus T$; components = CC(H[V_res])
allocate seeds per component with quotas ~ |C_i| and ≥S_min
for each component C_i:
    seeds(C_i) = FARTHest_FIRST_kcenter(C_i, dist = shortest_path(cost))
seeds = $\bigcup\_i$ seeds(C_i); precompute DISTS_SEED[s] for all seeds s on H
Sbar = |V_res|/|seeds|; S_lo = max(S_min, floor(beta_lo*Sbar)); S_hi = ceil(beta_hi*Sbar)
init labels $L(v)$ = −1 if $v \in T$ else nearest_seed_under_cap(v, S_hi, DISTS_SEED)
repeat ≤ R1 (anti-orphan):
    for each DMA k: reattach fragments to best neighbour (max boundary strength, else nearest seed)
repeat ≤ $R2$ (uniformity):
    grow undersized DMAs by pulling low-Δd boundary nodes (skip donor articulations; keep donors ≥ max(S_lo,S_min))
    shrink oversized DMAs donating low-Δd boundary nodes to undersized neighbours
for any DMA |V_k| < S_min: merge whole DMA into best neighbour (boundary strength, else nearest seed)
relabel DMAs to 0..K' − 1
E_col = { (u,v) ∈ E: [$u \in T$ XOR $v \in T$] AND [L(other) ≠ −1]}
compute PUI, HLG, DSR, BCC per DMA; aggregate
return L, E_col, metrics

3. RESULTS AND DISCUSSION

3.1. Case studies

The proposed framework is demonstrated on the D-town network (Figure 1(a)) (Marchi 2021), a medium network that is composed of one reservoir, seven tanks, 399 junctions, 443 pipes, and 11 pumps. It is also demonstrated on the Battle of the Water Sensor Networks (BWSN) Network 2 (BWSN2) (Figure 1(b)), a complex and large network consisting of 12,253 nodes, 14,822

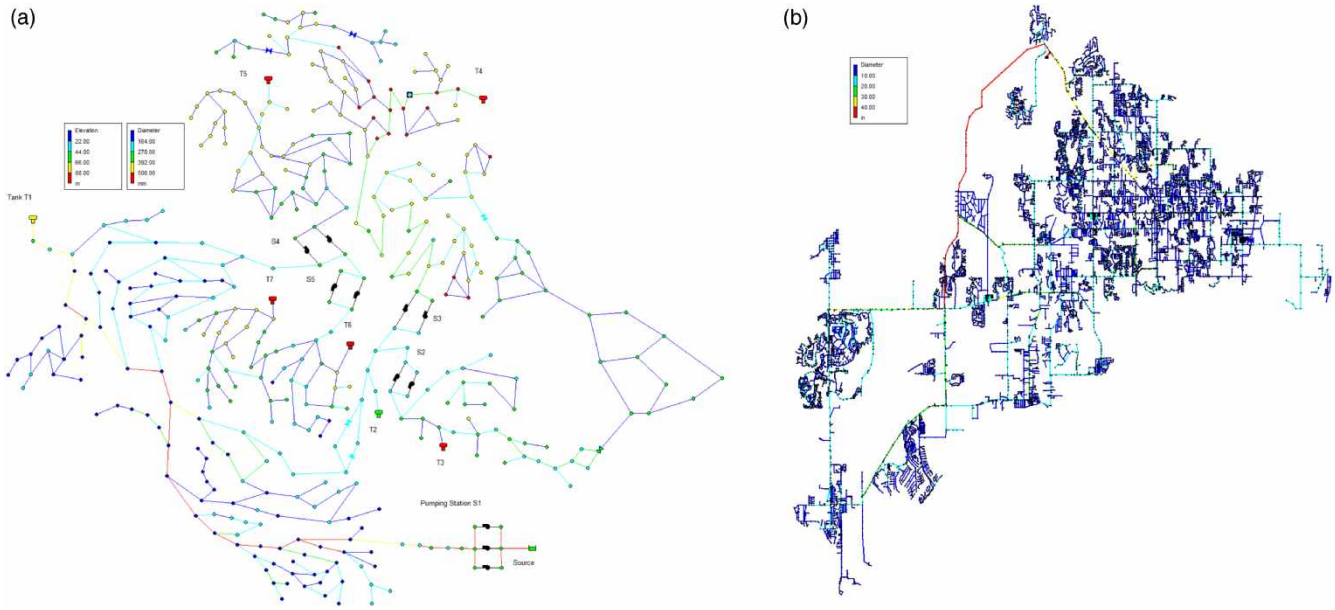


Figure 1 | (a) D-town and (b) BWSN2 raw water network layout.

pipes, two reservoirs, two tanks, five valves, and four pumps (Ostfeld *et al.* 2008). For the DWEBC calculation, pumps, pipes, and valves immediately after water sources were excluded, so that they do not form part of the criticality ranking and TN.

For each hourly timestep, nodal demands, pipe flows, and head losses were exported and used to compute the directed, DWEBC, and the associated hydraulic weights used in the partitioning procedure.

3.2. TN identification using DWEBC

For each t, DWEBC was computed on G_t, a graph that was obtained by orienting each pipe along the simulated flow direction. The 24-h edge scores were then averaged and mapped back onto the undirected physical graph, such that we had a robust DWEBC value for every pipe.

The TNs were obtained by thresholding the DWEBC distribution at quantiles, $q \in \{0.80,\ 0.85,\ 0.9,\ 0.95,\ 0.97\}$ for D-town and $q \in \{0.90,\ 0.92,\ 0.94,\ 0.96,\ 0.98\}$ for BWSN2, resulting in edges with DWEBC values above the threshold being designated as trunk edges with their endpoints forming the node set. This yielded TNs spanning the major corridor of the networks, aligning topology and hydraulics with the main supply routes from the reservoir and tanks to the demand clusters. As q increases, the TN shrinks to the most critical pipes, increasing the size of the residual graph available for partitioning. For the D-town network, a lower threshold, e.g., $q = 0.80$, created more extensive TN with many high-centrality pipes excluded from any DMA, whereas higher thresholds, e.g., $q = 0.97$, yielded a more compact TN but granted greater freedom for the residual partitioning. As can be observed, this trade-off directly affects both the achievable number of DMAs and their spatial coherence, which also affects the hydraulics of the WDN. For the BWSN2 case, high TN thresholds were used because of the type of network topology. BWSN2 consists of an elongated trunk pipeline with a large diameter and has a low loop density, an average node size degree of 2, with meshed connectedness of approximately 0.12. The network is classified as a distribution network with branch transmission and a sparse-grid (SG) distribution network topology (Hwang & Lansey 2017).

3.3. Partitioning design space

On the residual network, capacitated graph-Voronoi growth was performed for the target DMA $K \in \{5,\ 6,\ 7,\ 8,\ 9\}$ and $K \in \{22,\ 25,\ 28,\ 30,\ 35\}$ for D-town and BWSN2, respectively. Hence, for each $(q,\ K)$, seeds were allocated per component according to Equation (5) given, while enforcing a minimum DMA, $S_{\min} = 10$ nodes, and $S_{\min} = 350$, respectively. Farthest-first k-centre seeding using $w^{-}_{(s,t)}$ generated initial centres that were well separated in terms of energy and demand support. We set $\eta = 0.001$ to avoid elongated DMAs for both networks, while $\gamma_{(s,t)}$ was set to 6.0 for

D-town and 2.0 for BWSN2 to discourage DMAs growth on the TN. Anti-orphan consolidation removed residual fragments, and uniformity balancing ensured reduced size dispersion without degrading connectivity. The final networks highlight collector/connector edges at TN–DMA interfaces, aligning with expected metering locations. This resulted in the networks being partitioned 25 times, with each partitioned network having a different number of partitions. While the algorithm parameters were set up to ensure we do not get the isolated clusters and clusters with few nodes, the fixed trunk results in hard graph-theoretic feasibility limits, as the partitions must be on the residual graph. This sometimes gives away isolated nodes that must be added to fully grown partitions through inspection of the network; if necessary, this is especially true for the BWSN2 network owing to its topology.

3.4. Comparison of candidate configurations

For all the 25 configurations realised, the resulting DMAs were evaluated using the proposed PUI, HLG, DSR, and BCC metrics, with the results for D-town summarised in Figures 3–5. An analysis of the resulting visual network outputs and the metrics led to five configurations being identified as having desirable metrics results, namely configuration (i) $q = 0.80$, $k = 6$, (ii) $q = 0.80$, $k = 7$, (iii) $q = 0.95$, $k = 9$, and (iv) $q = 0.97$, $k = 9$, as shown in Figure 2, and configuration (v) $q = 0.85$, $k = 8$ (Figure 3).

