A Hybrid Approach for Considering Topography in Graph-Based Optimization of Water Distribution Networks

Sitzenfrei, Robert; Qiu, Mengning; Ostfeld, Avi; Savic, Dragan; Kapelan, Zoran

DOI
10.1061/9780784484852.078

Publication date
2023

Document Version
Final published version

Published in
World Environmental and Water Resources Congress 2023

Citation (APA)
Sitzenfrei, R., Qiu, M., Ostfeld, A., Savic, D., & Kapelan, Z. (2023). A Hybrid Approach for Considering Topography in Graph-Based Optimization of Water Distribution Networks. In S. Ahmad, & R. Murray (Eds.), *World Environmental and Water Resources Congress 2023: Adaptive Planning and Design in an Age of Risk and Uncertainty - Selected Papers from World Environmental and Water Resources Congress 2023* (pp. 831-841). American Society of Civil Engineers (ASCE). https://doi.org/10.1061/9780784484852.078

A hybrid approach for considering topography in graph-based optimization of water distribution networks

Robert Sitzenfrei[1*], Mengning Qiu[2,†], Avi Ostfeld[2], Dragan Savic[3,4], Zoran Kapelan[5]

[1]Unit of Environmental Engineering, University of Innsbruck, Technikerstrasse 13, Innsbruck 6020 Tirol, Austria.
**Corresponding author: Robert Sitzenfrei (robert.sitzenfrei@uibk.ac.at)*
[2]Technion-Israel Institute of Technology, Haifa, Israel.
[3]KWR Water Research Institute, Nieuwegein, The Netherlands
[4]Centre for Water Systems, University of Exeter, Exeter, United Kingdom
[5] Delft University of Technology, Faculty of Civil Engineering and Geosciences, Department of Water Management, Delft, The Netherlands.
†passed away on February 25, 2021

ABSTRACT

Water distribution networks (WDNs) are a vital component of urban water infrastructure. They transport water from production sites (sources) to spatially distributed consumers (sinks). Multi-objective optimization procedures are often used to minimize construction costs and at the same time maximize the resilience of such systems, which is usually a very computationally expensive task. Recently, highly efficient approaches based on complex network analysis (CNA) have been developed to solve this task more computationally efficiently. With CNA, very large WDNs can be optimized, considering network topology and demand distribution (using, e.g., demand edge betweenness centrality). However, existing CNA approaches do not consider network topography (i.e., height differences between sources and sinks). Comparing design solutions based on CNA with those found by Evolutionary Algorithms shows that the least-cost CNA design cannot compete with the latter. In this work, a hybrid approach is developed, where low-cost design CNA solutions are evaluated with a hydraulic solver (Epanet2), and subsequently the demand edge betweenness centrality distribution is iteratively altered for nodes with pressure deficits. This enhanced CNA-based optimization is tested on 2 different large case studies from the literature and shows promising results (2% costs increase). These solutions were obtained using significantly less computational effort (at least factor 1,000 faster), enabling solving very large WDN optimization problems (>150,000 decision variables).

Keywords: graph, multi-objective optimization, demand edge betweenness centrality, resilience, least cost design;

INTRODUCTION

Water distribution networks (WDNs) are a vital component of urban water infrastructure. They transport water from production sites (sources) to spatially distributed consumers (sinks). A major task of WDNs is also to reliably supply good quality water to consumers. While one aim is to design these systems cost-efficiently, there is also a need for resilience in the system (i.e., to cope with critical and abnormal conditions) [1]. This results in a multi-objective design problem with two contradicting objectives. Multi-objective optimization procedures are used to minimize construction costs and at the same time maximize the resilience of such systems, which is usually a very computationally expensive task [2]. Recently, an approach based on complex network analysis (CNA) has been developed to solve this task more computationally efficiently [3]. With CNA, very large WDNs can be optimized, considering network topology and the location of the sources and the sinks. With a centrality measure customized for WDN analysis (denoted ‘demand edge betweenness centrality’) and varying design velocities, a broad range of design solutions can be achieved with the CNA approach. A comparison of these design solutions to those obtained by evolutionary algorithms shows that a very relevant part of the Pareto front (around the knee bend) can be determined with the CNA approach. However, while design solutions are close to the least cost design, they cannot compete with those from evolutionary algorithms. The reason for it is that for the assessment of the performance of WDNs, there are more important parameters than just the network topology and the demand distribution. Specifically, the height differences between the sources and the demand nodes, and the flow paths are of utmost importance when determining the least-cost design of a WDN. However, in classical graph-based approaches, only the network topology is considered and the existing CNA-based methods are not able to consider the topography.

In the least-cost design of a WDN, the resilience is not considered and only the costs are minimized. In literature, many different approaches exist for the least-cost design of WDNs, e.g., optimal use of power surface [4], (hybrid) genetic algorithms [5] [6], hybrid discrete dynamically dimensioned search [7], nonlinear programming [8] or particle swarm optimization [9] to name a few but a CNA based approach is missing.

