Application of Fuzzy Cluster Analysis and Decision Making Algorithms for Choosing the Optimal Water Distribution Network Design

Vasan Arunachalam (✉ vasan@hyderabad.bits-pilani.ac.in)

BITS Pilani, Hyderabad Campus <https://orcid.org/0000-0001-7871-2831>

K Srinivasa Raju

Birla Institute of Technology and Science - Hyderabad Campus

M Naveen Naidu

Birla Institute of Technology and Science - Hyderabad Campus

Research Article

Keywords: DCVI, Decision Making, Fuzzy Cluster Analysis, NSGA-II, WDN

Posted Date: May 13th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-309401/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Application of Fuzzy Cluster Analysis and Decision Making Algorithms for Choosing the Optimal Water Distribution Network Design

A Vasana*, K Srinivasa Raju, M Naveen Naidu

Department of Civil Engineering, BITS Pilani, Hyderabad Campus, India

Email: vasan@hyderabad.bits-pilani.ac.in, ksraju@hyderabad.bits-pilani.ac.in,

p20170009@hyderabad.bits-pilani.ac.in

* Corresponding Author

Abstract

Non-dominated Sorting Genetic Algorithm (NSGA-II) is employed to facilitate multi-objective optimization in Water Distribution Network(s) (WDN) framework for a benchmark problem of Hanoi Network and a real-world problem, Pamapur Network, Telangana, India. Maximization of resilience, minimization of cost and minimization of leakages are considered in a multiobjective context which result in generation of Non-dominated WDN Strategies (NWDNS). In order to simplify the decision making process of engineers, Fuzzy Cluster Analysis (FCA) is employed to categorize NWDNS into groups. Thereafter, Dunn's Cluster Validity Index (DCVI) is used for identification of optimal cluster size. Representative NWDNS i.e. RNWDNS for each sub-cluster is based on the maximum membership of NWDNS in the respective sub-cluster. Ranking of RNWDNS is performed with three decision making algorithms, namely, Preference Ranking Organization METHOD for Enrichment of Evaluations-2 (PROMETHEE-2), Multicriterion Q-analysis-2 (MCQA-2) and VlseKriterijumska Optimizacija I Kompromisno Resenje (VIKOR). Additive ranking rule is also applied to facilitate obtained ranks in group decision making environment to arrive

at the optimal WDN. It is observed that 1020 NWDNS generated for Hanoi network are optimally classified into 18 clusters based on DCVI, and A13 representing RNWDNS 37 is found preferable. Whereas 272 NWDNS generated for Pamapur network are classified into 9 clusters where S6 is preferred (representing RNWDNS 203).

Keywords: DCVI, Decision Making, Fuzzy Cluster Analysis, NSGA-II, WDN

Introduction

Water Distribution Network(s) (WDN) need a structured approach to accommodate ever increasing demand in a sustainable manner. The design, construction, operation and maintenance of WDN create enormous challenge for engineers around the globe and require huge amount of resources to address the same. Therefore, a need is felt to enhance their effectiveness by minimizing leak and cost and simultaneously maximizing resilience. However, optimization of WDN is facing criticism because of 1) complex non-linear association linked to head loss and discharge 2) combinatorial nature that arises due to incorporation of discrete pipe diameters available in the market into modeling which leads to reduced quality of optimal solution. This necessitates robust modelling and filtering mechanisms. Accordingly, multiobjective optimization, clustering and decision making algorithms are employed in the present study due to their potential applicability in WDN (Bui *et al.* 2020). Brief but relevant literature reviews are as follows:

Prasad and Park (2004) applied NSGA-II to Hanoi network and two loop network with orientation towards network cost and reliability. Kapelan *et al.* (2005) employed RNSGA-II with design cost and robustness as objectives to New York tunnels reinforcement problem. Uncertainties in pipe roughness coefficients and future water consumption are considered. RNSGA-II identified robust solutions in spite of less computational time. Farmani *et al.*

(2006) considered minimization of total cost, maximization of minimum resilience index and minimization of maximum water age to Anytown's WDN in NSGA-II framework. Vasan and Simonovic (2010) combined differential evolution and EPANET with orientation towards minimization of cost. They applied proposed methodology to New York water supply system and Hanoi networks. Montalvo *et al.* (2010) employed improved Particle Swarm Optimization (PSO) to Hanoi network and a real world problem of Lima, Peru WDN to facilitate multiobjective optimization.

Fu and Kapelan (2011) proposed fuzzy probabilistic mechanism for optimal design as well as rehabilitation of WDN. System performance and design cost are considered and applied to Hanoi network and New York tunnels. They also employed NSGA-II along with other mathematical approaches. Wu *et al.* (2013) considered hydraulic reliability, economic cost, and greenhouse emissions for three studies (two benchmark and one case study in Australia).

Water System Multiobjective Genetic Algorithm (WSMGA) and enumeration schemes are applied. Marques *et al.* (2018) proposed hydraulic capacity, pressure deficit and undelivered demand, carbon emissions, and traditional cost for hypothetical WDN to develop flexible solutions employing a multi-stage approach. They also considered uncertainty in demand and employed enhanced simulated annealing in multiobjective context to solve multiple lacuna that can come in WDNs. Wang *et al.* (2019) studied interrelationships among the five key parameters of NSGA-II resulting 32 parameter combinations. Hanoi, New York tunnel and real world problem, Balerna irrigation networks are considered with focus on minimum cost and minimum head deficit. It is concluded that the impact of parameters is not equal on the outcome. Patil *et al.* (2020) developed augmented multiobjective PSO to five benchmark problems of WDN and compared with other approaches. They suggested a new criterion for comparing the efficacy of pareto fronts. Pankaj *et al.* (2020) applied self-adaptive cuckoo

search optimization technique to design WDN in single objective and multiobjective scenario. Two loop, Hanoi network and Pamapur network, India are considered. It is concluded that the present work is found to be effective, as compared to previous research works. Hu *et al.* (2020) applied modified NSGA-III to WDNs from the repository of ‘battle of the water sensor networks’ for water quality sensor placement.