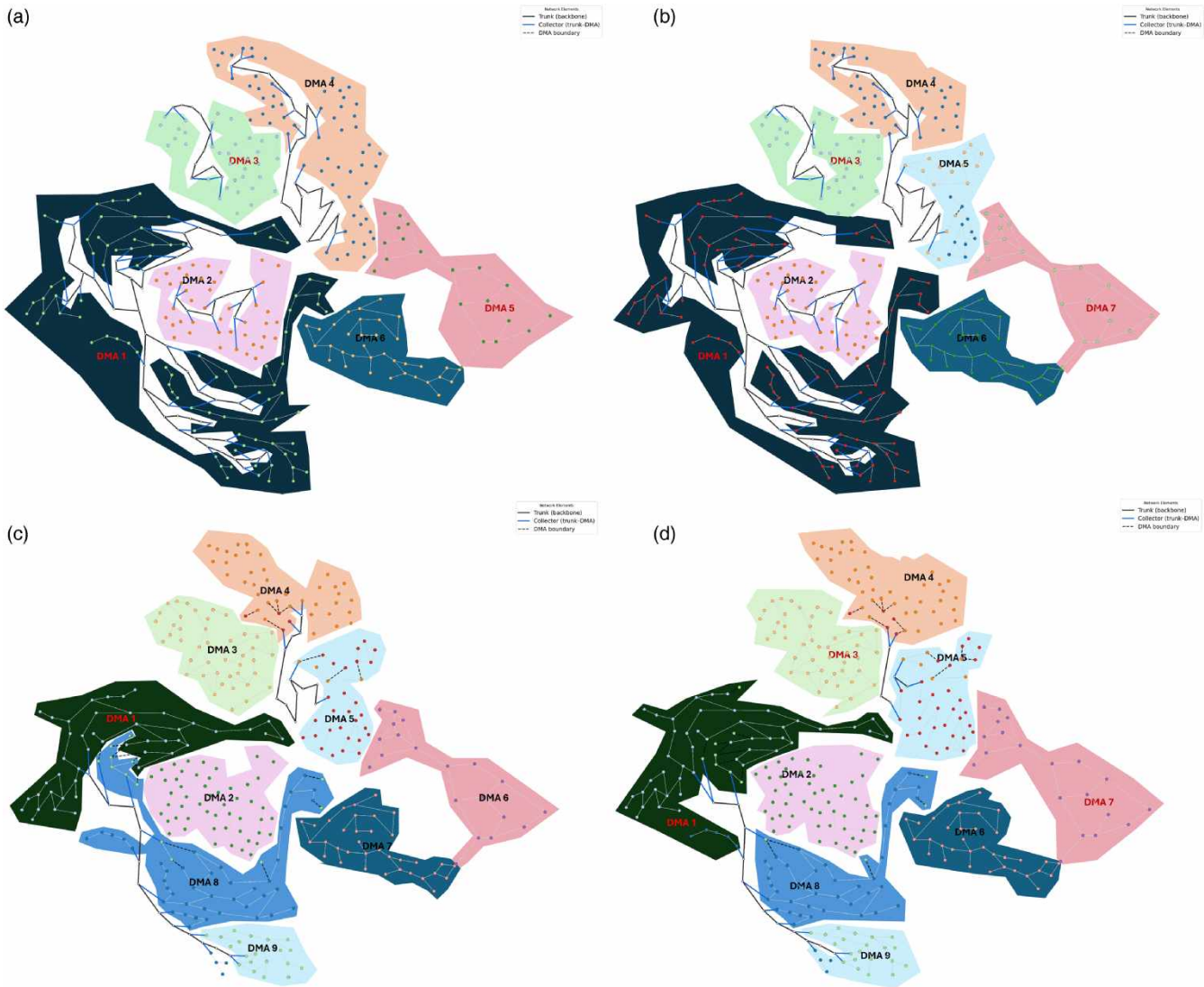


Figure 2 | D-town partitioned (a) $q = 0.80$, $k = 6$; (b) $q = 0.8$, $k = 7$; (c) $q = 0.95$, $k = 9$; and (d) $q = 0.97$, $k = 9$.

Table 1 reports summary statistics for these five configurations, from which it is observed that configuration (ii) achieves lower PUI and HLG, indicating more uniform pressure and lower average head-loss gradient (HLG) within DMAs, together with a lower BCC, implying simpler boundaries and fewer boundary nodes per DMA. Configuration (i) exhibits slightly better performance in terms of BCC but at the expense of a slightly higher HLG, and inferior PUI to (ii) in several partitions.

PUI, HLG, and BCC should be as low as possible for leakage control, energy efficiency, and operational tractability; hence, the metrics results collectively favour configuration (ii) ($q = 0.80, K = 7$). The better DSR value for $q = 0.80, K = 6$ does not compensate for the substantially higher values in the other two indicators. However, other methods could be utilised to decide the right configuration among the best identified, considering that configuration (v) as shown in Figure 4 is a special case that visually gave DMA no. 1 that could not be defined properly, but it resulted in the best PUI, HLG, and second best DSR as shown in Table 1.

Figures 3 and 5 present all experimental results, from which it can also be observed from the means that configuration (ii) with $q = 0.80, K = 7$ outperforms all other solutions in the design space as it offers the most balanced overall performance. The selected configuration is similar in structure to those obtained by the work of Diao *et al.* (2016) on a similar network (C-Town).

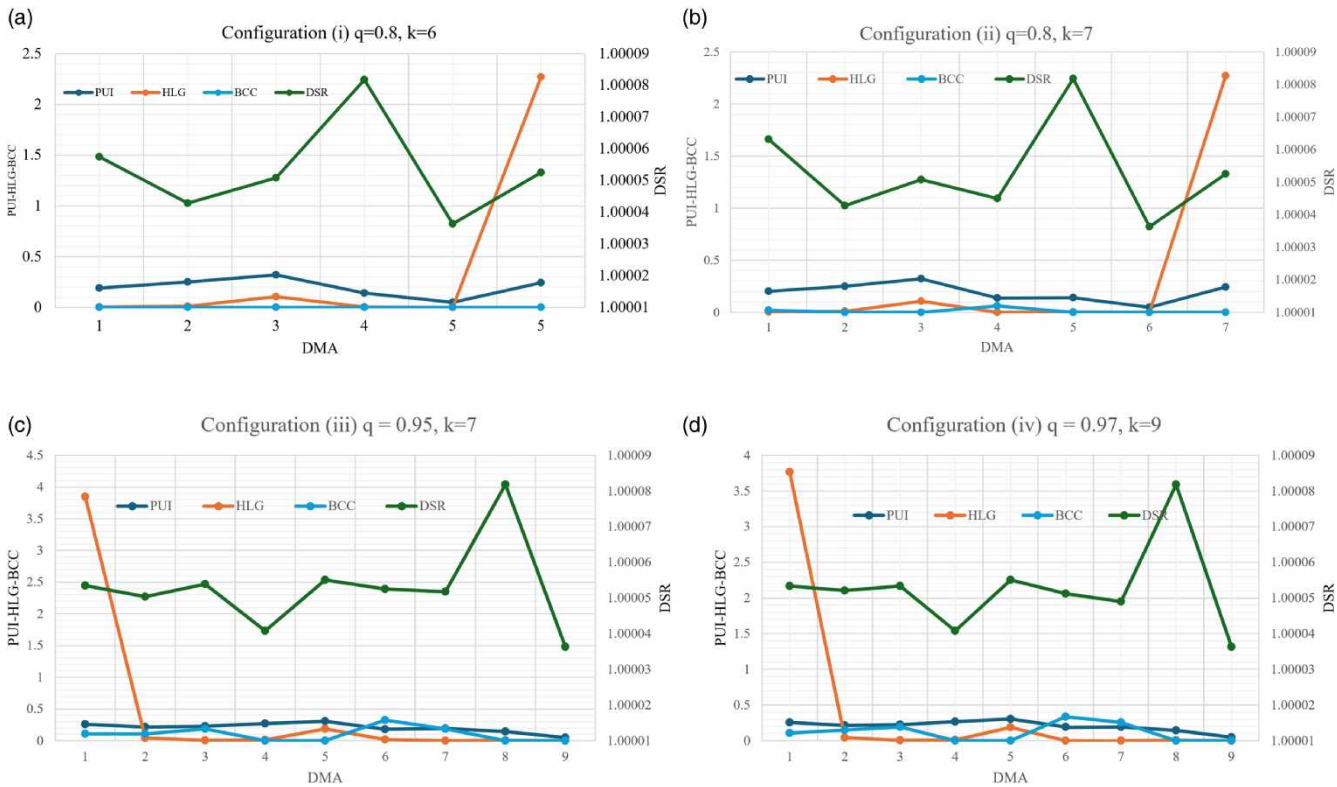


Figure 3 | D-town (a) configuration (i), (b) configuration (ii), (c) configuration (iii), and (d) configuration (iv) evaluation metrics.