This work aims to address this issue by enhancing a CNA-based approach to designing WDNs by considering topography. However, the topography cannot be tackled directly in the CNA approach. Therefore, a hybrid approach is required, in which critical nodes are identified using a hydraulic solver. In this work, a hybrid CNA approach is developed, where least-cost CNA design solutions are evaluated with a hydraulic solver (Epanet2), and subsequently the demand edge betweenness centrality distribution is iteratively modified before being used again for the CNA-based design. This enhanced CNA-based optimization is tested on 2 different large case studies from the literature (>150,000 decision variables). The results are then compared with least-cost designs from the literature.

METHODOLOGY

WDN design based on complex network analysis (CNA). The hydraulic behavior of WDNs cannot be easily understood. There are interacting hydraulic structures like nodes (e.g., sources and sinks) and links (e.g., pipes, valves and pumps). Therefore, the hydraulic behavior of WDNs is complex an can be determined with hydraulic solvers.

CNA can also provide valuable insight into the characteristics and performance of networks. Compared to, e.g., hydraulic solvers, CNA requires much less computational effort [10]. The network topology of a WDN can be represented by a network graph. A network graph is composed of a set of vertices (nodes) which are connected via a set of edges (pipes, pumps, etc). The connections can be represented by an adjacency matrix A which is symmetric whose dimensions depend on the number of vertices, V x V. Each matrix element a_{ij} of A indicates if there is a connection (pipe) between node i and node j ($a_{ij}=1$) or not ($a_{ij}=0$). The elements in the graph can be weighted. A common edge weight for analysis of WDNs is the Euclidean distance (i.e., pipe length), but can also be related to volume (e.g., length/diameter) or residence time (length/velocity) [11]. The advantage of CNA is that it directly uses the above matrix for evaluations, which is advantageous from a computational efficiency perspective.

An important analysis of graphs in this work is the shortest path between two nodes i and j ($\sigma_{i,j}$). It is defined as the minimum of the sum of edge weights along that path. For transport graphs, the edge betweenness centrality (EBC) is an important measure of transport capacity. It counts how often an edge is part of $\sigma_{i,j}$ between all node pairs. While this is a reasonable indicator for networks like streets or the Internet (where connections between all node pairs are of interest), for WDNs this is not appropriate. Therefore, customization of EBC is required to include the hydraulic behavior of WDNs in the methodology. To determine the flows in a WDN with CNA, edges in $\sigma_{i,j}$ connecting sources and demand nodes (instead of all node pairs) are determined. Further, instead of counting the number of $\sigma_{i,j}$ passing through an edge, the actual demand at a node is used to weigh the EBC values along $\sigma_{i,j}$. This customized measure is called demand edge betweenness centrality (EBC^Q) [3].

For an illustrative example of a WDN in Figure (a), the EBC^Q is calculated in Figure (b). For Node 1, the shortest path to the source node is determined ($\sigma_{1,S1}$). For all edges along that path, the demand of Node 1 (a demand parcel of 2.1 L/s) is added to the EBC^Q values. When repeating this for all demand nodes, the EBC^Q (L/s) values from Figure (b) are obtained. Applying the continuity equation with a design velocity (v_{design}) of, e.g., 0.8 m/s, the required diameters can be determined. When using a set of available discrete pipe diameters (e.g., 76.2, 101.6 and 152.4 mm) and corresponding costs per one-meter pipe (8, 11, and 16 \$/m), the diameters can be determined for all pipes (rounding up to the nearest available diameter) and the total construction costs can be calculated (ranging from \$4,264 to \$3,814 in Figure 1(d)). For different values for v_{design} (0.8, 1.0 and 1.2m/s), the diameters can be determined, starting with the lowest one (see Figure 1(d)) and then v_{design} is step-wise increased. When the pressure distribution is checked with Epanet2 (EN2), possible pressure deficits can be determined (e.g., <30m). If there are no pressure deficits, the design solution is feasible and the next (higher) design velocity is evaluated. If there are pressure deficits present, these solutions are enhanced to meet the pressure criterion. This is a departure from the previous CNA design approach, which discarded these designs and did not consider them in the Pareto front solutions.

