Exploring the Puzzles in Comparing Reliability Surrogate Measures for Water Distribution System Design: From the Dual Perspectives of Equal-Cost and Consistency

Yuehua Huang¹, Qianqian Zhou¹, Qi WANG¹, and Zhiwei Zhao¹
¹Guangdong University of Technology

December 27, 2023

Abstract

Various reliability surrogate measures (RSMs) have emerged over last three decades to design water distribution systems (WDSs). However, there has yet to be a consensus on their rationalities. In this work, we proposed a novel dual-perspective framework to examine RSMs’ rationalities, including the equal-cost perspective (ECP) by comparing RSMs’ performance at three comparable equal-cost levels and the traditional consistency perspective (TCP) widely used in other comparative studies. RSMs’ performance was evaluated by the mechanical reliability score (MRS) and the normalized resilience score (NRS). The former compared fewer pipe failures under a snapshot scenario, while the latter considered more comprehensive pipe failures using global resilience analysis. We compared five well-known RSMs from the literature, i.e., Network Resilience Index (NRI), Diameter-Sensitive Flow Entropy (DSFE), Redundancy (REDU), Pipe Hydraulic Resilience Index (PHRI) and Combined Entropy-Resilience Index (CERI). We verified their performance for the multi-objective design of two medium-scale, single-source WDSs, i.e., Hanoi and Fossolo. Results reveal that NRI is the most rational measure. The rankings from NRS are similar to those from MRS under the ECP; however, the two rankings vary significantly under the TCP. The ECP is more reasonable than the TCP because it measures network resilience and reliability from practical performance. We recommend using the MRS+ECP framework to compare RSMs since it reflects the critical aspect of global resilience with lower computational overhead. Furthermore, some meta information, e.g., nodal pressure head and demand, has the most significant impact on the rankings of RSMs, inspiring the development of new and better RSMs.

Figure 1: This is a caption
1. Select equal-cost solutions.
2. Run global resilience analysis.
3. Find comparable stress range.
4. Get normalized resilience score.

Figure 2: This is a caption

Figure 3: This is a caption

Figure 4: This is a caption
Figure 5: This is a caption
Figure 6: This is a caption
Figure 7: This is a caption
Exploring the Puzzles in Comparing Reliability Surrogate Measures for Water Distribution System Design: From the Dual Perspectives of Equal-Cost and Consistency

Yuehua Huang\textsuperscript{1}, Qianqian Zhou\textsuperscript{1,2}, Qi Wang\textsuperscript{1,2}, Zhiwei Zhao\textsuperscript{1,2}

\textsuperscript{1}School of Civil and Transportation Engineering, Guangdong University of Technology, Guangzhou 510006, China
\textsuperscript{2}Cross Research Institute of Ocean Engineering Safety and Sustainable Development, Guangdong University of Technology, Guangzhou 510006, China

Corresponding author: Qi Wang (q.wang@gdut.edu.cn)

Key Points:

- A dual-perspective framework to compare reliability surrogate measures was proposed, and the equal-cost perspective was more significant.
- Comparable stress range plays an essential role in global resilience analysis under the dual-perspective framework.
- Nodal head, nodal demand, and pipe diameter had a more significant impact on reliability surrogate measures’ performance.
Abstract

Various reliability surrogate measures (RSMs) have emerged over last three decades to design water distribution systems (WDSs). However, there has yet to be a consensus on their rationalities. In this work, we proposed a novel dual-perspective framework to examine RSMs’ rationalities, including the equal-cost perspective (ECP) by comparing RSMs’ performance at three comparable equal-cost levels and the traditional consistency perspective (TCP) widely used in other comparative studies. RSMs’ performance was evaluated by the mechanical reliability score (MRS) and the normalized resilience score (NRS). The former compared fewer pipe failures under a snapshot scenario, while the latter considered more comprehensive pipe failures using global resilience analysis. We compared five well-known RSMs from the literature, i.e., Network Resilience Index (NRI), Diameter-Sensitive Flow Entropy (DSFE), Redundancy (REDU), Pipe Hydraulic Resilience Index (PHRI) and Combined Entropy-Resilience Index (CERI). We verified their performance for the multi-objective design of two medium-scale, single-source WDSs, i.e., Hanoi and Fossolo. Results reveal that NRI is the most rational measure. The rankings from NRS are similar to those from MRS under the ECP; however, the two rankings vary significantly under the TCP. The ECP is more reasonable than the TCP because it measures network resilience and reliability from practical performance. We recommend using the MRS+ECP framework to compare RSMs since it reflects the critical aspect of global resilience with lower computational overhead. Furthermore, some meta information, e.g., nodal pressure head and demand, has the most significant impact on the rankings of RSMs, inspiring the development of new and better RSMs.

1 Introduction

Water distribution systems (WDSs) always face various uncertain threats, so improving system performance is of great significance. Performance indicators can be expressed in terms of reliability (Goulter, 1995), vulnerability (Haimes, Matalas, Lambert, Jackson, & Fellows, 1998), resilience (Saldarriaga, Ochoa, Rodriguez, & Arbeláez, 2008), availability (Zhuang, Lansey, & Kang, 2013), and robustness (Jung, Kang, Kim, & Lansey, 2014), etc. Among them, reliability is widely used in network optimization design.

Over the last three decades, researchers have proposed various reliability surrogate measures (RSMs) to avoid complex and time-consuming reliability calculations (Atkinson, Farmani, Memon, & Butler, 2014). Well-known RSMs include flow entropy (FE) (Tanyimboh & Templeman, 1993), resilience index (RI) (Todini, 2000), network resilience (NRI) (Prasad & Park, 2004), modified resilience index (MRI) (Jayaram & Srinivasan, 2008), diameter sensitive flow entropy (DSFE) (H. Liu et al., 2014), redundancy (REDU) (S. Liu, Wang, & Xin, 2014), available power index (API) and pipe hydraulic resilience index (PHRI) (H. X. Liu, Savic, Kapelan, Creaco, & Yuan, 2017), combined entropy-resilience index (CERI) (Sirsant & Reddy,
2020), to name a few. The reliability is usually affected by network topological structure, linking to two basic properties: energy redundancy (Di Nardo et al., 2017) and hydraulic uniformity (Saleh & Tanyimboh, 2014). Thus, existing RSMs can be categorized into three groups: (1) energy redundancy-based, including RI, MRI, REDU, API, and PHRI; (2) hydraulic uniformity-based, including FE and DSFE; and (3) mixed by energy redundancy and hydraulic uniformity, including NRI and CERI.