Table 1 | D-town evaluation metrics for configurations (i)–(v) at 75% quantile

Configuration	PUI	HLG	DSR	BCC
$q = 0.80, k = 6$	0.257	0.084	1.000056	0.0
$q = 0.80, k = 7$	0.246	0.059	1.000058	0.01
$q = 0.85, k = 8$	0.209	0.046	1.000054	0.11
$q = 0.95, k = 9$	0.258	0.040	1.000054	0.18
$q = 0.97, k = 9$	0.256	0.042	1.000053	0.19

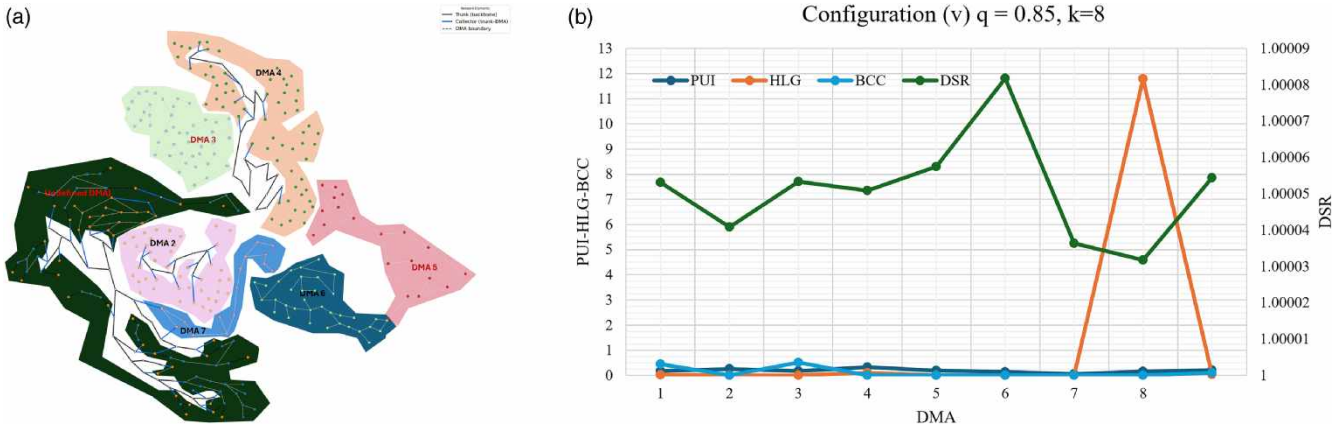


Figure 4 | (a) D-town partitioned configuration (v) and (b) configuration (v) evaluation metrics

Further to Figures 5, which gives the means of the evaluation metrics per partition, Figures 6 considers the entire WDN without partitions. For this entire WDN we plot the evaluation metrics alongside the means for the values in $q \in \{0.80,\ 0.85,\ 0.9,\ 0.95,\ 0.97\}$. For D-town, for each q (except for $q = 0.85$), increasing the number of DMAs reduces PUI, with $q = [0.80,\ 0.97]$ recording favourable metrics as before.

For the BWSN2, as a result of its network topology, DMA formation differs from the D-town. Feasible DMAs are largely constrained by the SG topology, as there are limited alternative paths and many subgraphs become disconnected upon the removal of critical edges. The proposed trunk-preserving DWEBC framework remains applicable in this regime, as high-DWEBC edges correspond predominantly to single-corridor conveyance routes whose failure would disconnect large downstream demand regions. As these edges are unassigned, it avoids placing DMA boundaries across global critical edges.

However, the sparse nature of the network introduces practical implications to apply the methodology. First, the achievable number of DMAs K_{actual} is often topology-limited, because the residual graph may split into small subgraphs after TN extraction, and components smaller than the minimum DMA size $S_{\min}$ must be merged by inspection. Second, enforcing tight uniformity constraints (β_{lo}, β_{hi}) or large TN penalties γ can over-fragment the residual network, leading to elongated or scattered micro-DMAs. This is a consequence of enforcing balanced partitioning on an SG graph with limited redundancy. These effects are consistent with the framework's stated sensitivity to q, γ, and the presence of small residual components.

To partly address this, for BWSN2, a high TN threshold in the range $q = \{0.94,\ 0.95,\ 0.96, 0.97, 0.98\}$ was used to keep the trunk compact and preserve residual connectivity. The trunk penalty was also reduced to $\gamma = 2$ to avoid suppressing the only available growth pipelines. A substantially larger minimum DMA size $S_{\min} = 350$ nodes is proposed to prevent hydraulically insignificant micro-DMAs. The best configurations for BSWN2 are shown in Figure 7.

The best two configurations obtained both have 22 DMAs, and are consistent with the results of Ferrari *et al.* (2014) on the same network. With considerably fewer DMAs, the method resulted in DMAs of the same structure as those obtained by Diao *et al.* (2013). BWSN2 is hence a particularly suitable case to demonstrate the framework interpretability for a larger network, as it correctly identifies TN and explicitly flags candidate metre/valve pipes and DMAs. In sparse networks where a small number of interfaces control large downstream subgraphs, the collector set is expected to be sparse and operationally meaningful, supporting practical DMA instrumentation and control. The evaluation metrics are presented in Figure 8. As can be observed, the configurations do well on HLG, as a result of the network's sparse topology, and the DSR is around unit as desired, while the complexity of most DMAs approaches zero.

From Figures 5, 6, and 8, we can observe that there is a negative correlation between BCC and the two metrics, namely DSR and HLG, while it is positively correlated with PUI.

The pattern in Figure 9 is consistent with the hydraulic meaning of the proposed four metrics. DMAs with several inlets and interfaces will find it difficult to maintain uniform local pressure transitions due to the large fraction of boundary nodes, which results in nodes at slightly different pressure levels being in the same DMA. This is the cause of the positive correlation between BCC and PUI. However, against practice, the negative correlation between BCC and DSR indicates that

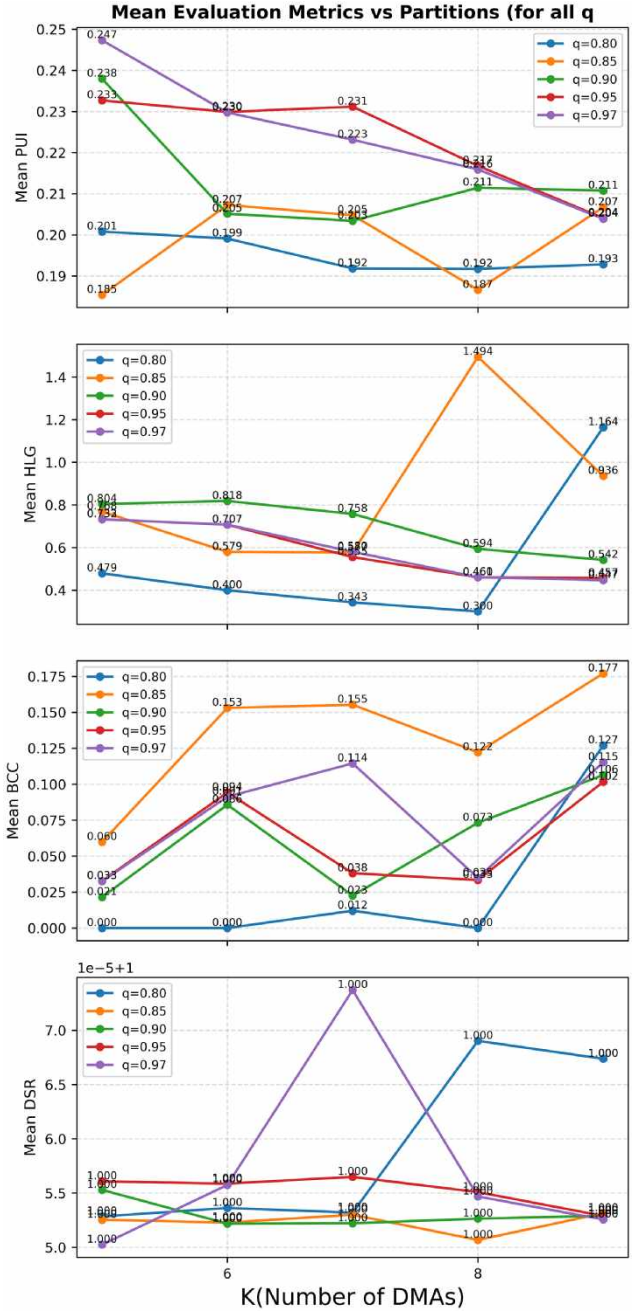


Figure 5 | Evaluation metrics for (q, K) set for D-town network.