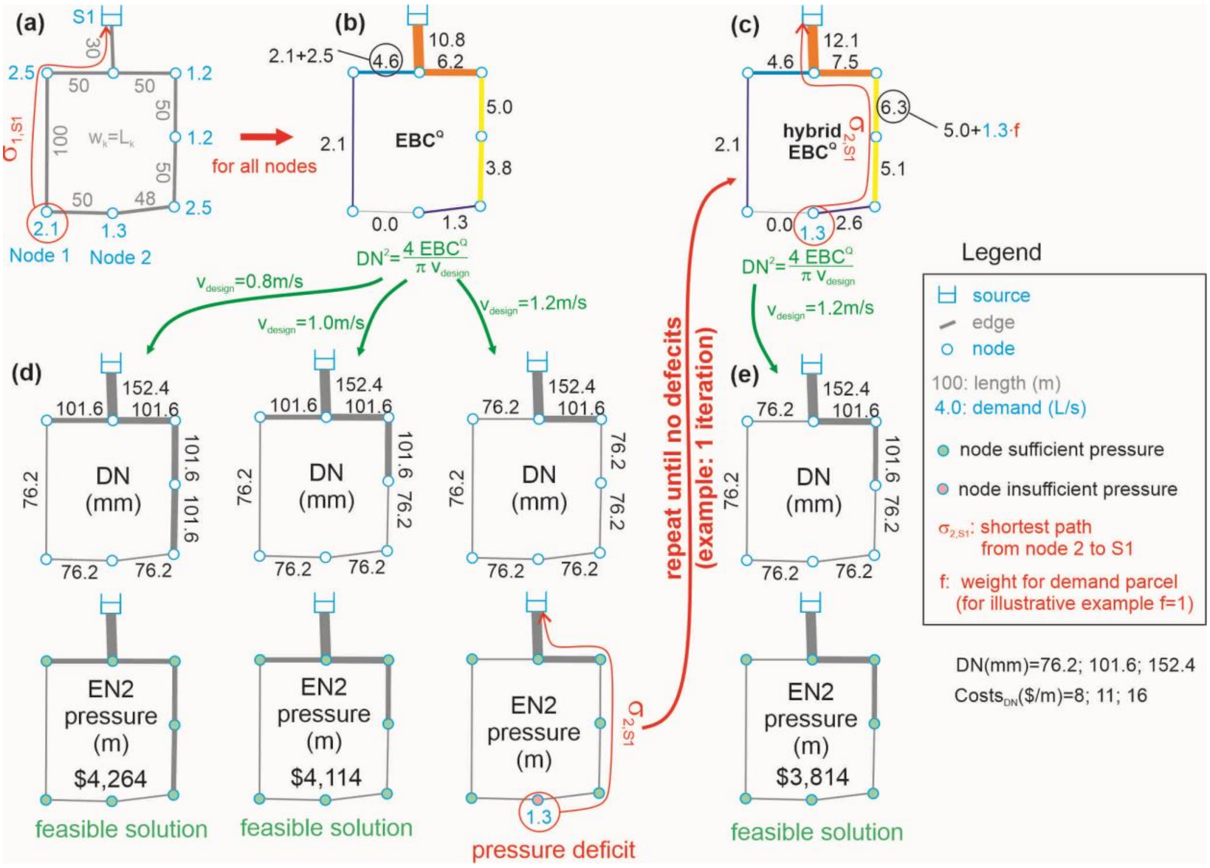


Figure 1. Illustrative example for (a) routing demand parcels (shortest path $\sigma_{i,j}$) to a source, (b) determining EBC^Q, (c) hybrid determining $EBC^{Q,}$ (d) determining diameters (DN) with EBC^Q, and (e) determining DN with hybrid EBC^Q

To improve the design solutions experiencing pressure deficits and reach pressure feasibility, the nodes with pressure deficits (node 2 in the illustrative example) are determined with Epanet2. This is followed by determining $\sigma_{i,j}$ to the closest source node ($\sigma_{2,S1}$ in the example Figure 1(d)). For the EBC^Q values along that path $\sigma_{i,j}$, the demand of the node with pressure deficit is added (i.e., 1.3 L/s in the example). The design procedure is repeated and the pressures are determined again with the new hybrid EBC^Q values. This procedure is repeated until there are no pressure deficits in the system (Figure 1(e)). In the next step, the design velocity is further increased and, if necessary, pressure deficits are again removed iteratively using the proposed hybrid approach.

For resilience assessment of the design solutions, a common resilience metric is used. Prasad and Park [12] determined how much excess energy (pressure above the required minimum pressure) is put in the system in comparison to (divided by) how much of that excess energy reaches the demand nodes. In addition, the numerator of that ratio is weighted with the actual demand and the uniformity of pipe connections [12]. A theoretical resilience value of 1 would mean that no pressure head is dissipated along the transport through the WDN, while a value of 0 would mean all demand nodes meet exactly the required minimum pressure.

When looking at the obtained Pareto front comprised of design solutions for the illustrative example (Figure 2), the proposed procedure (hybrid EBC^Q design) improves the solution space by making feasible solutions with higher flow velocities (v_{design} = 1.2, 1.4 m/s) and therefore lower resilience values and lower costs. The least-cost design solution is considered the one with the lowest cost (v_{design} = 1.4 m/s). As constraint the maximum number of iterations for hybrid EBC^Q can be considered. If no feasible solution is obtained within the maximum number of iterations, the optimization is enclosed. To obtain feasible solutions with less iterations, when calculating hybrid EBC^Q, also a weight f can be used (see Figure 1(c)). How to choose f and the maximum number of iterations depends on the WDN model. For example if there are many demand nodes with very small demands in a large network, a high weight f can be productive (f>10) potentially resulting in a smaller number of NFE.