However, finding better RSMs is still challenging for researchers. Raad et al. (2010) assessed FE, RI, NRI, and a mixed reliability index using three benchmarks, making correlation analysis on the Pareto optimal set from optimization. None of the measures were satisfactory to represent both hydraulic and mechanical reliability, although RI and NRI performed better than FE. Later, Liu et al. (2017) proposed API and PHRI and compared two new measures with RI, NRI, MRI, and DSFE under demand and pipe failure uncertainties. They found that energy-based measures performed better than uniformity-based ones in two networks. Wang et al. (2019) compared RI, NRI, API, PHRI, REDU, FE, and DSFE by establishing a many-objective optimization model. They considered single pipe failure mode in two benchmarks to analyze the correlation between mechanical reliability and RSMs, revealing that energy-based RSMs were more consistent with reliability and cost than uniformity-based ones. REDU was recommended due to satisfactory performance closer to other energy redundancy measures with much lower computational overhead. Sirsant et al. (2020) proposed two new measures, i.e., the combined entropy and resilience index (CERI) and the combined entropy and network resilience index (CENRI). They investigated the correlation between RSMs and reliability in four networks under demand fluctuations and single pipe failure, revealing that CERI performed well when facing uncertain threats in different network structures.

As mentioned above, previous studies only compare RSMs partially and comprehensively despite various comparative studies. First, they only considered single-pipe failure scenarios, neglecting multi-pipe failure scenarios due to low-temperature weather or earthquakes, which are Black Swan events but have incredibly significant impacts. Second, previous studies only focused on the consistency between RSMs and the reliability of Pareto optimal solutions. However, few studies compared RSMs’ absolute performance of equal-cost solutions on the Pareto front, in which the latter perspective significantly impacts the decision-making process in practice.

In recent years, resilience analysis has attracted increasing interest in academia. Resilience is an extension of reliability, considering more pipe failures and dynamic system performance over time. Diao et al. (2016) proposed the global resilience analysis (GRA) to acquire dynamic system performance over time in three failure modes. Later, Meng et al. (2018) enriched the GRA framework, considering a key aspect called the recovery stage during the
failure period. They extracted six strain indicators to reflect network performance, including time to strain (TTS), duration (DUR), failure rate (FR), recovery rate (RR), magnitude (MAG), and severity (SEV). The GRA framework extends the traditional comparative perspectives and adds value to decision-making.

The specific objective of this study is to assess and rank some typical RSMs in the GRA framework, compared with mechanical reliability calculation. We obtained a Pareto optimal solution set after bi-objective (cost and RSM) optimization. Then, we compared RSMs in both the equal-cost perspective (ECP) and the traditional consistency perspective (TCP) by calculating the mechanical reliability score (MRS) and normalized resilience score (NRS), respectively. As such, we identified the most potential RSM for network design.

This study also attempts to explore why some RSMs perform with distinction. It may be due to the unique topological and hydraulic information combinations to calculate RSMs. Different network information has varied influence on network performance. So, the critical calculation information of RSMs can be found by contrasting the TCP and ECP rankings. While some RSMs perform acceptably, there is considerable room for improvement in the more complex condition of networks. This study provides insights on developing better RSMs in the future.

2 Methodology

In this study, we proposed a dual perspective framework to fairly compare well-known RSMs calculated with different network meta information in the literature. The process contained three steps, including optimization, calculation, and comparison, to assess and rank RSMs, as shown in Figure 1.
First, we selected some typical RSMs in previous studies and established a set of multi-objective optimization models using each RSM and cost as objectives. Afterward, we solved the optimization models to get solutions on Pareto fronts by NSGA-II (Deb, Pratap, Agarwal, & Meyarivan, 2002) over 30 independent runs. Then, we adopted the fast, non-dominated sorting (Deb et al., 2002) to get the Pareto optimal solution set for each optimization model. Second, after choosing solutions from each solution set at three comparable cost levels (low, medium, and high), we calculated two scores (i.e., MRS and NRS) under two perspectives (i.e., ECP and TCP). Third, we combined the comparison perspectives and analyzed the RSM rankings to find the best RSM and identify important network meta information (Table S1) for designing RSMs in the future.

2.1 RSMs Selected for Comparison

Five RSMs are selected in the comparison framework from a dual perspective, including NRI, DSFE, REDU, PHRI, and CERI. Readers can find details in relevant studies mentioned in the Introduction.

Todini (2000) proposed the RI to describe energy transfer, including input, dissipation, and surplus power. In random disasters, a network requires more surplus energy at each demand node to improve its absorbing ability. RI ranges from 0 to 1, and higher RI means more energy and safer to deal with threats. The detailed formula can be expressed as follows.

\[ RI = \frac{\sum_{i=1}^{nn} Q_i (H_i - H_i^{req})}{\sum_{r=1}^{nr} Q_r H_r + \sum_{k=1}^{npu} \left( \frac{P_k}{\gamma} \right) - \sum_{i=1}^{nn} Q_i H_i^{req}} \]  

(1.)

Where \( nr, npu, \) and \( nn \) are the numbers of reservoirs, pumps, and demand nodes, respectively. \( Q_r \) and \( H_r \) are the discharge and the head of reservoir \( r \), respectively. \( P_k \) is the power input by pump \( k \), and \( \gamma \) is the water-specific weight. \( Q_i \) is the demand of node \( i \). \( H_i \) and \( H_i^{req} \) are the actual head and required head at node \( i \).

However, RI only considers energy redundancy, ignoring hydraulic uniformity. Prasad and Park (2004) modified it to NRI by adding a nodal uniformity coefficient, which reduces the difference in pipe diameter connecting to a junction. The following equations show how to calculate NRI.

\[ NRI = \frac{\sum_{i=1}^{nn} U_i Q_i (H_i - H_i^{req})}{\sum_{r=1}^{nr} Q_r H_r + \sum_{k=1}^{npu} \left( \frac{P_k}{\gamma} \right) - \sum_{i=1}^{nn} Q_i H_i^{req}} \]  

(2.)
\[ U_i = \frac{\sum_{j=1}^{np_i} d_{ij}}{np_i \times \max\{d_{ij}\}} \tag{3} \]

Where \( np_i \) and \( U_i \) represent the number of connected pipes and the uniformity of pipe diameters at node \( i \), respectively. \( d_{ij} \) is the pipe diameter from node \( i \) to \( j \).

Inspired by informational entropy (Shannon, 1948), FE (Tanyimboh & Templeman, 1993) was used in WDSs to follow a probability distribution to estimate the degree of uncertainty. Higher entropy means less uncertainty, making it easier to cope with unpredictable demands, random component failure, etc. The formula is as follows.

\[ FE = -\sum_{r=1}^{nr} \frac{Q_r}{T} \ln \left( \frac{Q_r}{T} \right) - \frac{1}{T} \sum_{i=1}^{nn} T_i \left[ \frac{Q_i}{T_i} \ln \left( \frac{Q_i}{T_i} \right) + \sum_{i=1}^{nps_i} \frac{q_{ij}}{T_i} \ln \left( \frac{q_{ij}}{T_i} \right) \right] \tag{4} \]

Where \( T_i \) is the sum of flow to node \( i \), and \( T \) is the total flow supplied from all sources. So, \( T_i / T \) represents the ratio of the flow of node \( i \) to the total flow of sources. \( nps_i \) means pipe set starting from node \( i \), and \( q_{ij} \) means pipe flow from node \( i \) to \( j \).