configurations with more complex boundaries tend to exhibit slightly lower demand satisfaction. The higher BCC implies DMA exchanges flow with the rest of the WDN through several control interfaces, which makes it easier for small head losses and valve throttling that result in mild deficits inside the DMA, lowering the DSR. BCC is negatively correlated with HLG because high-BCC DMAs in D-town are typically clusters that are located close to the TN, with several multiple supply paths and shorter average path lengths within the DMA. This distributed supply reduces the average head loss gradient per unit length. Low-BCC DMAs, by contrast, tend to be more radial and extend farther from the TN, with longer single-entry control points and longer average supply paths. Overall, Figure 6 confirms that BCC trades off against both DSR and PUI, while also interacting with the spatial position of each DMA to influence the average HLG in a DMA.

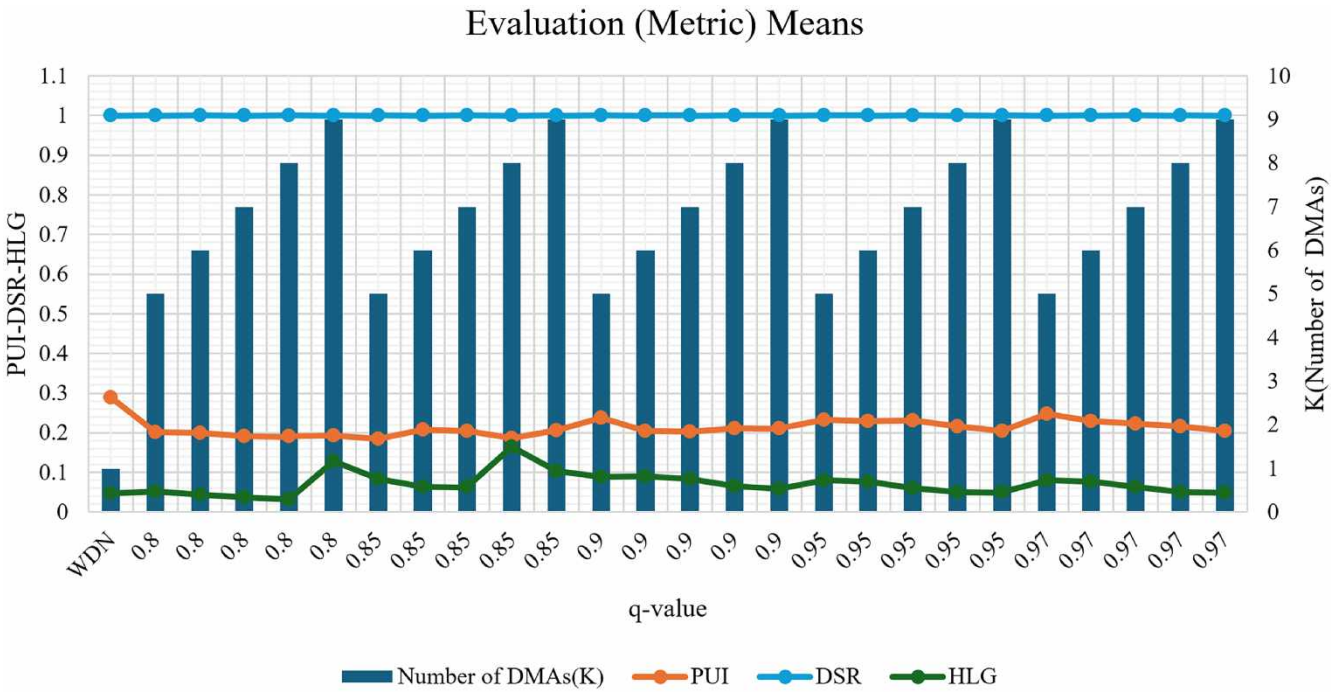


Figure 6 | Evaluation set (metric) means for D-town network.

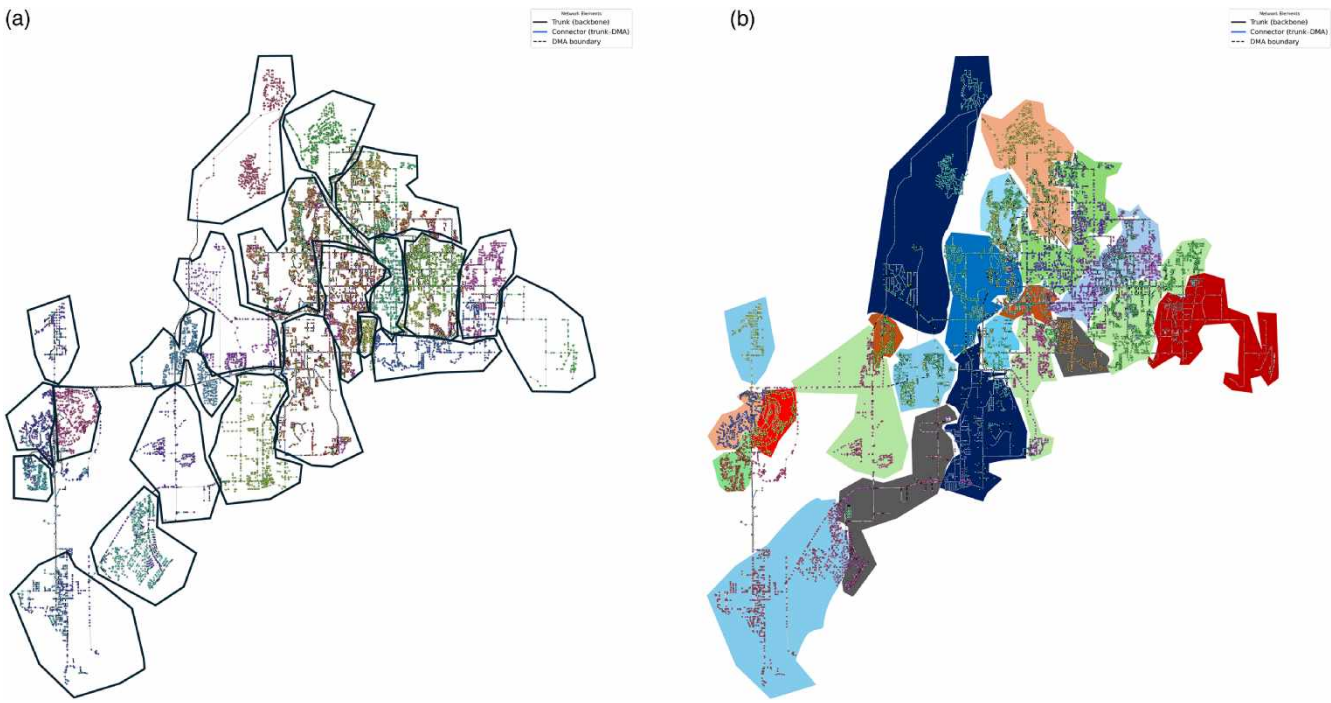


Figure 7 | BWSN2 (a) configuration (i) $q = 0.94$, $k = 22$ and (b) configuration (ii) $q = 0.98$, $k = 22$.