Chang *et al.* (2011) employed (a) K-means, (b) fuzzy c-mean and (c) self-organizing map (SOM) to group large WDNs. They considered a number of water quality indicators and employed Davies Bouldin Index. It is concluded that SOM is useful due to its ability for comprehensive analysis of WDNs. Liu and Han (2018) employed integrated methodology of spectral clustering and graph theory for design and analysis of district metered area (DMA) in simple network and Wolf-Cordera Ranch WDN. Genetic algorithm, in association with K-means, is used to enhance quality of clustering. They also employed multicriterion decision analysis to compare different solutions. Novarini *et al.* (2019) partitioned DMAs using hybrid model and K-means for D-Town network. They considered mathematical and topological criteria. PSO was applied to optimally locate DMA of clusters. It is concluded that K-means performed well and Calinski-Harabasz is found to be a promising approach for finding optimum clusters. Similar studies on classification are reported by Kirstein *et al.* (2015) and Lifshitz and Ostfeld (2018).

Gheisi and Naser (2015) applied weighted sum model, Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS) and weighted product model to rank twenty-two WDN design layouts for a hypothetical WDN. Salehi *et al.* (2018) assessed water supply failure risks in the WDN and employed fuzzy TOPSIS for rehabilitation needs of Qods Town, Iran. Cunha *et al.* (2019) employed Preference Ranking Organization METHod for Enrichment of Evaluation (PROMETHEE) and TOPSIS for Hanoi. Twenty network design alternatives

were evaluated on various criteria. Cunha *et al.* (2020) applied PROMETHEE and Simple Additive Weighting methods to New York tunnels. Ten design alternatives are proposed with four criteria. Brentan *et al.* (2021) employed NSGA-II for multiobjective optimization and ELECTRE-TRI to case study of D-town. Gaur *et al.* (2021) applied Analytic Element Method (AEM)-NSGA-II to part of the Dore river catchment, France in groundwater aspects. K-Means, Davies Bouldin Index, TOPSIS and VIKOR were other employed techniques to arrive at suitable groundwater strategy.

Limited studies are noticed on classifying generated pareto front of WDN strategies and related ranking. Accordingly, following objectives/methodology is formulated for the Hanoi Network & Pamapur Network focusing on:

1. To generate Non-dominated WDN Strategies (NWDNS) based on goals, namely, maximization of resilience, minimization of cost, and minimization of leakages with NSGA-II based multiobjective optimization
2. To categorize the generated NWDNS with Fuzzy Cluster Analysis (FCA)
3. To find the optimum cluster size with Dunn's Cluster Validation Index (DCVI)
4. To identify representative NWDNS (from now termed as RNWDNS) of each cluster based on maximum membership grade
5. To rank RNWDNS with decision making algorithms, PROMETHEE-2, Multicriterion Q-analysis-2 (MCQA-2) and VlseKriterijumska Optimizacija I Kompromisno Resenje (VIKOR)
6. To suggest suitable RNWDNS in additive ranking framework of group decision making

Details about NSGA-II, FCA, DCVI, PROMETHEE-2, MCQA-2 and VIKOR are presented in the next section.

Description of Methods

NSGA-II proposed by Deb *et al.* (2002) is a robust multiobjective optimization method which facilitates a trade-off among the number of objectives of any formulated mathematical model. The algorithm uses a fast, non-dominated sorting strategy to prioritize solutions, followed by selecting population members of the next generation based on pareto dominance and crowding distance. The main operators of NSGA-II are (a) Simulated Binary Crossover method to enable crossover between solutions, (b) a polynomial mutation approach to maintain population diversity and (c) a constraint-handling strategy to account for infeasible solutions. Detailed information on NSGA-II can be found in Deb *et al.* (2002).

FCA is a categorization mechanism in uncertain environment where each NWDNS can belong to a chosen cluster with certain degree of membership (between zero and one) and aggregation of values of membership for each NWDNS adds up to 1 (Raju and Nagesh Kumar, 2007; Ross, 2016). Brief methodological flow of FCA is presented in Fig.1.

Fig. 1 Flow chart of FCA methodology

DCVI (Dunn 1974) is a measure that minimizes intra-cluster distances $\Delta(X_i)$ and simultaneously maximizes inter-cluster distances $\delta(X_i, X_j)$ where X_i and X_j are clusters. The optimal number of clusters that maximizes DCVI is the decisive factor. Mathematically, it can be shown as:

$$DCVI(U) = \min_{1 \leq i \leq K} \left\{ \min_{\substack{1 \leq j \leq K \\ j \neq i}} \left[\frac{\delta(X_i, X_j)}{\max_{1 \leq k \leq K} (\Delta(X_k))} \right] \right\} \quad (1)$$

PROMETHEE-2 (Brans *et al.* 1986) is on the basis of outranking. Methodology includes computation of preference functions P_j between two RNWDNS a and b which build on the distance from a to b for selected feature and criterion function. Mathematical procedure of PROMETHEE-2 is as follows:

$$1. \text{ Multi Feature Preference Index (MFPI), } \pi(a, b) = \frac{\sum_{j=1}^J w_j P_j(a, b)}{\sum_{j=1}^J w_j}; w_j \text{ is weight of the}$$

$$j^{\text{th}} \text{ criterion} \quad (2)$$

$$2. \text{ Outranking index of RNWDNS } a \text{ in the N set, } \phi^+(a) = \frac{\sum \pi(a, b)}{(N-1)}; N \text{ are number of}$$

$$\text{RNWDNS}] \quad (3)$$

$$3. \text{ Outranked index of RNWDNS } a \text{ in the N set, } \phi^-(a) = \frac{\sum \pi(b, a)}{(N-1)} \quad (4)$$

$$4. \text{ Net ranking of } a \text{ is } \phi(a) = \phi^+(a) - \phi^-(a) \quad (5)$$

5. Higher $\phi(a)$ indicates suitable RNWDNS.

MCQA-2 (Duckstein and Nobe 1997) is based on value and outranking type indices. Methodology includes formulation of preference matrix, incidence matrix, computation of shared face matrix, indices related to project satisfaction, project concordance, and project discordance (PSI, PCI, PDI). Normalization of indices is used to compute L_p - metric values for the purpose of ranking. Lower L_p - metric value indicates suitable RNWDNS. Fig. 2 presents flowchart of MCQA-2.