Liu et al. (2014) proposed DSFE by adding pipe velocity into FE. The only change is an additional coefficient (i.e., \( v_C / v_{ij} \)) in the last term. The equation is as follows.

\[ DSFE = -\sum_{r=1}^{nr} \frac{Q_r}{T} \ln \left( \frac{Q_r}{T} \right) - \frac{1}{T} \sum_{i=1}^{nn} T_i \left[ \frac{Q_i}{T_i} \ln \left( \frac{Q_i}{T_i} \right) + \sum_{i=1}^{nps_i} \frac{v_C}{v_{ij}} \cdot \frac{q_{ij}}{T_i} \ln \left( \frac{q_{ij}}{T_i} \right) \right] \tag{5} \]

Where \( v_C \) is the velocity constant and equals 1 m/s. \( v_{ij} \) represents the flow velocity from node \( i \) to \( j \).

Sirsant and Reddy (2020) integrated FE and RI into CERI, considering both hydraulic uniformity and energy redundancy. It was a good substitute for mechanical and hydraulic reliability in different topological networks. The equation is as follows.

\[ CERI = w_1 \cdot \left( \frac{FE}{FE_{max}} \right) + w_2 \cdot \left( \frac{RI}{RI_{max}} \right) \tag{6} \]

Where \( FE_{max} \) depends on the specific network, \( RI_{max} \) is constantly equal to 1. The sum of weights \( w_1 \) and \( w_2 \) equals 1; both vary from 0 to 1. The recommended \( w_1 \) and \( w_2 \) are equal to 0.4 and 0.6, respectively.

Liu et al. (2014) used only nodal pressure to calculate REDU, directly revealing energy redundancy. The formula is as follows.
\[
REDU = \frac{1}{nn} \sum_{i=1}^{nn} \frac{P_i - P_{\text{min},i}}{P_{\text{max},i} - P_{\text{min},i}} \tag{7.}
\]

Where \( P_i, P_{\text{min},i}, \) and \( P_{\text{max},i} \) represent the actual, minimum, and maximum pressure at node \( i \), respectively.

Liu et al. (2017) adopted a pipe hydraulic gradient between upstream and downstream nodes to calculate PHRI, which reflects on friction losses from sources to users. Less energy consumption in the upstream is required for a more reliable network. So, more available head downstream is guaranteed. The equations are as follows.

\[
PHRI = \frac{\sum_{j=1}^{np} (H_{\text{ds},j} - H_{\text{req}}) L_{\text{pro},j}}{\sum_{j=1}^{np} (H_{\text{us},j} - H_{\text{req}}) L_{\text{pro},j}} \tag{8.}
\]

\[
L_{\text{pro},j} = \sqrt{L_j^2 - (Z_{\text{us},j} - Z_{\text{ds},j})^2} \tag{9.}
\]

Where \( np \) is the number of pipes; \( H_{\text{req}} \) is the required head of the system; \( H_{\text{ds},j} \) and \( Z_{\text{ds},j} \) mean nodal head and elevation at the downstream end of pipe \( j \); \( H_{\text{us},j} \) and \( Z_{\text{us},j} \) are corresponding values at the upstream end of pipe \( j \); \( L_j \) and \( L_{\text{pro},j} \) are the actual length and projected length of pipe \( j \), respectively.

The selection of RSMs needs explanations. RI represents energy redundancy, which guarantees nodal pressure and demand. FE represents hydraulic uniformity, which reflects the available multiple supply paths. When a specific path fails, another can transport water to meet the required nodal demand (Shuang, Liu, & Porse, 2019). Liu et al. (2017) found NRI performed better than RI, MRI, and API. So, we chose NRI based on RI.

Similarly, DSFE was also chosen as an improved RSM based on FE. In addition, REDU and PHRI were chosen with energy-related measures mentioned before. The last one, CERI, represents a mixed measure of redundancy and uniformity.

### 2.2 Optimization Model and Algorithm

A multi-objective optimization model can solve the network design problem. Decision variables are different commercial pipe diameters. The objectives are minimizing cost and maximizing each RSM. The constraints contain implicit constraints, such as mass balance at nodes and energy conservation in loops, and explicit constraints, such as minimum nodal heads. The problem definition can be expressed as follows.

\[
\text{minimize: } obj_1 = \sum_{j=1}^{np} u(d_j) \times L_j \tag{10.}
\]
\[ \text{maximize: } obj_2 = RSM\#(11.) \]
\[ \text{s.t. } \begin{align*} H_i &\geq H_i^{req} \\
d_j &\in D = d_1, d_2, \ldots, d_{max} \tag{12.} \end{align*} \]

Where \( u(d_j) \) is the unit cost depending on the pipe diameter, and \( L_j \) is the length for pipe \( j \). \( D \) denotes the collection of available pipe diameters commercially.

The NSGA-II algorithm (Deb et al., 2002) can solve the optimization problem effectively and provide alternative solutions in the trade-off between cost and each RSM. Wang et al. (2015) have proved that it uses less computation and finds a Pareto front closer to the true one than other algorithms such as AMLAGAM, Borg, etc. Some adjustable parameters influence optimization results, including population size, number of function evaluations, probability of Simulated Binary Crossover (SBX), probability of Polynomial Mutation (PM), and distribution indices of SBX and PM. Fine-tuning these parameters will approach the actual Pareto front (Wang, Wang, et al., 2019).

### 2.3 Assessment and Comparison of RSMs

#### 2.3.1 The Dual Perspectives

We evaluated various RSMs based on a dual-perspective framework. Specifically, the equal-cost perspective (ECP) and the traditional consistency perspective (TCP) were employed to assess and compare RSMs. The ECP and the TCP correspond to the equal-cost solution set and the entire optimal solution set, respectively. The reliability calculation utilized the MRS (Wang, Huang, et al., 2019); meanwhile, the resilience calculation followed the GRA framework (Diao et al., 2016; Meng et al., 2018). The GRA framework obtains stress-strain curves for pipe failure, excess water demand, and contaminant intrusion. In our study, we only focused on the pipe failure mode that is the most widely studied (Leštáková, Logan, Rehm, Pelz, & Friesen, 2023), namely comparing RSMs only at the mechanical reliability and resilience domains.

#### 2.3.2 Resilience Calculation

A detailed procedure was presented here for calculating resilience under ECP, as shown in Figure 2.
Figure 2. Flowchart of calculating resilience under ECP.

Step 1: Select equal-cost solutions.