3.5. Operational interpretation of the preferred partition

The selected configuration for D-town ($q = 0.80, K = 7$) and for BWSN2 ($q = 0.98, K = 22$) produced DMAs distributed around the DWEBC-derived TN. Visual inspection of Figures 2(b) and 7 shows that the TN edges form a coherent

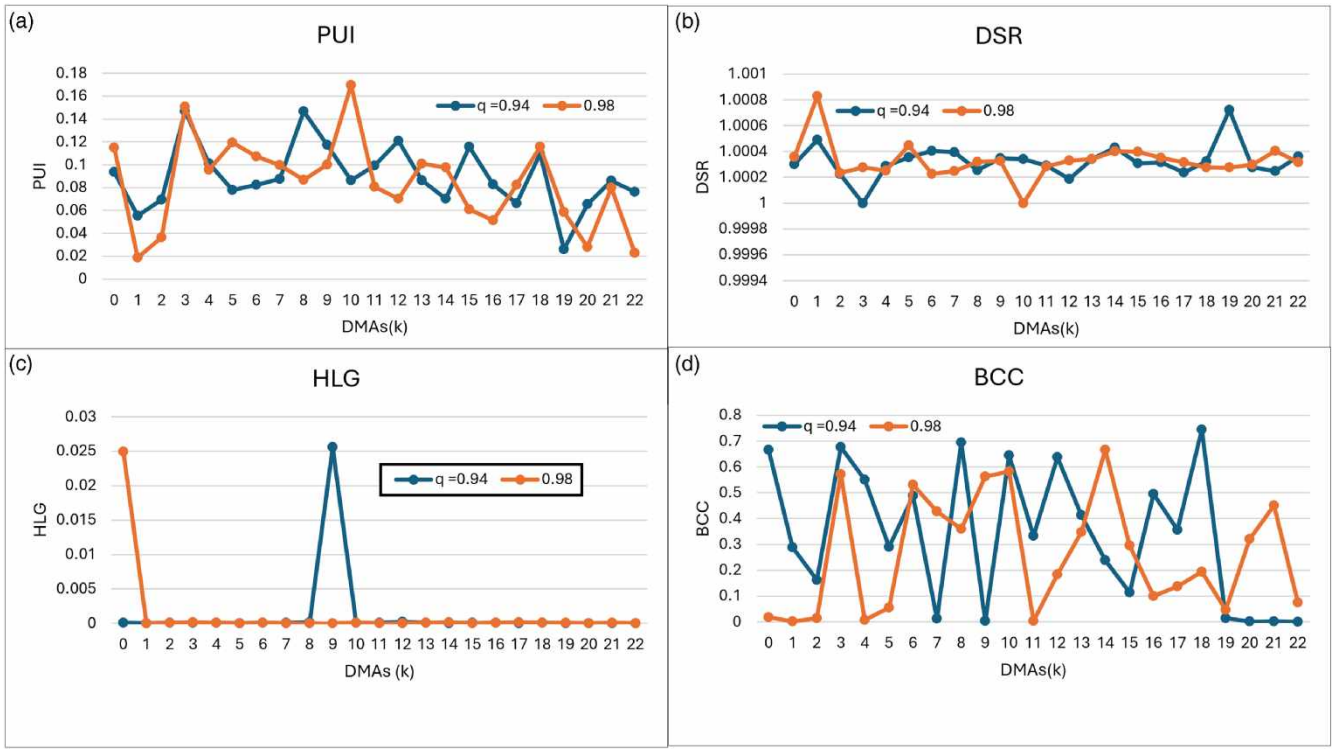


Figure 8 | BWSN2 configuration (i) and (ii) evaluation metric per DMA.

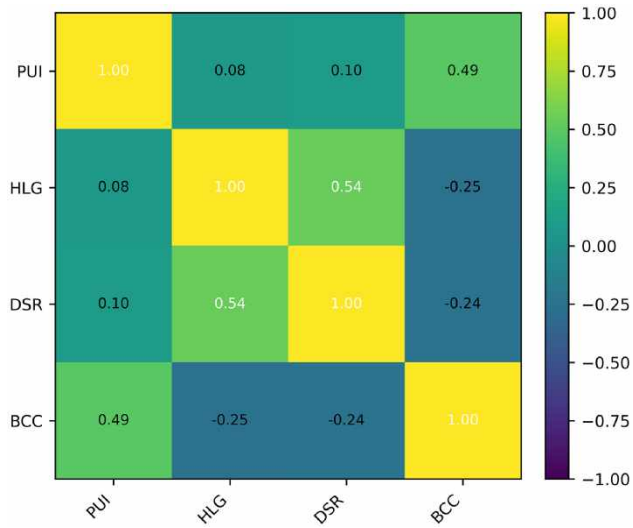


Figure 9 | Evaluation metrics correlation.

supply skeleton connecting the major inflow reservoir to downstream demand clusters, while remaining entirely outside any DMA as preferred. As such, the design choice preserves redundancy and avoids assigning DMA boundaries to pipes whose failure would cause widespread service disruption.

Collector edges, pipes with one node on the TN and the other inside a DMA, are explicitly identified and highlighted. These edges form a sparse and interpretable set of candidate locations for flow metres and isolation valves, which, from an operational perspective, reduce the BCC relative to purely modularity-based community detection, where high-centrality pipes are often cut and boundaries are more convoluted.

The PUI and HLG values indicate that the DMAs follow hydraulically coherent corridors, with relatively smooth pressure and head-loss profiles. This is a direct consequence of using hydraulic weight, derived from the DWEBC weights, rather than using purely topological distance in the k-centre growth. Meanwhile, the BCC metric result confirms that the resulting boundaries remain manageable for dynamic reconfiguration, which is critical for modern, dfc DMA requirements.

As alluded to, there are a few residual isolated nodes that fall outside the main DMAs and require manual reassignment for the BWSN2, which illustrates the method's transparency, as they are easy to visually diagnose and can be resolved with minimal local adjustments, without affecting the global trunk-preserving design of the DMAs.

3.6. Sensitivity analysis of uniformity parameters β_{lo} and β_{hi}

In the proposed trunk-preserving DMA partitioning framework, the parameters β_{lo} and β_{hi} (Equation (6)) control the admissible size range of each DMA during the uniformity balancing stage of the procedure. These parameters do not affect TN identification or DWEBC computation but influence the reassignment of nodes during the balancing phase.

A sensitivity analysis was conducted on the D-town network to evaluate the influence of $\beta_{lo}$and β_{hi} on the partitioning results. The backbone quantile was fixed at $q = 0.90$ and the target number of DMAs at $K_{\text{target}} = 12$. The parameters were varied over the ranges:

$$\beta_{lo} \in \{0.60, 0.70, 0.80, 0.90\},\ \beta_{hi} \in \{1.00, 1.10, 1.20, 1.30\},$$

subject to $\beta_{hi} \geq \beta_{lo}$. For each parameter combination, the algorithm was executed using identical hydraulic simulations and random seeds to isolate the effect of the uniformity bounds.

The actual number of DMAs K_{actual}, deviation $K_{\text{diff}} = K_{\text{actual}} - K_{\text{target}}$, mean BCC, mean (HLG), mean PUI, mean DSR, and coefficient of variation of DMA sizes (CV_{size}) were recorded for the D-town network.

3.6.1. Stability of partitions and service performance

Figures 10(a) and 10(b) show the heatmaps of K_{actual} and K_{diff} over the (β_{lo}, β_{hi}) grid. For all tested combinations, the algorithm consistently recovered $K_{\text{actual}} = 12$, yielding $K_{\text{diff}} = 0$. This demonstrates that the uniformity parameters do not compromise the feasibility or stability of the partitioning process.

Similarly, the mean DSR remained effectively equal to unity across the entire parameter space (Figure 11), indicating that variations in DMA size uniformity do not affect demand satisfaction or hydraulic feasibility. These results confirm that β_{lo} and β_{hi} act as structural tuning parameters rather than constraints that influence supply performance.