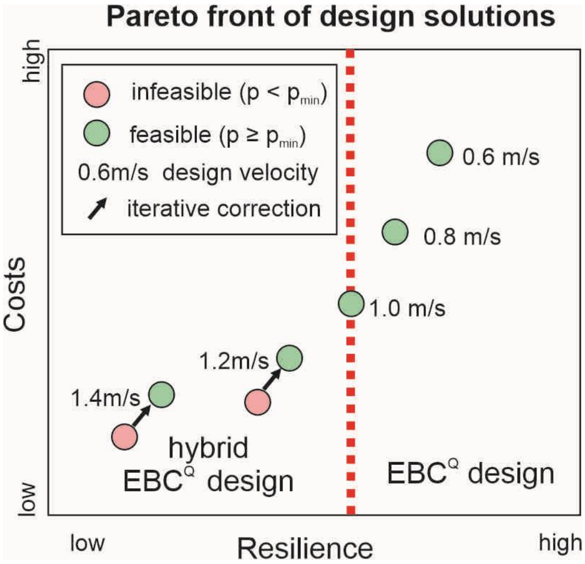


Figure 2. Illustrative example of the proposed hybrid approach

CASE STUDIES

In this work, two case studies from the literature are used. The first one is the Balerma case introduced by Reca and Martínez [5], which is classified as a large problem [6]. The layout and its nodal elevations are shown in Figure 3. It has four sources with a total demand of 1,104 L/s. There are 443 demand nodes and 454 pipes. There are 10 different diameters available starting from 113 to 581.8 mm with a unit cost ranging from 7.22 to 215.85 $/m (see also Figure 3). The minimum pressure criterion is 20 m. For the hybrid EBC^Q design, a weight for the demand parcels with pressure deficits f=25 is chosen.