The Pareto optimal solution set was obtained after 30 independent optimization runs and by the fast, non-dominated sorting. At the design stage, decision-makers favor more economical network solutions. So, we chose equal-cost (i.e., low, medium, and high) solutions within the cost overlap intervals on the Pareto front of RSMs. The medium cost is the mean of the low and the high. The equal-cost solutions fall at the nearest point where the Pareto front intersects the equal-cost line.

Step 2: Run GRA.

For selected solutions, we added an actual water demand pattern from the KY network in the United States (Jolly, Lothes, Bryson, & Ormsbee, 2014) to the hydraulic model, which enabled the GRA under extended period simulation. The three-hour peak demand period (19:00 to 22:00) was chosen for pipe failure simulation (Diao et al., 2016). The calculations were performed separately on each solution with the same demand pattern. Six strain indicators were calculated from the network performance curve, including TTS, DUR, FR, RR, MAG, and SEV (Meng et al., 2018). In multiple pipe failure scenarios, the strains depend on the random locations of burst pipes. After running GRA, we obtained three stress-strain curves for each strain indicator among solutions.

Step 3: Find CSR.

Pareto optimal solution set
Choose solutions on three vertical cost lines
Cost levels of low, medium, high
Equal-cost solution set
Node demand pattern
Set up extended period simulation
Network performance curve
Run GRA on pipe failure mode
Stress-strain curve
Consider all strain indicators
Stress-NRS curve in CSR
Radar chart of strain indicators
Average score at each strain
CSR of equal-cost solution set
Mean stress-strain curves
Comparable Stress Range
Critical Stress
Stress-NRS curve in CSR
Average NRS at each stress
NRS of equal-cost solution set
Three stress-strain curves were obtained by calculating the solution set, but only the mean curve was used to continue afterward analyses. For low, medium, or high costs, we obtained five mean curves of RSMs, which showed a convergence phenomenon. As stress increases, the initial divergent strains gradually converge into the same one (Figure 2). Consequently, the solution set has a CSR where the strains of RSMs can be distinguished to compare.

The stress and strain are denoted as $\sigma_{ij}$ and $\epsilon_{ij}$, where $i$ equals the number of failed pipes and $j$ represents each strain indicator. The relative range ($rr$) was used to describe the convergence degree of a numerical series by the minimum, the mean, and the maximum (Fekete, Vörösmarty, Roads, & Willmott, 2004; Hongyu et al., 2021). When it is below or equal to 5%, it indicates sufficient convergence within the dataset (Morgan, 2018). The formulas are as follows.

$$rr\{\epsilon_{ij}\}_{RSM} = \frac{\max\{\epsilon_{ij}\}_{RSM} - \min\{\epsilon_{ij}\}_{RSM}}{\text{mean}\{\epsilon_{ij}\}_{RSM}} \times 100\% \tag{13.}$$

$$\begin{cases} rr\{\epsilon_{ij}\}_{RSM} \geq 5\%, & 0 < \sigma_{ij} \leq \sigma_{cj} \\ rr\{\epsilon_{ij}\}_{RSM} < 5\%, & \sigma_{cj} < \sigma_{ij} \leq 100\% \end{cases} \tag{14.}$$

$$\sigma_c = \min\{\sigma_{cj}\}_{cost} \tag{15.}$$

Where $\{\epsilon_{ij}\}_{RSM}$ represents a strain set of RSMs at the same cost level. $\sigma_{ij}$ means the percentage of burst pipes to a total number of pipes, ranging from 0 to 1. $\sigma_{cj}$ equals the critical (maximum)stress within CSR of strain indicator $j$. The relative range falls below 5% when the stress exceeds $\sigma_{cj}$. $\{\sigma_{cj}\}_{cost}$ is a critical stress set for different strain indicators and cost levels. $\sigma_c$ is the minimum to reflect the overlapping CSR (i.e., including $c$ stresses) for comparison. We excluded $\sigma_{cj}$ less than 5%, meaning no comparative significance for strain $j$ (e.g., TTS).

Step 4: Get NRS.

Next, we calculated resilience from strain to stress to a single solution. Since we got the strain values, they can be converted to strain scores. The radar chart shows five strain scores at each stress of CSR, ranging from 0 to 100. The higher strain means poorer network resilience, so high strain values correspond to low scores. The strain score of 0 is at the critical stress. The equations are as follows.

$$SS_{ij} = 100 \times \left(1 - \frac{\epsilon_{ij}}{\epsilon_{cj}}\right) \tag{16.}$$

$$\text{NRS}_i = \frac{1}{5} \sum_{j=1}^{5} \frac{SS_{ij}}{100} \tag{17.}$$
\[ NRS = \frac{1}{c} \sum_{i=1}^{c} NRS_i \]  #(18.)

Where \( \varepsilon_{ij} \) and \( \varepsilon_{cj} \) represent the values of the strain at specific stress \( i \) and critical stress \( c \), respectively. \( SS_{ij} \) is the score within 0 to 100 of the strain at the stress \( i \). \( NRS_i \) equals the resilience score at stress \( i \). \( NRS \) averages the resilience scores within the CSR for a single solution, ranging from 0 to 1.

However, the four-step process was different under TCP. We run GRA among the entire set of Pareto optimal solutions. Then, we found a new CSR within the optimal set, regardless of cost levels and strain indicators. Finally, the calculation of NRS was similar to ECP.

### 2.3.3 Reliability Calculation

We also employed MRS (Wang, Huang, et al., 2019) to calculate the mechanical reliability of individual solutions under both ECP and TCP. The MRS for single-pipe failure may be insufficient if a network has many loops. So, we use the meshed-ness coefficient (MC) (Yazdani & Jeffrey, 2011), reflecting the loop degree of the network, to determine the maximum number of burst pipes (i.e., \( f_{\text{max}} \)). The equations (19~22) consider the number of computational scenarios; the equations (23~25) calculate mechanical reliability based on the scenarios.