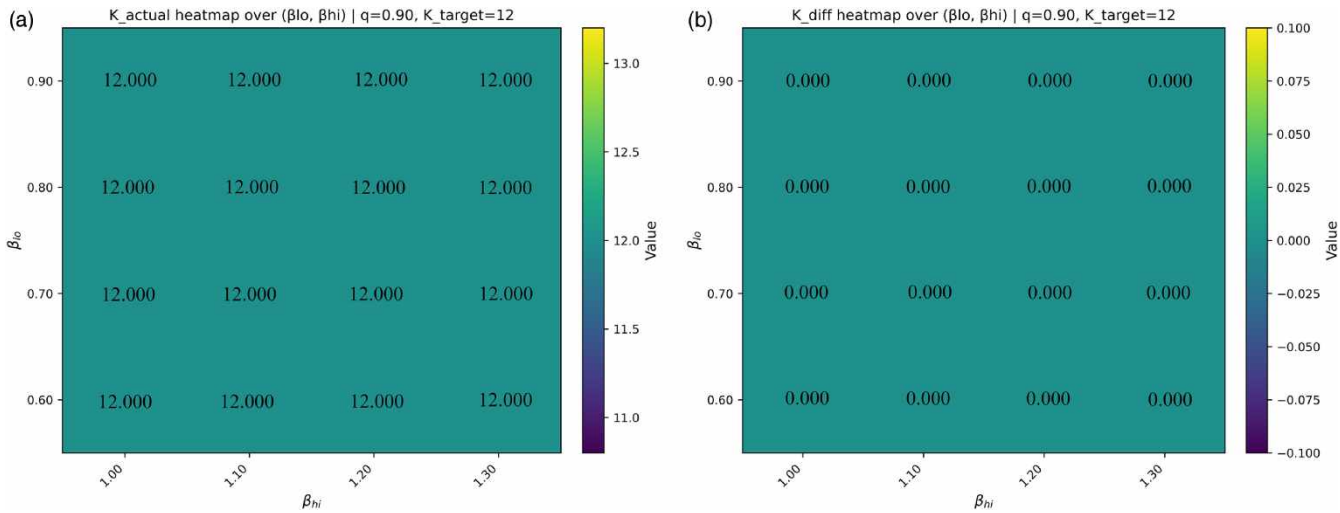


Figure 10 | (a) Actual partitions (k_{actual}) heatmap and (b) difference in partitions realised heatmap over (β_{lo}, β_{hi}).

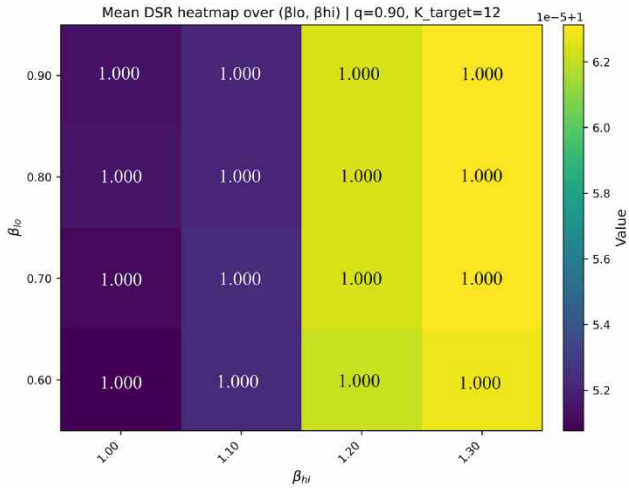


Figure 11 | Mean DSR over (β_{lo}, β_{hi}).

3.6.2. Influence on hydraulic coherence and boundary connectivity

In contrast, Figures 12(a) and 12(b) reveal a sensitivity of internal DMA structure to the choice of β_{lo} and β_{hi}. The mean HLG decreases systematically as the admissible size range widens (lower β_{lo}, higher β_{hi}), indicating improved hydraulic coherence within DMAs. This suggests that relaxed uniformity constraints allow partitions to align more closely with hydraulics, reducing internal head loss gradients.

The effect is even more pronounced for boundary connectivity. As shown in Figure 12(ii), tight uniformity bands (low β_{hi}) lead to significantly higher BCC values, reflecting an increase in boundary nodes. Conversely, increasing β_{hi} to 1.20–1.30 reduces BCC, particularly when combined with lower β_{lo}. This indicates that excessive enforcement of size uniformity forces non-hydraulic node reassignments, degrading DMA separability.

3.6.3. Trade-off with DMA size uniformity

Figure 13(a) illustrates the expected increase in CV_{size} as β_{hi} increases and β_{lo} decreases. While wider bounds reduce boundary connectivity and hydraulic gradients, they also permit greater size heterogeneity among DMAs. This confirms that β_{lo} and β_{hi} effectively control size dispersion but highlights a clear trade-off between strict uniformity and hydraulic/topological optimality.

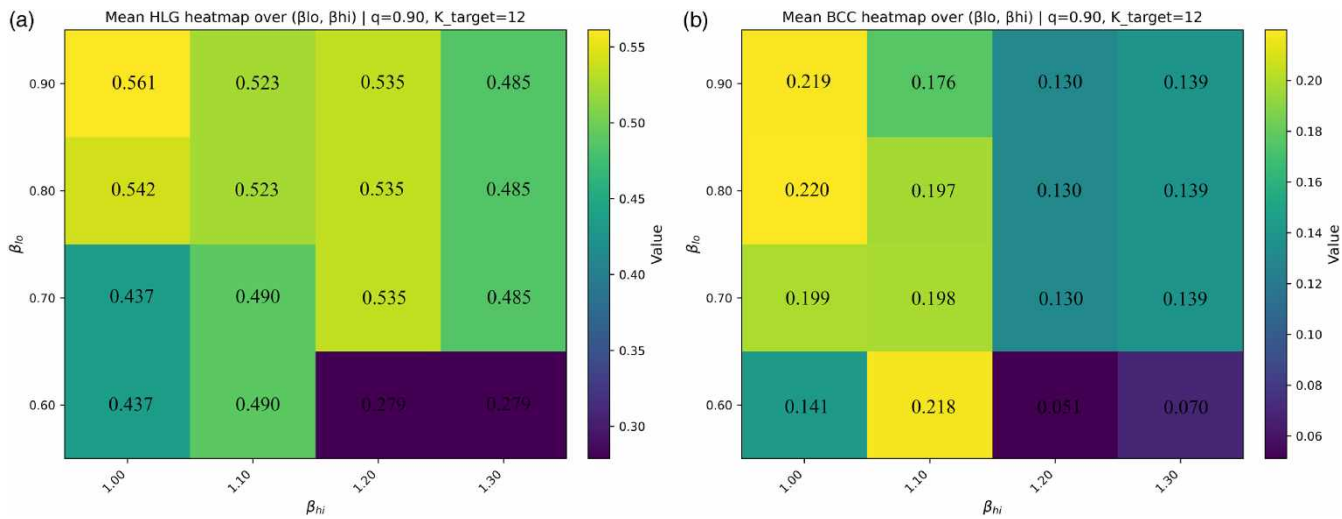


Figure 12 | (a) Mean HLG over (β_{lo}, β_{hi}) and (b) mean BCC over (β_{lo}, β_{hi}).

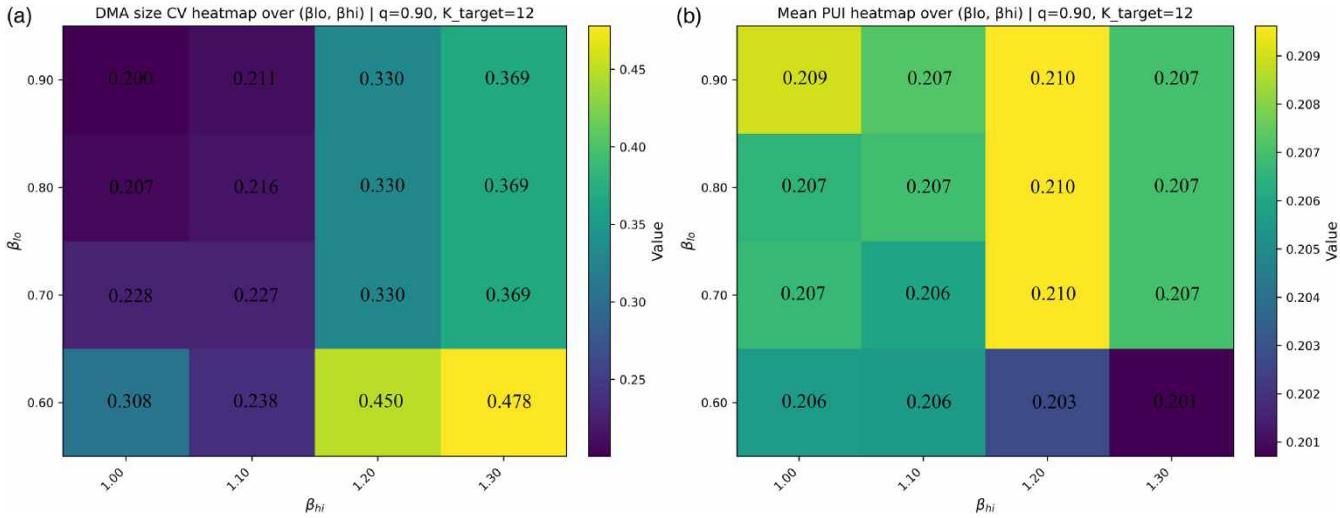


Figure 13 | (a) DMA size CV heatmap and (b) mean PUI heatmap over (β_{lo}, β_{hi}).