Fig. 2 Flow chart of MCQA-2 Methodology

VIKOR is based on the distance concept, which uses utility and regret measures (Opricovic and Tzeng 2004; Raju *et al.* 2018). S_i and R_i represent utility and regret measures. V and

(1- ν) are weights of strategy of ‘majority of criteria’ and individual regret function.

Mathematical progression is as follows:

1. Estimate the values S_i and R_i $i=1,2,...,N$ $\left[S_i = \sum_{j=1}^J w_j \left| \frac{f_j^* - f_j(a)}{M_j - m_j} \right| \right]; R_i = \text{Max}[S_i];$

f_j^* , M_j , m_j = ideal value, maximum & minimum values of criterion j ; $f_j(a)$ =

value of RNWDNS a for criterion j

2. Compute the measure $Q_i = \nu \frac{(S_i - S^{**})}{(S^* - S^{**})} + (1 - \nu) \frac{(R_i - R^{**})}{(R^* - R^{**})}$;

3. $R^* = \text{Max}(R_i, i = 1, 2, \dots, N)$; $R^{**} = \text{Min}(R_i, i = 1, 2, \dots, N)$

4. $S^* = \text{Max}(S_i, i = 1, 2, \dots, N)$; $S^{**} = \text{Min}(S_i, i = 1, 2, \dots, N)$. Range of ν is 0- 1.

5. Lower Q_i indicates suitable RNWDNS.

Details about case studies, methodology, results and discussion and conclusions are part of the manuscript.

Case Study

Hanoi WDN and Pamapur WDN (located in Telangana, India) are chosen for demonstrating the proposed methodology. Hanoi WDN comprises of thirty-two nodes, thirty-four pipes linking them, arranged in three loops as shown in Fig. 3 (Fujiwara and Khang 1990). The Pamapur WDN consists of one hundred and twenty-two pipes, one hundred and two demand nodes, and three tanks as shown in Fig. 4. The detailed information of the WDN is available in Pankaj *et al.* (2020).

Fig. 3 Layout of Hanoi WDN

Fig. 4 Layout of Pamapur WDN

Mathematical Modelling

The main purpose of mathematical modelling is to find out the optimal set of diameters of the WDN while balancing the network hydraulically, and satisfying the constraints such as minimum head requirements, and conservation of mass and energy. Chosen objectives for both the case studies are minimization of network cost, maximization of network resilience and minimization of network leakages. In addition, pipe diameters should be chosen from the commercially available set. Mathematical representations of each objective are given below:

Cost of pipe is based on diameter and length, expressed as:

$$\text{Network Cost } (c) = \sum_{j=1}^{nop} (ucp)_j \times (L)_j \quad (6)$$

where nop = number of pipes in network, ucp_j = unit cost of j^{th} pipe in network, L_j = length of j^{th} pipe in network.

Resilience is the capacity of system to meet unexpected demands which require surplus power. Todini (2000) used this concept to characterize the resilience of WDN. Prasad and Park (2004) introduced uniform coefficient (C_j) for computing resilience index that facilitates reliable loops.

$$\text{Network resilience } (I_{nr}) = \frac{\sum_{j=1}^{nn} C_j Q_j (H_j - H_j^{min})}{\left(\sum_{k=1}^{nr} Q_r H_r + \sum_{i=1}^{npu} \frac{P_i}{\gamma} \right) - \sum_{j=1}^{nn} Q_j H_j^{min}} \quad (7)$$

$$\text{where } C_j = \frac{\sum_{i=1}^{npj} D_i}{npj \times \text{maximum}(D_i)} \quad (8)$$

where nn = number of nodes, H_j = available head at demand nodes, Q_r = discharge at reservoir, H_r = head at reservoir, H_j^{min} = minimum head required at demand node, Q_j = demand at nodes, npu = number of pumps, npj = number of pipes linked to node, D_i = diameter of i^{th} pipe, and γ = specific weight of water.

The leakage is determined based on the pipe pressure and calculated as follows (Tucciarelli *et al.* 1999)

$$Leakages = (H_{i,t})^{\alpha} \times \sum_{j=1}^{npj} \frac{\pi}{2} * D_i * L_{ij} * \theta_{ij} \quad (9)$$

where L_{ij} = length of pipe linking nodes i and j , θ_{ij} = leak per unit surface of the pipe linking nodes i and j , α = loss exponent, $H_{i,t}$ = calculated pressure at node i at time step t .

Results and Discussion

NSGA-II - EPANET 2.2 framework is established for accomplishing optimal design of WDN (Rossman 2000). Continuity and energy conservation constraints are fulfilled externally via EPANET 2.2 whereas remaining is accommodated by NSGA-II utilizing exterior penalty mechanism. NSGA-II parameters are selected after a careful trial and error for both the case studies. NSGA-II was run for ten independent trials with a population size of 2000, 70000 generations, random multipoint crossover with value of 0.8, and Gaussian mutation with value of 0.8. In this study, random multipoint crossover operator is used instead of double point or single point crossover operator as it has shown promising results in WDNs.

Hanoi WDN

Six commercially available pipes are considered and minimum needed pressure head at each node taken is 30 m. Fig. 5(a) presents generated three dimensional pareto front consisting of 1020 NWDNS that shows interdependency between all the chosen objectives. X, Y, Z axes represent resilience, leakages and cost. Among the three objectives, the cost varies from 6.09×10^6 \$ to 10.96×10^6 \$, network resilience varies from 0.1758 to 0.3529, network leakages vary from 1.076×10^{-3} to 2.076×10^{-3} m³/s. It is observed from results that network cost is (a) directly proportional to network leakages up to some cost (b) indirectly proportional to network leakages up to certain cost, and (c) indirectly proportional to network resilience. This

shows that network cost, network resilience, network leakages are dependent on four decision variables, namely, diameter, length of pipe, pressure and discharge at nodes of WDN. Also observed is that most of the NWDNS are very closely spaced and requires facilitation of FCA for clustering similar NWDNS for effective decision making.