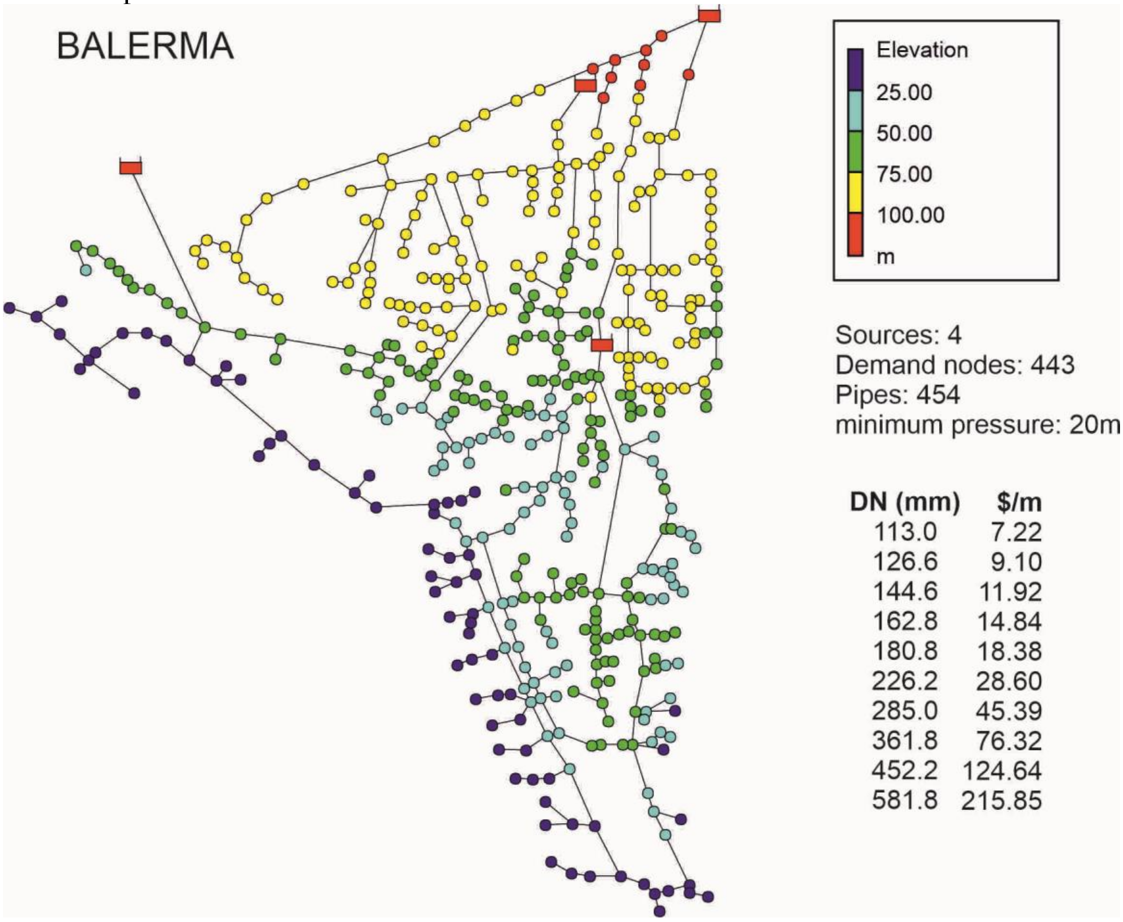


Figure 3. The layout of the Balerma irrigation case study

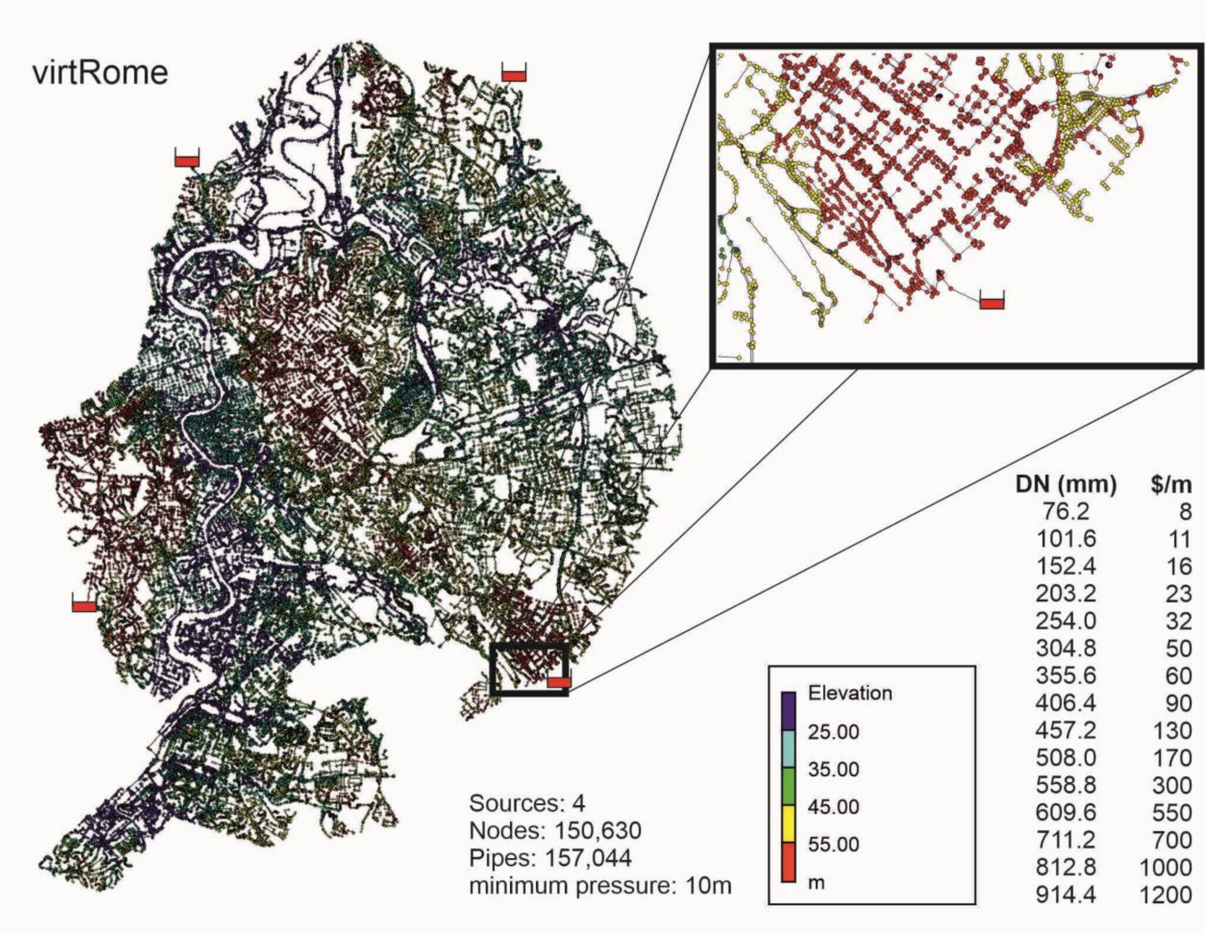
The summary of previous results obtained by some literature approaches is shown in Table 1. The least cost found is $1.921M. There are many more studies who identified optimal solutions with the cost above $2.0M [4] which are not listed here for clarity reasons.