Figure 13(b) shows that uniformity constraints should not be tuned for pressure reasons, as PUI is weakly sensitive to (β_{lo}, β_{hi}).

3.6.4. Recommended parameter range

The sensitivity results indicate a Pareto-optimal region around $\beta_{lo} \approx 0.60–0.90$, $\beta_{hi} \approx 1.20$, which preserves the target number of DMAs, maintains full demand satisfaction, minimises boundary connectivity, and improves hydraulic coherence while keeping size variability within acceptable limits. These findings support the treatment of uniformity constraints as *soft bounds*, rather than strict requirements, within the trunk-preserving DMA formulation process.

We followed similar steps in the Exnet case study to identify the TN. The TNs were obtained by thresholding the DWEBC distribution at quantiles, $q \in \{0.75,\ 08,\ 0.85,\ 0.9,\ 0.95\}$, and the DMAs were obtained for $K \in \{20,\ 22,\ 25,\ 27,\ 30\}$. An analysis of the resulting visual network outputs and the metrics led to two configurations being identified with desirable metrics results, namely configuration (i) $q = 0.85$ and $k = 22$ and configuration (ii) $q = 0.95$ and $k = 25$, as shown in Figure 14.

The results show that the framework can also be used for a network with a number of pipes in the range of thousands and still be tractable.

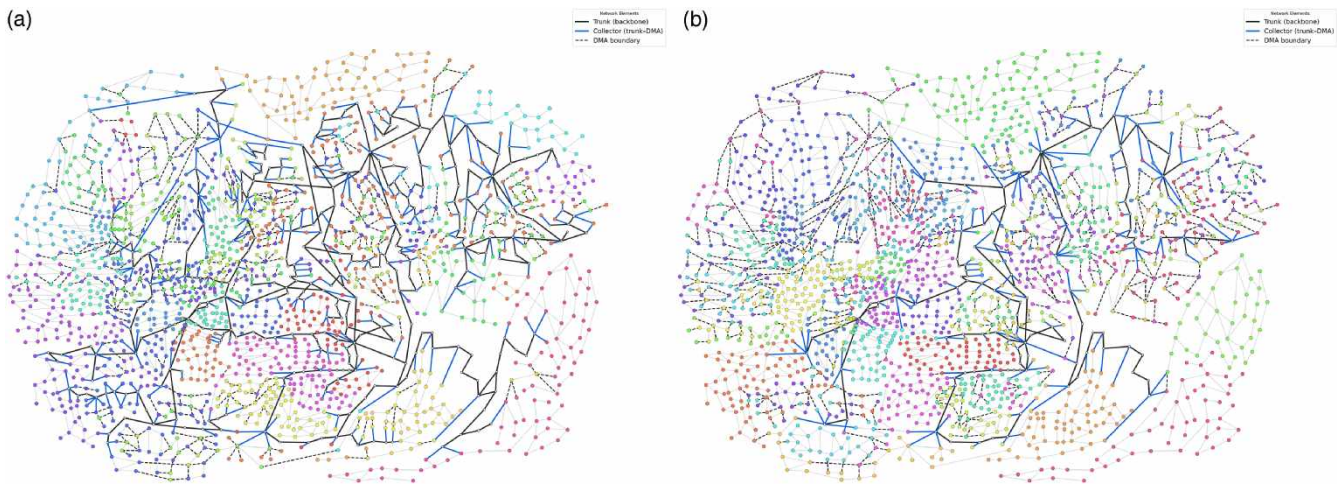


Figure 14 | Exnet (a) configuration $q = 0.85$, $k = 22$ and configuration (b) $q = 0.95$, $k = 25$.

3.7. Limitations

The method is greedy; hence, it trades optimality for robustness and speed. Where planning requires explicit Pareto exploration of [PUI, HLG, 1 – DSR, BCC], the uniformity-balanced partition can serve as a first step in a multi-objective optimiser that swaps boundary nodes or merges/splits DMAs under the same trunk constraint and distance/cut penalties to improve the evaluation metrics.

The study also relies on steady-hourly snapshots (24 h) without transient hydraulics, such that extreme events are not explicitly modelled. In addition, TN threshold q and penalties γ, η require sensitivity analysis; very small residual components constrain the achievable K. Increasing the trunk penalty γ further discourages cross-trunk growth and tightening (β_{lo}, β_{hi}) narrows the DMA size distribution, while adjusting q changes the extent of the TN and hence the degrees of freedom for the residual graph partitioning.

4. CONCLUSIONS

In this paper, a trunk-preserving, hydraulically informed partitioning framework for WDNs that combines a time-robust DWEBC metric with capacitated graph-Voronoi growth has been introduced. Compared with purely topological community detection, this method explicitly keeps a hydraulically significant TN unassigned to any DMA, grows DMAs on the residual graph using a hydraulic weight derived from DWEBC, and repairs connectivity and size imbalances through bounded refinement.

The methodology was demonstrated on the D-town and BWSN2 as the main case studies and EXNET benchmark network for TN identification, under PDD conditions, in which the framework consistently identified TNs that align with major supply corridors on all networks and generated connected DMAs with controlled size dispersion and limited boundary complexity on the D-town. We considered several DWEBC thresholds and target DMA counts. The evaluation metrics highlighted preferred configurations at $q = 0.85$, $K = 22$ for the Exnet case study, $q = 0.80$, $K = 7$ for the D-town case study, and $q = 0.98$, $K = 22$ for the BWSN2. These configurations offer a favourable balance of PUI, HLG, DSR, and BCC. The identification of collector edges at TN–DMA interfaces provides a set of candidate locations for metering and isolation, thereby supporting dfc and dynamic reconfiguration approaches.

The framework remains computationally tractable as it relies on single-source shortest-path computations and bounded refinement loops, without heavy multilevel coarsening. This makes it compatible with existing EPANET/WNTR workflows and suitable for practical planning studies in medium-sized WDNs. For both case studies, the dominant computational cost arises from the hydraulic simulation and the repeated shortest-path computations required for DWEBC evaluation. Using a 24-h EPS under PDA, the full DWEBC computation scales linearly with the number of simulated timesteps and network size, as it relies on repeated single-source shortest-path calculations. In practice, the complete DWEBC evaluation and TN extraction were completed in less than 120 seconds and 1,935 seconds for Exnet and BWSN2, respectively, on a laptop equipped with an Intel 12th Gen Intel® Core™ i7-12700H 2.60 GHz processor, 32 GB RAM, and NVIDIA GeForce RTX 3060 6 GB. The subsequent partitioning stage is comparatively lightweight because the partitioning and the bounded refinement loops operate on the residual network and do not require repeated precomputation of the DWEBC. In addition, the refinement iterations are explicitly controlled, and the runtime remains predictable and avoids the convergence issues sometimes encountered in modularity-based or multilevel clustering methods. Overall, the framework exhibits good scalability with respect to both network size and the number of target DMAs, and its modular structure allows straightforward comparison across timesteps or across candidate (q, K) configurations. This makes the approach compatible with scenario analysis, sensitivity studies, and integration into digital-twin environments where repeated re-partitioning may be required.

As discussed in the limitations, the approach trades global optimality for speed and robustness and currently relies on steady-hourly simulations; nevertheless, it offers a reproducible starting point that can be systematically improved.

Future work can focus on integrating the trunk-preserving partition as a first step in a multiphase multi-objective optimisation to enable explicit Pareto exploration of trade-offs among PUI, HLG, DSR (FDI), and BCC, as well as economic indicators such as valve and metre installation costs. Extending the DWEBC formulation to account for transient events and real-time telemetry would allow the method to support digital-twin-based emergency response and adaptive pressure management, which is an area worth exploring as well. Additionally, exploiting the identified collector edges and DMA

boundaries for sensor placement, leak localisation, and resilience enhancement represents a promising avenue for integrating graph-theoretic partitioning with operational decision support.