\[
\begin{align*}
  f_{\text{max}} &= \begin{cases} 
    1, & 0 < MC \leq 1 \\
    2, & 1 < MC \leq 2 \\
    3, & 2 < MC \leq 3 \\
    4, & MC > 3 
  \end{cases} \tag{19.} \\
  MC &= \frac{np - nn + 1}{2 \cdot nn - 5} \tag{20.} \\
  s &= \begin{cases} 
    np, & f_{\text{max}} = 1 \\
    np + s_f, & f_{\text{max}} \geq 2 
  \end{cases} \tag{21.} \\
  s_f &= \sum_{f=2}^{f_{\text{max}}} \left[ \frac{p(1-p) \sum_{i=1}^{np} C_{np}^i}{p(1-p) + (\sum_{i=1}^{np} C_{np}^i - 1) \left( \frac{CI}{Z_\alpha} \right)^2 + 2(np - f + 1)} \right] \tag{22.} \\
  DS_i &= \begin{cases} 
    1, & P_i \geq P_{\text{min},i} \\
    \frac{P_i}{P_{\text{min},i}}, & 0 < P_i < P_{\text{min},i} \tag{23.} \\
    0, & P_i \leq 0 
  \end{cases} \\
  MRS &= \frac{1}{s} \sum_{j=1}^{s} \sum_{i=1}^{nn} w_i \cdot DS_i \tag{24.}
\end{align*}
\]
Where \( f \) is the number of failed pipes, \( DS_i \) is the demand score at node \( i \), depending on the nodal actual pressure \( P_i \) and the required \( P_{\text{min},i} \). \( w_i \) is the proportion of water demand at node \( i \). \( s \) is the number of failure scenarios for calculation, including at least single-pipe failure (i.e., \( np \)). However, it may contain multiple pipe failures (i.e., \( s_f \)). The number of scenarios \( s_f \) refers to the GRA framework. The simulation success probability \( p \) is equal to 0.5. At 95% confidence level, \( CI \) and \( Z_{\alpha} \) equal ±5% and 1.96, respectively.

### 3 Case Studies

#### 3.1 Benchmark Networks

We chose two benchmark WDSs to assess and rank NRI, DSFE, REDU, PHRI, and CERI. The network layouts are shown in the Supporting Information (Figure S1).

The first well-known case is the Hanoi network (HAN) from Fujiwara and Khang (1990), which contains 34 pipes organized in three loops, 31 demand nodes, and one reservoir with a fixed head of 100 m. The Hazen-Williams roughness coefficients of all pipes are equal to 130. All nodes require a head above the ground elevation at least 30 m. Six commercial pipe sizes range from 304.8 mm to 1016.0 mm. Therefore, the problem has a vast search space equal to \( 6^{34} \approx 2.87 \times 10^{26} \) discrete combinations.

The second case is a real-world network adapted from a town named Fossolo (FOS) in Italy (Bragalli, D'Ambrosio, Lee, Lodi, & Toth, 2012). The network includes 58 pipes, 36 demand nodes, and one reservoir with a fixed head of 121.0 m. All the polyethylene pipes have a considerably high roughness coefficient of 150. All the demand nodes have the same minimum pressure head requirement equal to 40 m, but each node has a different maximum pressure head. In addition, the flow velocity in each pipe is enforced to be no more than 1 m/s. There are 22 commercially available diameters ranging from 16.0 mm to 409.2 mm; thus, the search space equals \( 22^{58} \approx 7.25 \times 10^{77} \) discrete combinations.

#### 3.2 Simulation of Pipe Failure Scenarios

For single-source networks like HAN and FOS, we did not consider the failure of pipes connected to the water source directly, which would inevitably lead to the breakdown of the entire system (Wang et al., 2019a). Under pipe failures, we run pressure-dependent demand hydraulic simulation instead of demand-driven to obtain more accurate nodal demand (L. Jun & Guoping, 2013). The valves were assumed to be installed following the N-rule, meaning each pipe has two valves at both ends and can ultimately be isolated from the network (H. Jun & Loganathan, 2007).

To calculate MRS, a snap-shot simulation was run without a demand pattern. Due to fewer loops in HAN, we only simulated single-pipe failure scenarios. More pipe failure scenarios
(i.e., one to four pipe failures) were simulated in FOS to compare RSMs. To calculate NRS, we add the demand pattern of KY cases (Jolly et al., 2014) to make an extended period simulation. Then, we run GRA to simulate failure scenarios of a single pipe, multiple, or all pipes. Under ECP, absolute values of MRS and NRS are used to compare RSMs. Under TCP, we calculated the Spearman correlation coefficient for MRS-Cost and NRS-Cost.

3.3 Experimental Setup

For HAN and FOS networks, optimization models were established for each RSM. To eliminate the effect of a random initial population, 30 optimization runs were carried out to get an aggregate set of Pareto front solutions. Then, we used the fast non-dominated sorting (Kalyanmoy, 2002) to generate the Pareto optimal solution set of each RSM. The intersected points of the cost lines and the Pareto fronts are chosen as equal-cost solutions (Figure 3). The WNTR (Klise, Bynum, Moriarty, & Murray, 2017) was used to run hydraulic simulations under scenarios, and the Pymoo (Blank & Deb, 2020) was used to complete the optimization calculation.

**Figure 3.** Pareto fronts and equal-cost lines for two networks: (a) HAN and (b) FOS. In HAN, low = 6.75 M$, medium = 7.25 M$, high = 7.75 M$. In FOS, low = 0.06 M€, medium = 0.18 M€, high = 0.30 M€.

NSGA-II’s parameters need to be fine-tuned (Wang, Wang, et al., 2019) to improve the optimal solution set’s quality. We shared some standard parameters for both HAN and FOS: the population size of 100, the SBX probability of 0.9, the SBX distribution index of 20, and the PM distribution index of 20. The PM probabilities relate to the number of pipes in two cases, equal to 1/34 and 1/58, respectively. Due to different search spaces, the generations were set to be 500 and 1000 for HAN and FOS, respectively.
4 Results and Discussion

4.1 Comparison under ECP

In cases of HAN and FOS, the equal-cost solution sets of the five RSMs were calculated by MRS and NRS. Under ECP, we made the comparison in three scopes of the strain, the stress, and the solution for NRS, but only the solution scope for MRS. So, we mainly focused on resilience calculation in Section 0.

After running GRA to get three stress-strain curves (Figure S2), we obtained the CSR by the mean curve. In HAN and FOS, the CSR equals 3.125~15.625% and 1.75%~40.35%, respectively (Figure 4). A broader CSR links to a higher loop degree of the network. We considered only the stress within the CSR to compare RSM under ECP.

Figure 4. The CSR for evaluating RSMs in two networks: (a) HAN and (b) FOS. Considering cost levels and strain indicators, the grey area is the overlapping CSR for a single network.

Critical stress varied with strain indicators and cost levels in both cases. The TTS was excluded due to the critical stress lower than 5%. Simultaneously, RR and DUR had the highest and lowest critical stress in the two networks, respectively, but at different cost levels. The CSR depends on DUR. Since RR has high critical stress, it may be an essential aspect of resilience to compare RSMs, which should be addressed.

4.1.1 The Strain Scope

We calculated the remaining five strain indicators’ scores at three cost levels, as shown in the radar charts in the Supporting Information (Figure S3, S4). All the strain indicators on the same radar chart facilitate the resilience comparison among RSMs. The larger the area enclosed by the pentagon, the better the corresponding RSM performed. Regardless of costs and stresses, the strain scores decreased simultaneously in five dimensions without crossovers. This reflects that the five strains are harmonized in the downward resilience trend. However, the decline
degrees of strains vary as the stress increases. The decline in DUR was the most pronounced, while SEV was the least pronounced.