(a) Generated Pareto Front (1020)

(b) Clustered Pareto Front with ranks (18)

Fig. 5 Pareto front of Hanoi WDN

Matlab-based FCA using the function fcm was employed. Inputs are payoff matrix (1020×3), number of iterations, and lowest & highest number of clusters leading to the optimum size, and tolerance value. Optimal size of clusters, membership value of each NWDNS and the representative NWDNS for each cluster are the outputs. The FCA code was run for minimum 3 and maximum of 10 clusters (represented as 3-10), 3-15, 3-20, 3-25, 3-30, 3-35, 3-40, 3-45, 3-50 with maximum number of iterations 10^6 and tolerance value of 1×10^{-6} . Respective optimal cluster sizes obtained on the basis of DCVI are 7 (3-10), 13 (3-15), 18 (3-20), 18 (3-25), 18 (3-30), 33 (3-35), 39 (3-40), 43 (3-45), 46 (3-50). After deliberations, the optimum cluster size chosen was 18 in the cluster segment of (3-30).

Fig. 6 presents squared error value and DCVI for (3-30) segment. It is noted that squared error varies from 77.08 (cluster 3) to 7.08 (cluster 30) and corresponding DCVI from 0.6210 (cluster 3) to 0.2055 (cluster 30). It is noticed that squared error decreases from cluster 3 to cluster 17, increases again up to cluster 18 and thereafter shows a decreasing trend. Interestingly, at cluster 18, DCVI is minimum i.e. 0.1409, indicating that it is optimal cluster.

Fig. 6 Squared error value and DCVI for clusters 3 to 30

The membership value in each partition implies possibility of the NWDNS to be clustered in that particular partition. NWDNS having maximum membership in each of the 18 partitions is the representative NWDNS i.e. RNWDNS. It is also observed from 1020 NWDNS that in most of the cases, higher membership value of NWDNS is unique and distinct as compared to membership values in other clusters.

Table 1 presents the number of NWDNS falling in cluster groups 1–18 totaling 1020, representative NWDNS i.e. RNWDNS and corresponding membership. It also presents notation for RNWDN, network cost, resilience and leakages corresponding to each RNWDNS. The lowest and highest numbers of NWDNS per group are 21 and 97. Average

NWDNS per group is 57. Similarly, minimum and maximum membership values as evident from Table 1 are 0.8887 and 0.9984.

Table 1 Representative NWDNS i.e. RNWDNS (Hanoi WDN)

Cluster	No. of NWDNS in each cluster	Representative NWDNS i.e. RNWDNS	Membership corresponding to RNWDNS	Notation for RNWDNS Data set	Network Cost (10^6 \$)	Network Resilience	Network Leakages (10^{-3} m ³ /sec)
1	41	238	0.9535	A1	7.7788	0.3206	0.6139
2	89	819	0.9939	A2	6.4949	0.2491	1.0952
3	78	195	0.9936	A3	8.7613	0.3388	1.0952
4	58	494	0.9763	A4	6.6155	0.2568	0.2784
5	38	124	0.9806	A5	9.6546	0.3483	1.0952
6	52	914	0.9905	A6	6.3422	0.2279	1.0952
7	97	534	0.9882	A7	7.0036	0.2908	1.0952
8	92	676	0.9853	A8	6.7050	0.2745	1.0952
9	21	41	0.9405	A9	9.3039	0.3442	0.2720
10	75	281	0.9900	A10	7.8413	0.3251	1.0952
11	24	945	0.9661	A11	6.1994	0.1974	0.5810
12	36	85	0.9245	A12	9.2536	0.3440	0.6545
13	31	37	0.8887	A13	7.8818	0.3194	0.2024
14	65	652	0.9587	A14	6.6937	0.2715	0.6581
15	48	409	0.9910	A15	7.1435	0.2960	0.5981
16	58	990	0.9984	A16	6.1587	0.1946	1.0952
17	42	816	0.9745	A17	6.4030	0.2405	0.5862
18	75	406	0.9836	A18	7.3210	0.3064	1.0952
Min.	21		0.8887		6.1587	0.1946	0.2024
Max.	97		0.9984		9.6546	0.3483	1.0952

Table 2 presents results corresponding to PROMETHEE-2, MCQA-2 and VIKOR and are discussed as follows:

- PROMETHEE-2: ϕ^+ , ϕ^- , net ϕ for all RNWDNS are based on usual criterion function.
- Net ϕ varies between 0.2942 and -0.1765 spread over 9 ranks.

- 293 • A9 representing RNWNDS 41 has the highest net ϕ value of 0.2942 and ranked 1st
294 position, succeeded by A13 & A4 corresponding to RNWNDS 37 & 494 having net ϕ
295 values of 0.255.
- 296 • Eight strategies (A2, A3, A5, A6, A7, A8, A16, and A18) were given same rank due to
297 their same net ϕ value of -0.1765.
- 298 • It is also interesting to note that out of 18 strategies, 9 strategies have positive net ϕ values
299 whereas remaining have negative net ϕ values, and interestingly, the average net ϕ value
300 is zero. It clearly demonstrated the preferred and not preferred RNWDNS. Accordingly,
301 some of these can be discarded for further usage.
- 302 • MCQA-2: Normalized project satisfaction index, normalized project concordance index,
303 and normalized project discordance index (PSIN, PCIN, PDIN) are used for computing
304 L_p metric value for each RNWDNS. p value chosen is 2. Five slicing levels were
305 considered i.e. 0.2, 0.4, 0.6, 0.8, and 1, for computing shared face matrix after several
306 trials.
- 307 • Range of L_p values is 0 -1.5274 spread over 11 ranks. It is also interesting to note that
308 PCIN values are zero except for A13; PDIN values are zero except for A5 and A16.
- 309 • A13 occupied the first position (with L_p value of 0) and A9 with second position (L_p
310 value of 1.0183). A16 and A5 occupied last positions (10th and 11th) with respective L_p
311 values of 1.2762, 1.5274.
- 312 • It is noted that most L_p values are very close to each other indicating similar RNWDNS,
313 that is, A1, A4 with L_p value of 1.0356 has 3rd rank; A7, A10, A14 with L_p value of

1.0714 has 4th rank; A12, A17 with L_p value of 1.1014 has 6th rank; and A2, A3, A8, A11 with L_p value of 1.118 has 7th rank.