Table 1: Least cost designs for Balerma case

least cost design (M $)	Number of evaluations	method	reference
2.015	833	optimal power use surface	[4]
2.002	254,400	genetic algorithm	[5]
1.999$^{*)}$	400,000	hybrid GA	[6]
1.940	30,000,000	hybrid discrete dynamically dimensioned search	[7]
1.923	1,427,850	non-linear programming	[8]
1.921	217,400	modified particle swarm optimization	[9]

$^{*)}$ least cost solution from a multi-objective optimization

The second case study is the virtRome network, which is an extremely large example of a WDN studied first by Sitzenfrei, et al. [3]. It was generated based on the street network and the topography of the city of Rome, Italy [13]. The model has more than 150,000 nodes and 157,000 pipes. It has four reservoirs and the total demand is 6,000 L/s. There are 15 different available diameters ranging from 76.2 to 914.4 mm with unit costs ranging from 8 to 1200 $/m (see Figure 4). The minimum pressure requirement is 10 m. For the hybrid $\mathrm{EBC}^{\mathrm{Q}}$ design, a weight for the demand parcels with pressure deficits f=50 is chosen.



Sources: 4
Nodes: 150,630
Pipes: 157,044
minimum pressure: 10m

DN (mm)	$/m
76.2	8
101.6	11
152.4	16
203.2	23
254.0	32
304.8	50
355.6	60
406.4	90
457.2	130
508.0	170
558.8	300
609.6	550
711.2	700
812.8	1000
914.4	1200

Figure 4. The layout of the virtRome case study

For virtRome there exists only one design approach so far in literature [3]. That approach considered two objectives (costs vs. resilience) and is based on CNA. The topography of virtRome is relatively flat with small height differences close to the reservoirs and in the middle of the case study area (see Figure 4, red areas). The minimum pressure criterion of 10 meters was therefore achieved in the CNA approach with a relatively low design velocity. The reported least-cost design from the literature is $129.68M which was achieved without considering information about this specific topography.

RESULTS AND DISCUSSION

For the Balerma case, Wang, et al. [6] determined a Pareto front of design solutions (costs vs. resilience) which is shown in Figure 5 (black dots) with a hybrid genetic algorithm (GA). Although this work focuses primarily on the least-cost design (because in that case, the topography plays the most important role), higher resilience designs are also presented and discussed. The red circles in Figure 5 indicate the results obtained with EBC^Q. With a total of 36 function evaluations (NFE) these were found. It can be noted that the GA outperforms the EBC^Q approach, however, only marginally, but with significantly higher computational effort (400,000 NFE [6] compared to 36 NFE, the EBC^Q requires much less computational budget than the NFE based on Epanet2 [3]). The design solutions based on the hybrid EBC^Q approach are shown as red dots in Figure 5. Some solutions are dominated and for a Pareto front, non-dominated solutions have to be determined. However, to better illustrate the proposed approach, all solutions (including dominated) are plotted. For the least cost design, a cost value of $1.96M was achieved with 182 NFE. This is only 2% above the best-known solutions with 217,400 NFE [9], but achieved with significantly lower computational effort (factor 1,000 faster).

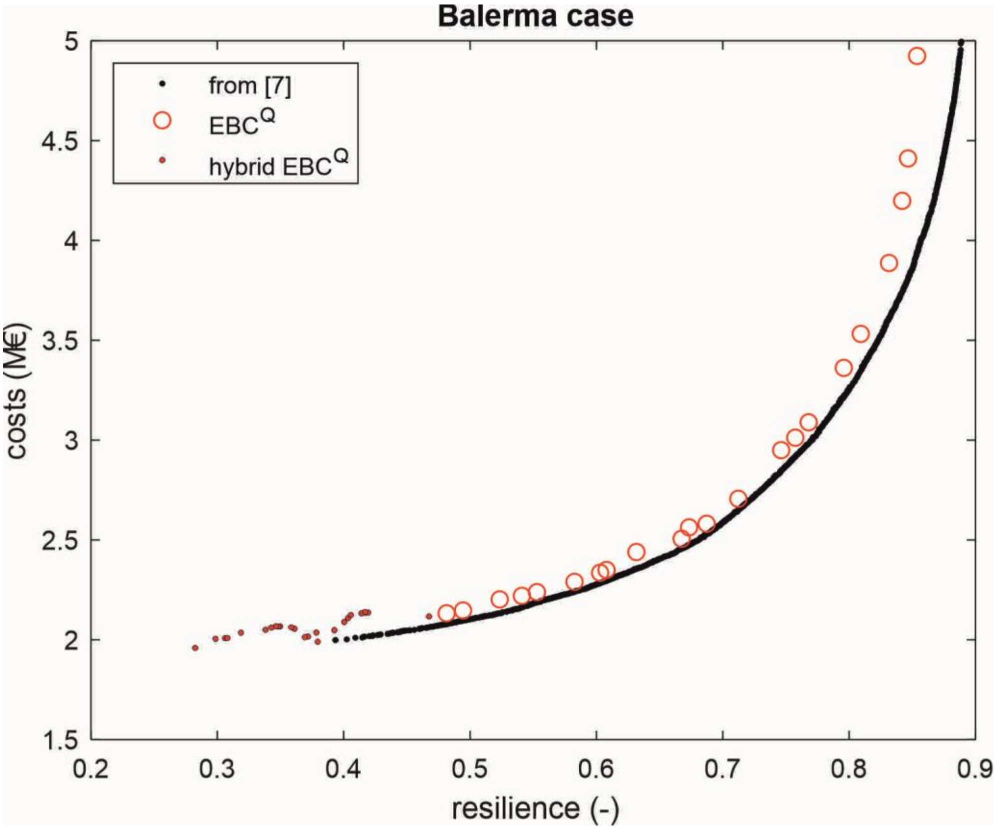


Figure 5. Design solutions for the Balerma case obtained with EBC^Q (red circles), hybrid EBC^Q (red dots), and a GA (black dots)