Overall, the proposed framework provides a hydraulically meaningful and operationally interpretable basis for the design and management of DMAs and can be readily transferred to real-world WDNs.

FUNDING

The work has been supported by the National Key R&D Program of China (grant number 2022YFC3203804), the State Key Laboratory of Urban Water Resource and Environment (Harbin Institute of Technology) (grant number 2024TS26), and the Key Research and Development Program of Heilongjiang Province of China (grant number 2022ZX01A06).

DATA AVAILABILITY STATEMENT

All relevant data are included in the paper or its Supplementary Information.

CONFLICT OF INTEREST

The authors declare there is no conflict.

REFERENCES

Alvisi, S. & Franchini, M. (2014) A procedure for the design of district metered areas in water distribution systems, *Procedia Engineering*, **70**, 41–50.

Bergamini, E. & Meyerhenke, H. (2016) Approximating betweenness centrality in fully dynamic networks, *Internet Mathematics*, **12**, 281–314.

Brandes, U. (2008) On variants of shortest-path betweenness centrality and their generic computation, *Social Networks*, **30**, 136–145.

Campbell, E., Izquierdo, J., Montalvo, I., Ilaya-Ayza, A., Pérez-García, R. & Tavera, M. (2015) A flexible methodology to sectorize water supply networks based on social network theory concepts and multi-objective optimization, *Journal of Hydroinformatics*, **18**, 62–76.

Creaco, E., Giudicianni, C. & Mottahedin, A. (2023) Improved community detection for WDN partitioning in the dual topology based on segments and valves, *Journal of Hydroinformatics*, **25**, 1341–1357.

Diao, K., Zhou, Y. & Rauch, W. (2013) Automated creation of district metered area boundaries in water distribution systems, *Journal of Water Resources Planning and Management*, **139**, 184–190.

Diao, K., Fu, G., Farmani, R., Guidolin, M. & Butler, D. (2016) Twin-hierarchy decomposition for optimal design of water distribution systems, *Journal of Water Resources Planning and Management*, **142**, C4015008.

Di Nardo, A., Di Natale, M., Santonastaso, G. F., Tzatchkov, V. G. & Alcocer-Yamanaka, V. H. (2014) Water network sectorization based on graph theory and energy performance indices, *Journal of Water Resources Planning and Management*, **140**, 620–629.

Di Nardo, A., Di Natale, M., Giudicianni, C., Greco, R. & Santonastaso, G. F. (2017) Weighted spectral clustering for water distribution network partitioning, *Applied Network Science*, **2**, 19.

Dongarra, J., Foster, I., Fox, G., Gropp, W., Kennedy, K., Torczon, L. & White, A. (2003) *Sourcebook of Parallel Computing*. San Francisco, CA: Morgan Kaufmann Publishers Inc.

Ferrari, G., Savic, D. & Becciu, G. (2014) Graph-Theoretic approach and sound engineering principles for design of district metered areas, *Journal of Water Resources Planning and Management*, **140**, 04014036.

Giudicianni, C., Herrera, M., Di Nardo, A. & Adeyeye, K. (2020) Automatic multiscale approach for water networks partitioning into dynamic district metered areas, *Water Resources Management*, **34**, 835–848.

Giustolisi, O. & Ridolfi, L. (2014) A novel infrastructure modularity index for the segmentation of water distribution networks, *Water Resources Research*, **50**, 7648–7661.

Gomes, R., Marques, A. S. & Sousa, J. (2012) Decision support system to divide a large network into suitable District Metered Areas, *Water Science and Technology*, **65**, 1667–1675.

Gonzalez, T. F. (1985) Clustering to minimize the maximum intercluster distance, *Theoretical Computer Science*, **38**, 293–306.

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

Khoa Bui, X., Marlim, M. S. & Kang, D. (2020) Water network partitioning into district metered areas: a state-of-the-art review, *Water*, **12**, 1002.

Liu, H., Zhao, M., Zhang, C. & Fu, G. (2018) Comparing topological partitioning methods for district metered areas in the water distribution network, *Water*, **10**, 368.

Marchi, A. (2021) *05 Long Term Improvement*. Available at: https://uknowledge.uky.edu/wdst_models/5.

Morrison, J., Stephen, T. & Roger, D. (2007) *District Metered Areas: Guidance Notes. Technical Report 1*. London, UK: International Water Association (IWA).

Newman, M. E. J. (2005) A measure of betweenness centrality based on random walks, *Social Networks*, **27**, 39–54.

Ostfeld, A., Uber, J. G., Salomons, E., Berry, J. W., Hart, W. E., Phillips, C. A., Watson, J.-P., Dorini, G., Jonkergouw, P., Kapelan, Z., Pierro, F. D., Khu, S.-T., Savic, D., Eliades, D., Polycarpou, M., Ghimire, S. R., Barkdoll, B. D., Gueli, R., Huang, J. J., Mcbean, E. A., James, W., Krause, A., Leskovec, J., Isovitsch, S., Xu, J., Guestrin, C., Vanbriesen, J., Small, M., Fischbeck, P., Preis, A., Propato, M., Piller, O., Trachtman, G. B., Wu, Z. Y. & Walski, T. (2008) The Battle of the Water Sensor Networks (BWSN): a design challenge for engineers and algorithms, *Journal of Water Resources Planning and Management*, **134**, 556–568.

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

Perelman, L. & Ostfeld, A. (2011) Topological clustering for water distribution systems analysis, *Environmental Modelling & Software*, **26**, 969–972.

Pirouz, M. (2021) Balanced weighted label propagation. In: Arai, K. (ed.) *Intelligent Computing*, Cham: Springer International Publishing, pp. 1–12.

Savić, D. & Ferrari, G. (2014) Design and performance of district metering areas in water distribution systems, *Procedia Engineering*, **89**, 1136–1143.

Schaeffer, S. E. (2007) Graph clustering, *Computer Science Review*, **1**, 27–64.

Sharma, A. N., Dongre, S. R., Gupta, R. & Ormsbee, L. (2022) Multiphase procedure for identifying district metered areas in water distribution networks using community detection, NSGA-III optimization, and multiple attribute decision making, *Journal of Water Resources Planning and Management*, **148**, 04022040.

Shekofteh, M. R., Yousefi-Khoshqalb, E. & Piratla, K. R. (2023) An efficient approach for partitioning water distribution networks using multi-Objective optimization and graph theory, *Water Resources Management*, **37**, 5007–5022.

Tzatchkov, V. G. & Alcocer-Yamanaka, V. H. (2019) Graph theory based single and multiple source water distribution network partitioning, *Instituto Mexicano de Tecnología del Agua*, **10**, 197–221.

Tzatchkov, V. G., Alcocer-Yamanaka, V. H. & Ortíz, V. B. (2012) Graph theory based algorithms for water distribution network sectorization projects, *Water Distribution Systems Analysis Symposium*, **2006**, 1–15.

Ugurlu, O. (2022) Comparative analysis of centrality measures for identifying critical nodes in complex networks, *Journal of Computational Science*, **62**, 101738.

Ulusoy, A.-J., Stoianov, I. & Chazerain, A. (2018) Hydraulically informed graph theoretic measure of link criticality for the resilience analysis of water distribution networks, *Applied Network Science*, **3**, 31.

Ulusoy, A.-J., Mahmoud, H. A., Pecci, F., Keedwell, E. C. & Stoianov, I. (2022) Bi-objective design-for-control for improving the pressure management and resilience of water distribution networks, *Water Research*, **222**, 118914.

Wright, R. J. (2016) *Water Distribution Networks with Dynamic Topology.* Doctor of Philosophy, Imperial College London.

Wright, R., Stoianov, I., Parpas, P., Henderson, K. & King, J. (2014) Adaptive water distribution networks with dynamically reconfigurable topology, *Journal of Hydroinformatics*, **16**, 1280–1301.

Yu, X., Wu, Y., Meng, F., Zhou, X., Liu, S., Huang, Y. & Wu, X. (2024) A review of graph and complex network theory in water distribution networks: mathematical foundation, application and prospects, *Water Research*, **253**, 121238.

Zarghami, S. A., Gunawan, I. & Schultmann, F. (2018). 'Vulnerability Analysis of Water Distribution Networks Using Betweenness Centrality and Information Entropy', *2nd IEOM European Conference on Industrial Engineering and Operations Management*. Paris, France, pp. 1811–1820.

First received 10 December 2025; accepted in revised form 2 March 2026. Available online 23 March 2026