4.1.2 The Stress Scope

After averaging the five strain scores, we got the NRS at each stress within the CSR, as shown in Figure 5. The higher the NRS, the better the system performance. However, the decline trends were slightly different in HAN and FOS.

![Figure 5](image_url)

**Figure 5.** The decline of NRS within CSR at three cost levels in HAN (subplots a, b, c) and FOS (subplots d, e, f).

In HAN, the extent of CSR is small (between 3.125% and 15.625%) due to fewer loops, resulting in a nearly linear decline. Under single pipe failure, the NRS of all RSM solutions is close to 1 at the three cost levels. As the stress increases to the critical one (i.e., 15.625%), the average NRS of five RSMs at three cost levels decrease to 0.26, 0.39, and 0.39, respectively.

The higher loop degree in FOS makes a larger CSR (between 1.754% and 40.351%), resulting in a nonlinear decline, especially at medium-cost and high-cost levels. The NRS decreases slowly at a low-stress interval. However, once the stress exceeds about 10%, the NRS declines rapidly and shows evident differences among the RSM solutions. At three cost levels, the average NRS of RSM solutions are almost one under single-pipe failure. As the stress reaches the critical one (40.351%), the average NRS of RSMs at three cost levels decrease to
0.23, 0.33, and 0.41, respectively. As the cost level increases, NRS significantly improves near the critical stress.

The stress-resilience curves from different RSMs have the overlap phenomenon, meaning some RSMs perform almost the same at a specific cost. Specifically, considering the high cost of HAN, we observed the overlap of two groups: (1) NRI, CERI, and REDU, and (2) DSFE and PHRI. In FOS, the stress-resilience curves of REDU and PHRI overlap at all three cost levels. This phenomenon reveals the similarity in the resilience of RSM solutions when the cost is equal (or comparable).

Furthermore, the stress-resilience curves have another surpassing phenomenon. Generally, a better RSM means the curve is situated in a higher position. At equal cost, the relative positions of RSMs remain unchanged. However, the cost increase leads to the phenomenon: one RSM’s curve is lower than another RSM’s curve at a lower cost, but it becomes higher at a higher cost.

In HAN, the REDU’s curve (green line in Figure 5a) is slightly lower than the PHRI’s curve (yellow line in Figure 5a) at a low cost but surpasses it at a medium cost, and the gap widens at a high cost. The NRI’s curve is lower than the CERI’s curve at low and medium costs but surpasses it at a high cost.

In FOS, we can observe that the overlapping curves of REDU and PHRI are lower than CERI’s curve at a low cost, close to it at a medium cost, and surpass it at a high cost. Surprisingly, the position of CERI decreases as the cost increases in FOS, indicating the inconsistency in CERI between resilience and cost.

The resilience improvement of RSM solutions differs as the cost increases. One RSM, which initially has a lower NRS but improves fast, may surpass another RSM with a higher NRS but improves slowly. This phenomenon also suggests that it is a challenging task to compare various RSMs unbiasedly. Therefore, a holistic viewpoint and a fair comparison framework are significant.

4.1.3 The Solution Scope

In HAN and FOS, the NRS was averaged by stresses within CSR. Then, we ranked the RSMs by the NRS at three cost levels. Additionally, we employed the MRS for a complementary comparison perspective. Relative deviations were used to quantify the gap of RSMs, as shown in the Supporting Information (Table S2 and Table S3). We found that the relative deviations of RSMs by NRS are more significant than by MRS. The ranking results were presented in the histogram (Figure 6), where the mean line represents the average value of MRS or NRS across all RSMs.
Figure 6. ECP results of MRS and NRS in HAN (subplots a, c) and FOS (subplots b, d).

Overall rankings of RSMs show consistency in the same network regarding reliability (measured by MRS) and resilience (measured by NRS). The NRS highlights the performance differences (i.e., more relative deviation) than MRS among RSMs. However, the ranking by NRS is similar to that by MRS because both consider demand satisfaction rate. MRS can effectively differentiate among RSMs. Since NRS was calculated in extended period simulation by more pipe failure scenarios, the computational burdens surpassed significantly than MRS. However, the overall rankings of RSMs by MRS and NRS are inconsistent in the two networks. As the cost increases, both rankings of MRS and NRS certainly change. The overall ranking is closer at a lower cost and differs from that at a higher cost.
The NRI exhibits excellent performance at different cost levels by both calculation methods. Several studies also proved (Sirsant & Reddy, 2020) that NRI is a better indicator, especially within medium-sized networks. The CERI performs well in HAN but poorly at medium and high costs in FOS, revealing it is sensitive to network topology. The rankings of REDU and PHRI are similar but not as good as those of NRI and CERI. The DSFE performs worst, which is not recommended for future research.

### 4.2 Comparison under TCP

Besides assessing various RSMs using selected solutions within CSR, it is necessary to explore their performance overall with Pareto’s optimal solutions. Under TCP, we calculated the MRS and NRS of all solutions. Then, the Spearman correlation coefficients were calculated to assess the consistency of MRS-Cost and NRS-Cost.

As shown in Figure 7, the TCP focuses on the upward trend of RSM consistency curves, while the ECP focuses on the relative positions of curves. The dual perspectives are complementary for comparison. A superior RSM should have two characteristics: (1) a higher position on the y-axis (i.e., higher MRS and NRS); (2) a consistent upward trend on the curve (i.e., increasing costs can improve MRS and NRS).
The NRI occupies the highest position in the two networks and demonstrates a prominent upward trend, having the two characteristics mentioned above. Conversely, DSFE takes the lowest position and suffers a fluctuating upward trend, indicating poorer performance. The remaining RSMs fall between these two extremes. We also observed similarities in the performance of two groups of indicators: (1) CERI and NRI; (2) PHRI and REDU. This could be attributed to these measures containing similar network meta information. We will further discuss this aspect in Section 0.

Specially, traditional consistency can be flawed. In FOS, the Spearman’s coefficient of NRI is even lower than REDU under TCP. That is because NRI performance is almost constant at minimal costs (i.e., 0.15 M€). Maintaining the curve horizontally affects the correlation coefficient and may result in poorer coefficients. However, if the position of the curve is higher...
than others, the corresponding RSM should be considered better. Combining ECP to compare RSMs is essential because this perspective directly reflects the performance level of the solution.

4.3 Comparison under the Dual Perspectives

As shown in Table 1, the final rankings based on the HAN and FOS cases are consistent with previous studies. Liu et al. (2017) discovered a ranking series of “NRI ≻ PHRI ≻ DSFE” based on uncertain scenarios involving node demands and pipe failures. Wang et al. (2019) proposed a high-dimensional optimization method to generate some solutions of RSMs, and calculated the Spearman coefficients to compare more measures, finding “NRI ≻ REDU ≻ PHRI ≻ DSFE.” Sirsant and Reddy (2020) concluded that both CERI and NRI were ideal measures in a network of small or medium size, considering the relationship between RSMs and hydraulic and mechanical reliability.