The design solutions for the virtRome case are shown in Figure 6. The design solutions based on EBC^Q from Sitzenfrei, et al. [3] are shown as red circles in Figure 6. Eight NFE were required for the EBC^Q solutions to be obtained. The design solutions from the hybrid EBC^Q are shown as red dots. For the least cost design, an NFE of 1,583 was required and minimal cost of $57.08M was achieved. Note again, that with a combination of EBC^Q and hybrid EBC^Q the latter NFE number, an entire Pareto front of design solutions was obtained. Note that one single run with Epanet2 for virtRome takes 23.8 seconds (13.2 seconds without loading the file) on a desktop PC (Intel® Core™ i5–6,500 CPU @ 3.2 GHz) making it very hard to solve for e.g., evolutionary algorithms which would require houndred thousands or even millions of NFEs. The optimization with hybrid EBC^Q took 8.66 hours (5.8 hours for NFEs with Epanet2 and 2.86 hours for graph analysis).

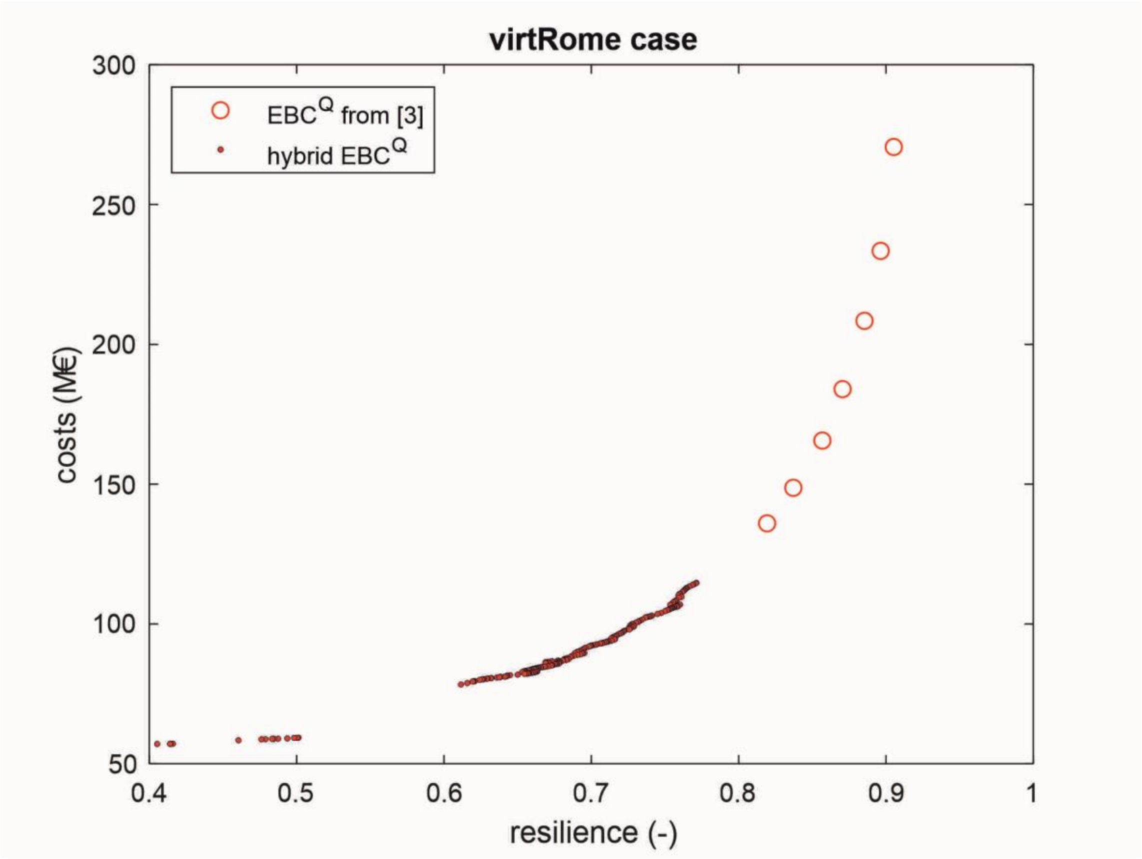


Figure 6. Design solutions for the virtRome case with EBC^{Q} (circles) and hybrid EBC^{Q} (dots)