- VIKOR: ν value is taken as 0.5 (representing the average value) and corresponding range of Q_i measure varies from 0.0446 to 1 with average Q_i value of 0.6329.
- A13 and A15 occupy first and second positions due to their low Q_i values i.e. 0.0446 and 0.1457.
- Ties in case of VIKOR are very less compared to PROMETHEE-2 and MCQA-2. These are only A5 and A16 with 17th rank.

Fig. 5(b) presents pareto front and the corresponding final ranking pattern (as an outcome of additive ranking approach). For example, A13 occupied 2nd, 1st, 1st positions respectively by PROMETHEE-2, VIKOR, MCQA-2 (as shown in Table 2). Summations of these ranks are $2 + 1 + 1 = 4$. Accordingly, A13 was found to be the preferred as shown in Fig. 5(b). Procedure is repeated for all RNWDNS. It is observed that A13 is found to be preferred and promising representing RNWDNS 37.

It is observed from Table 1 that RNWDNS (A13) ranked first has minimum leakage of 0.2024×10^{-3} m³/sec, good resilience of 0.3194 in comparison with the maximum resilience of 0.3483. Slightly higher network cost of $\$7.8818 \times 10^6$ than the minimum network cost of $\$6.1587 \times 10^6$ is also observed.

333

Table 2 ranking pattern obtained by PROMETHEE-2, MCQA, VIKOR (Hanoi WDN)

RWDNS	PROMETHEE-2				MCQA-2					VIKOR			
	ϕ^+	ϕ^-	net ϕ	Rank	PSIN	PCIN	PDIN	Lp	Rank	S	R	Q	Rank
A1	0.5686	0.4314	0.1373	5	0.7308	0	0	1.0356	3	1.1041	0.4634	0.1785	3
A2	0.3333	0.5098	-0.1765	9	0.5	0	0	1.118	7	1.741	1	0.9019	14
A3	0.3333	0.5098	-0.1765	9	0.5	0	0	1.118	7	1.806	1	0.9265	15
A4	0.6275	0.3725	0.2550	2	0.7308	0	0	1.0356	3	0.8107	0.595	0.1856	4
A5	0.3333	0.5098	-0.1765	9	0.4231	0	1	1.5274	11	2	1	1	17
A6	0.3333	0.5098	-0.1765	9	0.4231	0	0	1.1545	9	1.8358	1	0.9378	16
A7	0.3333	0.5098	-0.1765	9	0.6154	0	0	1.0714	4	1.6156	1	0.8543	11
A8	0.3333	0.5098	-0.1765	9	0.5	0	0	1.118	7	1.6365	1	0.8622	13
A9	0.6471	0.3529	0.2942	1	0.8077	0	0	1.0183	2	1.004	0.8997	0.5325	7
A10	0.3529	0.4902	-0.1373	8	0.6154	0	0	1.0714	4	1.6322	1	0.8606	12
A11	0.6079	0.3921	0.2157	3	0.5	0	0	1.118	7	1.4175	0.982	0.7631	9
A12	0.5294	0.4706	0.0588	6	0.5385	0	0	1.1014	6	1.4195	0.8853	0.677	8
A13	0.6275	0.3725	0.2550	2	1	1	0	0	1	0.6807	0.4929	0.0446	1
A14	0.5098	0.4902	0.0196	7	0.6154	0	0	1.0714	4	1.1627	0.5103	0.2429	5
A15	0.5686	0.4314	0.1373	5	0.5769	0	0	1.0858	5	1.0651	0.4433	0.1457	2
A16	0.3333	0.5098	-0.1765	9	0.3846	0	0.5	1.2762	10	2	1	1	17
A17	0.5883	0.4117	0.1765	4	0.5385	0	0	1.1014	6	1.2009	0.7011	0.4287	6
A18	0.3333	0.5098	-0.1765	9	0.4615	0	0	1.1358	8	1.6052	1	0.8504	10

334

335

336

The optimal WDN design represented by A13 is shown in Table 3. As the network trades off between network cost, resilience and leakages, it is observed that 16 out of 34 pipes have maximum diameter of 1016 mm, whereas only one pipe (13) has the diameter of 406.4 mm. Fig. 7 represents the nodal pressure of optimal network design. It is also observed that node 1 has a very high pressure head of 97.14 m as it is the first node of the network. Pressure head in the remaining nodes ranges from 40.73 m to 61.67 m.

Table 3 Pipe Diameters of Optimal Hanoi WDN design (A13)

Pipe Number	Number of Pipes with same Diameter	Pipe Diameter (mm)
30, 31, 32	3	304.8
13	1	406.4
14, 15, 33	3	508
12, 25, 26, 27, 29, 34	6	609.6
10, 11, 22, 24, 28	5	762
1, 2, 3, 4, 5, 6, 7, 8, 9, 16, 17, 18, 19, 20, 21, 23	16	1016

Fig. 7 Node Pressure of Optimal Hanoi WDN (A13)

Pamapur WDN

All the pipes are manufactured of polyvinyl chloride, and Hazen-Williams roughness coefficient of 130 is used in all pipes. Minimum pressure head of all demand nodes chosen is 6 m. However, it is 5.75 m for node 22 (Pankaj *et al.* 2020).

Fig. 8(a) presents pareto front representing 272 of NWDNS. The FCA code was run for 3-10, 3-15, 3-20, 3-25, 3-30. Respective optimal cluster sizes obtained on the basis of DCVI are 6, 9, 19, 24 and 30 and accordingly, optimal cluster size identified is 9. Table 4 presents RNWDNS where Fig. 8(b) presents pareto front for 9 RNWDNS along with their respective ranks. Parameters employed in Hanoi network have been used for decision making aspects of Pamapur. Some of the relevant observations are as follows:

- Net ϕ varies between 0.3334 and -0.3334 spreading over 9 ranks in case of PROMETHEE-2. Best RNWDNS, S6 and S8 (representing 203 and 47) having net ϕ values of 0.3334 and 0.2501 respectively.
- Interestingly, S1, S4, S5 and S9 have negative net ϕ values and S2 has zero net ϕ value which indicates these RNWDNS are not worth considering. However, there is no tie in the ranking pattern.
- L_p values range from 0 -1.0487 spreading over 7 ranks in case of MCQA-2.
- RNWDNS, S6 and S8 occupy first and second ranks with respective L_p metric values of 0 and 0.1053.
- Tie occurred for S4 & S9 at rank 6 with L_p metric value of 1.034; similarly, tie for S1 & S7 at rank 7 is observed with L_p metric value of 1.0487.
- Q_i measure varies from 0.0255 to 1 in case of VIKOR. There is no tie in the ranking pattern.