<table>
<thead>
<tr>
<th>RSMs</th>
<th>HAN ECP MRS</th>
<th>HAN TCP MRS</th>
<th>FOS ECP MRS</th>
<th>FOS TCP MRS</th>
<th>Mean Rank</th>
<th>Final Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>NRI</td>
<td>2 2</td>
<td>1 1</td>
<td>1 1</td>
<td>1 2</td>
<td>1.375</td>
<td>1</td>
</tr>
<tr>
<td>DSFE</td>
<td>5 5</td>
<td>5 5</td>
<td>5 5</td>
<td>4 4</td>
<td>4.750</td>
<td>5</td>
</tr>
<tr>
<td>REDU</td>
<td>3 3</td>
<td>4 4</td>
<td>3 3</td>
<td>3 1</td>
<td>3.000</td>
<td>3</td>
</tr>
<tr>
<td>PHRI</td>
<td>4 4</td>
<td>3 3</td>
<td>4 4</td>
<td>2 3</td>
<td>3.375</td>
<td>4</td>
</tr>
<tr>
<td>CERI</td>
<td>1 1</td>
<td>2 2</td>
<td>2 2</td>
<td>5 5</td>
<td>2.500</td>
<td>2</td>
</tr>
</tbody>
</table>

The dual perspectives are a more reasonable framework to compare RSMs. The rankings of RSMs from two perspectives are consistent. This suggests the existence of a best RSM under both ECP and TCP. However, comparing RSMs using the ECP may be superior to the TCP. Different network cases can influence the ranking under TCP, which is confusing for decision-makers. Ranking under ECP is more stable in different networks. The choice of computational method (MRS or NRS) does not significantly impact RSM ranking, but NRS requires more computational overhead. Therefore, we recommend using ECP+MRS to compare RSMs.

4.4 Importance of Network Meta Information

The question, “How can we improve RSMs effectively?” has troubled the academic community for decades. A reasonable conjecture is that it is related to the information required to
compute RSMs from WDSs. The complex network model can be mathematically understood as
the coupling of topological and hydraulic information (Giustolisi, Ridolfi, & Simone, 2019).

Some hydraulic information requires no simulation, such as nodal required demand, pipe
length, and diameter. Others need simulation to extract, such as actual nodal heads and pipe
flows, which can only be obtained through simulation. The number of nodes and pipes is the
basis of topological indicators, such as loops, connectivity density, and grid coefficient. Then, we
listed some network meta information in the Supporting Information (Table S1). The RSM final
ranking was compared with the network meta information, as shown in Table 2.

Table 2. Final ranking of RSMs compared with network meta information.

<table>
<thead>
<tr>
<th>RSMs</th>
<th>Final Rank</th>
<th>Node Information</th>
<th>Link Information</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>(mn) (Q) (P(H))</td>
<td>(np) (q) (d) (v) (L)</td>
</tr>
<tr>
<td>NRI</td>
<td>1</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>CERI</td>
<td>2</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>REDU</td>
<td>3</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>PHRI</td>
<td>4</td>
<td>√</td>
<td></td>
</tr>
<tr>
<td>DSFE</td>
<td>5</td>
<td>√</td>
<td></td>
</tr>
</tbody>
</table>

Note: \(mn\) = number of nodes; \(Q\) = required or actual demand; \(P(H)\) = pressure (head); \(np\) =
number of pipes; \(q\) = flowrate; \(d\) = diameter; \(v\) = velocity; \(L\) = length.

We explored the relationship between RSM rankings and the network meta information.
Among five RSMs, NRI performed the best, as its computation involved a comprehensive set of
information, including the node count, nodal head (pressure), water demand, pipe count, and pipe
diameter. Among this information, the nodal head is significant as it directly represents the
energy redundancy of the network. The measure DSFE does not utilize a nodal head, which
results in poor mechanical reliability and resilience performance. The performance differences
reflect some vital information to design the RSMs.

In order to represent network performance as accurately as possible, we should
effectively combine topological and hydraulic information due to the coupling characteristic.
However, individual pieces of information may overlap in reflecting the same aspect of network
performance. Therefore, it is essential to identify the key information with minimal overlap. By
comparing the ranking, we observed that DSFE performed the poorest because of a lack of nodal
head information, indicating that nodal head is one of the key factors. As for the other measures,
PHRI ranks second only to DSFE due to the absence of node count information, which is also
indispensable. It is worth noting that REDU only utilizes two pieces of information, nodal head
and node count, which is fewer than DSFE and PHRI. However, REDU outperformed DSFE and PHRI by capturing more crucial information. NRI and CERI incorporate these two critical pieces of information while considering additional information. The information discrepancies between the two measures include pipe diameter and flow rate. It suggests that the pipe diameter information is more critical, as proved in Table 2.

Jeong and Kang (2020) proposed a measure called hydraulic uniformity index (HUI), whose calculation only relies on pipe information such as the pipe count, flowrate, length, and head loss, but without any node information. We have tested HUI under the same analytical framework proposed in this study and found that it failed to achieve a Pareto front over a wide range of costs, with its solutions concentrated around a specific cost value. Existing studies find that nodal hydraulic information is vital in resilience evaluation for WDSs (Yu, Wu, Zhou, & Liu, 2023). The exclusion of the node information is unreasonable because it discards crucial information to evaluate network performance. Although NRI performed nearly flawlessly, we observed instances along the Pareto front where the cost increased while the performance decreased, indicating the potential lack of certain link information in NRI.

Based on the discrepancies between RSM ranking and network meta information, we can prioritize the importance of network information as follows, “nodal head ≻ node demand ≻ pipe diameter ≻ pipe flowrate.” The node information is more critical than the pipe information. Due to limitations inherent to the selected indicators, it is currently challenging to determine the relative importance of certain information, and we consider them approximately equal. The selected RSMs primarily utilize information including node count, demand, head or pressure, and pipe count. In future development of new and improved RSMs, it is recommended to consider combining key meta-information to approach network resilience.