SUMMARY AND CONCLUSIONS

With complex network analysis, very large WDNs can be optimized, considering network topology and demand distribution. However, these approaches have not been able to consider the topography of these pipe networks. This work addressed this shortcoming by introducing a hybrid approach that uses complex network analysis and hydraulic simulations. Identified pressure deficits are subsequently addressed iteratively with complex network analysis. By that enhanced (hybrid) approach, the solutions space obtained from complex network analysis is increased towards the least cost design. For least-cost design, the topography is of utmost interest and, therefore, the least-cost designs obtained in this study are compared to the least-cost designs from the literature. For a well know large benchmark case (the Balerma irrigation network), the obtained least-cost design is 2% above the best-known solution. However, it was obtained using only 182 function evaluations (i.e., runs of the hydraulic solver to assess if the pressure criterion is met) compared to 400,000 NFE [6]. For a very large case study (virtRome) with more than 150,000 decision variables, the least-cost design solution ($57,08M) was determined with 1,583 function evaluations. One single run (NFE) with Epanet2 for virtRome takes 23.8 seconds. This shows the need for high computational efficiency.

The proposed hybrid approach is capable of identifying a wide range of multi-objective design solutions (from the least-cost design to the knee bend of the Pareto front) with significantly less computational effort compared to evolutionary approaches from the literature. The obtained results are very close in terms of costs to those from the literature. Therefore, the proposed approach can be used for studies where a great number of scenarios are required to evaluate and compare the solutions or to seed the evolutionary algorithm starting population and improve their computational efficiency. Further, it was shown, that with the proposed approach, WDN optimization problems can be solved (>150,000 decision variables), which are virtually insolvable with approaches like evolutionary algorithms.

ACKNOWLEDGEMENTS

This study was funded by the Austrian Science Fund (FWF): P 31104-N29.

REFERENCES

[1] Q. Zhang, F. Zheng, Q. Chen, Z. Kapelan, K. Diao, K. Zhang, *et al.*, "Improving the Resilience of Postdisaster Water Distribution Systems Using Dynamic Optimization Framework," *Journal of Water Resources Planning and Management*, vol. 146, p. 04019075, 2020.

[2] H. Mala-Jetmarova, N. Sultanova, and D. Savic, "Lost in optimisation of water distribution systems? A literature review of system operation," *Environmental Modelling & Software,* vol. 93, pp. 209-254, 7// 2017.

[3] R. Sitzenfrei, Q. Wang, Z. Kapelan, and D. Savić, "Using Complex Network Analysis for Optimization of Water Distribution Networks," *Water Resources Research,* vol. 56, p. e2020WR027929, 2020.

[4] J. Saldarriaga, D. Páez, N. León, L. López, and P. Cuero, "Power use methods for optimal design of WDS: history and their use as post-optimization warm starts," *Journal of Hydroinformatics,* vol. 17, pp. 404-421, 2014.

[5] J. Reca and J. Martínez, "Genetic algorithms for the design of looped irrigation water distribution networks," *Water Resources Research,* vol. 42, 2006.

[6] Q. Wang, M. Guidolin, D. Savic, and Z. Kapelan, "Two-Objective Design of Benchmark Problems of a Water Distribution System via MOEAs: Towards the Best-Known Approximation of the True Pareto Front," *Journal of Water Resources Planning and Management,* vol. 141, p. 04014060, 2015.

[7] B. A. Tolson, M. Asadzadeh, H. R. Maier, and A. Zecchin, "Hybrid discrete dynamically dimensioned search (HD-DDS) algorithm for water distribution system design optimization," *Water Resources Research,* vol. 45, 2009.

[8] F. Zheng, A. R. Simpson, and A. C. Zecchin, "A combined NLP-differential evolution algorithm approach for the optimization of looped water distribution systems," *Water Resources Research,* vol. 47, 2011.

[9] X. Qi, K. Li, and W. D. Potter, "Estimation of distribution algorithm enhanced particle swarm optimization for water distribution network optimization," *Frontiers of Environmental Science & Engineering,* vol. 10, pp. 341-351, 2016/04/01 2016.

[10] R. Sitzenfrei, M. Oberascher, and J. Zischg, "Identification of Network Patterns in Optimal Water Distribution Systems Based on Complex Network Analysis," in *World Environmental and Water Resources Congress 2019*, ed, 2019, pp. 473-483.

[11] R. Sitzenfrei, "Using complex network analysis for water quality assessment in large water distribution systems," *Water Research,* vol. 201, p. 117359, 2021/08/01/ 2021.

[12] T. D. Prasad and N.-S. Park, "Multiobjective Genetic Algorithms for Design of Water Distribution Networks," *Journal of Water Resources Planning and Management,* vol. 130, pp. 73-82, 2004.

[13] R. Sitzenfrei, "A review on network generator algorithms for water supply modelling and application studies," in *World Environmental and Water Resources Congress 2016*, ed, 2016, pp. 907-916.