- RNWDNS, S4 and S3 occupy first and second positions due to their low Q_i values i.e. 0.0255 and 0.1453.
- Additive ranking clearly indicates that RNWDNS, S6 is the most preferred WDN design. Ranking pattern is presented in Fig.9.

Table 4 Representative NWDNS i.e. RNWDNS (Pamapur WDN)

Cluster	No. of NWDNS in each cluster	Representative NWDNS i.e. RNWDNS	Membership corresponding to RNWDNS	Notation for RNWDNS Data set	Network Cost (10^6 Rs)	Network Resilience	Network Leakages (10^{-3} m ³ /sec)
1	29	16	0.9783	S1	1.2831	0.2902	0.7740
2	39	100	0.9901	S2	1.7571	0.6040	0.7212
3	28	166	0.9781	S3	2.4844	0.7752	0.6924
4	28	74	0.9932	S4	1.5785	0.5399	0.9012
5	36	147	0.9620	S5	2.2060	0.7354	0.8600
6	24	203	0.9582	S6	3.5159	0.8367	0.5317
7	26	258	0.9957	S7	6.0978	0.8802	0.6629
8	27	47	0.9624	S8	1.4190	0.4305	0.6512
9	36	227	0.9897	S9	4.0442	0.8547	0.8318
Max.	39		0.9957		6.0978	0.8802	0.9012
Min.	24		0.9582		1.2831	0.2902	0.5317

(a) Generated Pareto Front (272)

(b) Clustered Pareto Front with ranks (9)

Fig. 8 Pareto front of Pamapur WDN

It is observed from Table 4 that RNWDNS (S6) ranked first has minimum leakage of $0.5317 \times 10^{-3} \text{ m}^3/\text{sec}$, good resilience of 0.8367 in comparison with the maximum resilience of 0.8802 with a slightly higher network cost of Rs. 3.5159×10^6 million rupees than the minimum network cost of Rs. 1.2831×10^6 .

Fig. 9 Ranking Pattern obtained by PROMETHEE-2, MCQA, VIKOR and Final Rank using Additive Ranking (Pamapur WDN)

The optimal WDN design represented by S6 is shown in Table 5. As the network trades-off between network cost, resilience and leakages, it is observed that 52 out of 122 pipes have diameters of 98.6 mm (25) and 112 mm (27). It is observed that nodal pressure head varies in the range of 7.67 m to 19.04 m.

393

Table 5 Pipe Diameters of Optimal Pampaur WDN design (S6)

Pipe Number(s)	Number of Pipes with same Diameter	Pipe Diameter (mm)
1, 2, 7, 16, 25, 26, 29, 33, 44, 50, 51, 59, 119	13	56.2
58, 72, 96, 97, 98, 110, 113, 116	8	66.8
19, 73, 99, 111, 115, 122	6	80.4
6, 8, 11, 12, 18, 23, 28, 34, 35, 37, 41, 47, 48, 49, 53, 56, 64, 77, 80, 83, 88, 89, 100, 101, 112	25	98.6
3, 9, 14, 15, 17, 20, 21, 22, 24, 36, 38, 39, 43, 46, 52, 57, 60, 61, 62, 65, 67, 71, 82, 86, 109, 117, 121	27	112
4, 13, 27, 32, 45, 55, 69, 76, 79, 81, 84, 85, 87, 91, 120	15	125.4
5, 31, 40, 75, 90	5	143.4
30, 68, 70, 74, 78, 92, 93, 94, 95, 103, 105, 107, 118	13	161.4
54, 66, 102, 106	4	179.4
104	1	201.8
42, 63	2	224.4
10, 108, 114	3	251.4

394

395 **Conclusions**

396 The present study applied NSGA-II, FCA, DCVI, PROMETHEE-2, MCQA-2 and VIKOR
 397 for Hanoi Network and real-world case study, namely, Pamapur WDN to determine the
 398 optimal design in a multiobjective scenario. Specific conclusions drawn from the study are:

- 399 • PROMETHEE-2 ranks A9, A4 & A13 in the first and second positions; MCQA-2 ranks
 400 A13 and A9 in the first and second positions; VIKOR ranks A13 and A15 the first and
 401 second positions. However, additive ranking prefers A13 followed by A4 which
 402 simplifies the decision making process.
- 403 • It is observed that network cost (10^6 \$), network resilience, network leakages (10^{-3}
 404 m^3/sec) respectively are: RNWDNS 37 (representing A13 that occupied first position) i.e.

7.8818, 0.3194, 0.2024; and RNWDNS 494 that (representing A4 that occupied second position) i.e. 6.6155, 0.2568, 0.2784 as noted from additive ranking approach for Hanoi WDN.

- Additive ranking based MCDM approach clearly indicates that RNWDNS, S6 (NWDNS 203) and S8 (NWDNS 47) are potential network designs for Pamapur WDN. Network cost (10^6 Rs.), network resilience, and network leakages (10^{-3} m³/sec) respectively are 3.5159, 0.8367, and 0.5317 for S6; and 1.4190, 0.4305, and 0.6512 for S8.
- Best ranked RNWDNS explicitly showcased an excellent trade-off across the three objectives (network cost, network resilience and leakages) to converge to a single optimal network design.
- Variation in the number of chosen pipe diameters in the optimal network design is reasonably broad in Pamapur WDN in comparison to Hanoi WDN.
- Implementing fuzzy cluster analysis and DCVI on generated pareto front by NSGA-II paved the path for effective decision making for such complex WDNs.

The present work can be extended in multiple directions. In this study, NSGA-II is employed as an initial exercise. However, work can be extended in NSGA-III back ground or with other suitable evolutionary algorithms. In addition, other clustering algorithms and validation indices are one of the more promising extensions that can be considered. Methodology proposed in the work is tested on one benchmark problem and one real world case study. However, potentiality of the proposed methodology will be interesting with more complex real-world case studies targeted as further studies.

Declarations

Funding (information that explains whether and by whom the research was supported)

429 Not applicable

430

431 **Conflicts of interest/Competing interests** (include appropriate disclosures)

432 None

433

434 **Availability of data and material** (data transparency)

435 Data not presented in the manuscript would be available on request

436

437 **Code availability** (software application or custom code)

438 Not applicable

439

440 **Additional declarations for articles in life science journals that report the results of**
441 **studies involving humans and/or animals**

442 Not applicable

443

444 **Ethics approval** (include appropriate approvals or waivers)

445 Not applicable

446

447 **Consent to participate** (include appropriate statements)

448 Yes

449

450 **Consent for publication** (include appropriate statements)

451 Yes

452

References

- Brans, J.P., Vincke, Ph., & Mareschal, B. (1986). How to select and how to rank projects: The PROMETHEE method. *European Journal of Operational Research*, 24: 228–238. [https://doi.org/10.1016/0377-2217\(86\)90044-5](https://doi.org/10.1016/0377-2217(86)90044-5)
- Brentan, B., Carpitella, S., Barros, D., Meirelles, G., Certa, A., & Izquierdo, J. (2021). Water quality sensor placement: a multi-objective and multi-criteria approach. *Water Resources Management*, 35: 225–241. <https://doi.org/10.1007/s11269-020-02720-3>
- Bui, X.K., Marlim, M.S., & Kang, D. (2020). Water network partitioning into district metered areas: a state-of-the-art review. *Water*, 12: 1002. <https://doi.org/10.3390/w12041002>
- Chang, K., Gao, J.L., Wu, W.Y., & Yuan, Y.X. (2011). Water quality comprehensive evaluation method for large water distribution network based on clustering analysis. *Journal of Hydroinformatics*, 13(3): 390-400. <https://doi.org/10.2166/hydro.2011.021>
- Cunha, M., Marques, J., Creaco, E., & Savić, D. (2019). A dynamic adaptive approach for water distribution network design. *Journal of Water Resources Planning and Management*, 145(7): 04019026. [https://doi.org/10.1061/\(ASCE\)WR.1943-5452.0001085](https://doi.org/10.1061/(ASCE)WR.1943-5452.0001085)
- Cunha, M., Marques, J., & Savić, D. (2020). A flexible approach for the reinforcement of water networks using multi-criteria decision analysis. *Water Resources Management*, 34: 4469–4490. <https://doi.org/10.1007/s11269-020-02655-9>
- Deb, K., Pratap, A., Agarwal, S., & Meyarivan, T. (2002). A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE transactions on evolutionary computation*, 6(2): 182-197. DOI: 10.1109/4235.996017
- Duckstein, L., & Nobe, S.A. (1997). Q-analysis for modeling and decision making. *European Journal of Operational Research*, 103(3): 411-425. [https://doi.org/10.1016/S0377-2217\(97\)00308-1](https://doi.org/10.1016/S0377-2217(97)00308-1)

477 Dunn, J.C. (1974). Well separated clusters and optimal fuzzy partitions. *Journal of*
478 *Cybernetics*, 4: 95-104. <https://doi.org/10.1080/01969727408546059>

479 Farmani, R., Walters, G., & Savic, D. (2006). Evolutionary multi-objective optimization of
480 the design and operation of water distribution network: total cost vs. reliability vs. water
481 quality. *Journal of Hydroinformatics*, 8(3): 165-179.
482 <https://doi.org/10.2166/hydro.2006.019b>

483 Fu, G., & Kapelan, Z. (2011). Fuzzy probabilistic design of water distribution networks.
484 *Water Resources Research*, 47, W05538, <https://doi.org/10.1029/2010WR009739>

485 Fujiwara, O., & Khang, D.B. (1990). A two-phase decomposition method for optimal design
486 of looped water distribution networks. *Water Resources Research*, 26(4): 539–549.
487 <https://doi.org/10.1029/WR026i004p00539>

488 Gaur, S., Raju, K.S., Kumar, S.N., & Bajpai, M. (2021) Multicriterion decision making in
489 groundwater planning, *Journal of Hydroinformatics*, <https://doi.org/10.2166/hydro.2021.122>

490 Gheisi, A., & Naser, Gh. (2015). Multistate reliability of water-distribution systems:
491 comparison of surrogate measures. *Journal of Water Resources Planning and Management*,
492 141(10): 04015018. [https://doi.org/10.1061/\(ASCE\)WR.1943-5452.0000529](https://doi.org/10.1061/(ASCE)WR.1943-5452.0000529)

493 Hu, C., Dai, L., Yan, X., Gong, W., Liu, X., & Wang, L. (2020). Modified NSGA-III for
494 sensor placement in water distribution system. *Information Sciences*, 509: 488–500.
495 <https://doi.org/10.1016/j.ins.2018.06.055>

496 Kapelan, Z. S., Savic, D.A., & Walters, G.A. (2005). Multiobjective design of water
497 distribution systems under uncertainty. *Water Resources Research*, 41: W11407,
498 [doi:10.1029/2004WR003787](https://doi.org/10.1029/2004WR003787). <https://doi.org/10.1029/2004WR003787>

499 Kirstein, J. K., Albrechtsen, H.J., & Rygaard, M. (2015). Topological clustering as a tool for
 500 planning water quality monitoring in water distribution networks. *Water Science and*
 501 *Technology: Water Supply*, 15(5): 1011-1018. <https://doi.org/10.2166/ws.2015.056>
 502 Lifshitz, R., & Ostfeld, A. (2018). Clustering for analysis of water distribution systems.
 503 *Journal of Water Resources Planning and Management*, 144(5): 04018016.
 504 [https://doi.org/10.1061/\(ASCE\)WR.1943-5452.0000917](https://doi.org/10.1061/(ASCE)WR.1943-5452.0000917)
 505 Liu, J., & Han, R. (2018). Spectral clustering and multicriteria decision for design of district
 506 metered areas. *Journal of Water Resources Planning and Management*, 144(5): 04018013.
 507 [https://doi.org/10.1061/\(ASCE\)WR.1943-5452.0000916](https://doi.org/10.1061/(ASCE)WR.1943-5452.0000916)
 508 Marques, J., Cunha, M., & Savić, D. A. (2018). Many-objective optimization model for the
 509 flexible design of water distribution networks. *Journal of Environmental Management*, 226:
 510 308–319. <https://doi.org/10.1016/j.jenvman.2018.08.054>
 511 Matlab 2018 <http://www.mathworks.com>.
 512 Montalvo, I., Joaquín, I., Schwarzeb, S., & Pérez-García, R. (2010). Multi-objective particle
 513 swarm optimization applied to water distribution systems design: An approach with human
 514 interaction. *Mathematical and Computer Modelling*, 52: 1219-1227.
 515 <https://doi.org/10.1016/j.mcm.2010.02.017>
 516 Novarini, B., Brentan, B.M., Meirelles, G., & Junior, E.L. (2019). Optimal pressure
 517 management in water distribution networks through district metered area creation based on
 518 machine learning. *Brazilian Journal of Water Resources*, 24: e37.
 519 <https://doi.org/10.1590/2318-0331.241920180165>
 520 Opricovic, S. & Tzeng, G. (2004). Compromise solution by MCDM methods: A comparative
 521 analysis of VIKOR and TOPSIS. *European Journal of Operational Research*, 156: 445-455.
 522 [https://doi.org/10.1016/S0377-2217\(03\)00020-1](https://doi.org/10.1016/S0377-2217(03)00020-1)

523 Pankaj, B.S., Naidu, M.N., Vasani, A., & Varma, M.R.R. (2020). Self-Adaptive Cuckoo
 524 Search algorithm for optimal design of water distribution systems. *Water Resources*
 525 *Management*, 34: 3129-3146. <https://doi.org/10.1007/s11269-020-02597-2>
 526 Patil, M.B., Naidu, M.N., Vasani, A., & Varma, M.R.R. (2020). Water distribution system
 527 design using multi-objective particle swarm optimisation, *Sādhanā*, 45: 21.
 528 <https://www.ias.ac.in/article/fulltext/sadh/045/0021>
 529 Prasad, T.D., & Park, N.S. (2004). Multiobjective genetic algorithms for design of water
 530 distribution networks. *Journal of Water Resources Planning and Management*, 130(1): 73-
 531 82. [https://doi.org/10.1061/\(ASCE\)0733-9496\(2004\)130:1\(73\)](https://doi.org/10.1061/(ASCE)0733-9496(2004)130:1(73))
 532 Raju, K.S., & Kumar, D.N. (2007). Classification of Indian meteorological stations using
 533 cluster and fuzzy cluster analysis, and Kohonen artificial neural networks. *Nordic Hydrology*,
 534 38, : 303–314. <https://doi.org/10.2166/nh.2007.013>
 535 Raju, K.S., Kumar, & D.N., Anmol, J. (2018). Prioritization of sub-catchments of a river
 536 basin using DEM and fuzzy VIKOR. *H2Open Journal*, 1(1): 1-11.
 537 <https://doi.org/10.2166/h2oj.2017.001>
 538 Ross, T.J. (2016). *Fuzzy logic with engineering applications*. Wiley
 539 Rossman, L.A. (2000). *EPANET 2 user's manual*, National Risk Management Research
 540 Laboratory. US Environmental Protection Agency, Cincinnati
 541 Salehi, S., Tabesh, M., & Ghazizadeh, M.J. (2018). HRDM method for rehabilitation of pipes
 542 in water distribution networks with inaccurate operational-failure data. *Journal of Water*
 543 *Resources Planning and Management*, 144(9): 04018053.
 544 [https://doi.org/10.1061/\(ASCE\)WR.1943-5452.0000943](https://doi.org/10.1061/(ASCE)WR.1943-5452.0000943)
 545 Todini, E. (2000). Looped water distribution networks design using a resilience index based
 546 heuristic approach. *Urban Water*, 2(2): 115–122. <https://doi.org/10.1016/S1462->

547 0758(00)00049-2

548 Tucciarelli, T., Criminisi, A., & Termini, D. (1999). Leak analysis in pipeline systems by
549 means of optimal valve regulation. *Journal of Hydraulic Engineering*, 125 (3): 277-285.
550 [https://doi.org/10.1061/\(ASCE\)0733-9429\(1999\)125:3\(277\)](https://doi.org/10.1061/(ASCE)0733-9429(1999)125:3(277))

551 Vasan, A., & Simonovic, S.P. (2010). Optimization of water distribution network design
552 using differential evolution. *Journal of Water Resources Planning and Management*, 136 (2):
553 279-287. [https://doi.org/10.1061/\(ASCE\)0733-9496\(2010\)136:2\(279\)](https://doi.org/10.1061/(ASCE)0733-9496(2010)136:2(279))

554 Wang, Qi., Wang, L., Huang, W., Wang, Z., Liu, S., & Savić, D.A. (2019). Parameterization
555 of NSGA-II for the optimal design of water distribution systems, *Water*, 11: 971.
556 <https://doi.org/10.3390/w11050971>

557 Wu, W., Maier, H.R., & Simpson, A.R. (2013). Multiobjective optimization of water
558 distribution systems accounting for economic cost, hydraulic reliability and greenhouse gas
559 emissions. *Water Resources Research*, 49(3): 1211-1225, <https://doi.org/10.1002/wrcr.20120>

Figures

Figure 1

Flow chart of FCA methodology

Figure 2

Flow chart of MCQA-2 Methodology

Figure 3

Layout of Hanoi WDN

Figure 4

Layout of Pamapur WDN

(a) Generated Pareto Front (1020)

(b) Clustered Pareto Front with ranks (18)

Figure 5

Pareto front of Hanoi WDN

Figure 6

Squared error value and DCVI for clusters 3 to 30

Figure 7

Node Pressure of Optimal Hanoi WDN (A13)

(a) Generated Pareto Front (272)

(b) Clustered Pareto Front with ranks (9)

Figure 8

Figure 9

Ranking Pattern obtained by PROMETHEE-2, MCQA, VIKOR and Final Rank using Additive Ranking (Pamapur WDN)