5 Conclusions

This paper investigated the relationship between various RSMs and the system reliability and resilience of WDSs. To this end, we proposed a novel dual-perspective framework, including TCP and ECP, for benchmarking RSMs. Under TCP, RSMs were evaluated based on the consistency between costs and system performance regarding MRS and NRS. In contrast, under ECP, RSMs’ MRS and NRS were assessed from the decision-makers’ preference, focusing on the cost-benefit analysis. Five well-known RSMs were selected from the literature, including NRI, REDU, PHRI, DSFE, and CERI, representing the hydraulic energy redundancy based, hydraulic uniformity based, and hybrid RSMs. These RSMs were compared in the multi-objective optimal design of two medium-scale, single-source WDSs, i.e., HAN and FOS. The main conclusions are drawn as follows.
(1) Benchmarking RSMs under the dual perspective, i.e., from ECP and TCP, is significant in comparative studies. ECP reflects the practical preference of decision-makers, while TCP demonstrates the intrinsic consistency between an RSM and cost. ECP extends and complements TCP by emphasizing that an increase in cost investment should enhance the reliability and resilience performance of WDSs. Moreover, ECP amplifies the differences between MRS and NRS, which is beneficial in comparative studies; however, this aspect may be overlooked under TCP.

(2) Within the GRA framework, CSR plays a key role in ranking RSMs, which facilitates the comparison among RSMs intuitively and reasonably. Within CSR, it is vital to examine the variation characteristics of NRS associated with RSMs. A less fluctuating magnitude with a larger NRS indicates better performance. Beyond CSR, a convergence phenomenon occurs, where RSMs’ performance becomes similar, making it difficult and unreasonable to distinguish them.

(3) Regarding effectiveness and efficiency, computing MRS under ECP is a more appropriate way to compare RSMs. For HAN and FOS networks, the rankings of RSMs in MRS and NRS are reasonably consistent. However, resilience calculation considers more pipe failures and a more extended simulation period to assess dynamic performance, requiring huge computational overheads. Despite MRS or NRS, ECP gives a more consistent ranking of RSMs than TCP. Moreover, the ranking under ECP is less sensitive to network topological structure.

(4) We obtained the final ranking, “NRI > CERI > REDU > PHRI > DSFE,” by the proposed dual-perspective framework. Among the five RSMs, NRI remains a preferable choice, making it a valuable tool for decision-making in WDS design and optimization. Usually, the performance of RSM solutions will improve as costs increase from low to high. However, it can be observed that network performance may decrease with an increasing cost (e.g., CERI for FOS). Therefore, decision-makers must select a consistent and reliable indicator like NRI to measure the performance of their networks.

(5) Last but not least, we identified the relationship between the competency of RSMs and the network meta information used in calculating these RSMs (i.e., topological and hydraulic information of WDSs). By linking the final ranking of RSMs with the meta information, we established the relative importance of these information, “nodal head > nodal demand > pipe diameter > pipe flowrate.” In short, nodal information proved to be more critical than pipe information. Consequently, a superior RSM stems from a reasonable combination of the abovementioned meta information and the specific formula to calculate. It also implies that using combined RSMs like CERI may be problematic as fixed weights between RI and FE are prone to deteriorate when dealing with highly looped WDSs.
The present study has certain limitations that can be improved in the future. Case selection has a significant impact on the performance of RSMs. In this study, we used two single-source, medium-scaled WDSs. Some researchers found that entropy-based metrics like FE had superior performance in multiple water source networks (Prasad & Tanyimboh, 2008) and large-scale networks (Sirsant & Reddy, 2020). Therefore, it is crucial to consider more network cases with varying topological features. It is also necessary to incorporate cases with complex hydraulic components, including pumps, valves, and DMAs.

Acknowledgments

The work reported was funded by the Guangdong Basic and Applied Basic Research Foundation (Grant No: 2022A1515011179 & 2023A1515030126).

Open Research

The benchmark design problems of water distribution systems to compare reliability surrogate measures can be found at the website of Centre for Water Systems, University of Exeter (https://www.exeter.ac.uk/research/centres/cws/resources/benchmarks/pareto). The used benchmark input files, strain radar charts, stress-strain curve charts derived from this study are available on GitHub (https://github.com/Boater4/WDSResilience). Other code and data from this study are available from the authors upon request.

References


**Step A: Optimization**

A1. Select RSMs for comparison
A2. Set up optimization model for each RSM
A3. Solve the model 30 times with NSGA-II
A4. Get Pareto optimal solution set by fast non-dominated sorting

**Step B: Calculation**

- Pareto optimal solution set under TCP
- Equal-cost solution set under ECP
- Calculate network loop degree
- Determine the number of failed pipes
- Add network demand pattern
- Run GRA under pipe failures
- Calculate MRS within failure range
- Calculate NRS within CSR

**Step C: Comparison**

C1. Compare RSMs under ECP
C2. Compare RSMs under TCP
C3. Compare RSMs under the dual perspectives
C4. Identify effective RSMs and key network meta information
1. Select equal-cost solutions.
2. Run global resilience analysis.
3. Find comparable stress range.
4. Get normalized resilience score.

**Pareto optimal solution set**

Choose solutions on three vertical cost lines.

**Network performance curve**

**Stress-strain curve**

**Stress-NRS curve in CSR**

**CSR of equal-cost solution set**

**NRS of equal-cost solution set**

**Radar chart of strain indicators**

Average score at each strain

Average NRS at each stress

**Node demand pattern**

Set up extended period simulation

Consider all strain indicators

**Mean stress-strain curves**

Comparable Stress Range

Critical stress

Relative range < 5%

**Equal-cost line**

**Mean Strain**

**Network performance curve**

**Stress-strain curve**

**Network performance curve**

**Stress-strain curve**

**Node demand pattern**

Set up extended period simulation

Consider all strain indicators

**Mean stress-strain curves**

Comparable Stress Range

Critical stress

Relative range < 5%

**Equal-cost line**

**Mean Strain**

**Network performance curve**

**Stress-strain curve**

**Node demand pattern**

Set up extended period simulation

Consider all strain indicators

**Mean stress-strain curves**

Comparable Stress Range

Critical stress

Relative range < 5%

**Equal-cost line**

**Mean Strain**

**Network performance curve**

**Stress-strain curve**

**Node demand pattern**

Set up extended period simulation

Consider all strain indicators

**Mean stress-strain curves**

Comparable Stress Range

Critical stress

Relative range < 5%

**Equal-cost line**

**Mean Strain**

**Network performance curve**

**Stress-strain curve**

**Node demand pattern**

Set up extended period simulation

Consider all strain indicators

**Mean stress-strain curves**

Comparable Stress Range

Critical stress

Relative range < 5%

**Equal-cost line**

**Mean Strain**

**Network performance curve**

**Stress-strain curve**

**Node demand pattern**

Set up extended period simulation

Consider all strain indicators

**Mean stress-strain curves**

Comparable Stress Range

Critical stress

Relative range < 5%

**Equal-cost line**

**Mean Strain**

**Network performance curve**

**Stress-strain curve**

**Node demand pattern**

Set up extended period simulation

Consider all strain indicators

**Mean stress-strain curves**

Comparable Stress Range

Critical stress

Relative range < 5%

**Equal-cost line**

**Mean Strain**

**Network performance curve**

**Stress-strain curve**

**Node demand pattern**

Set up extended period simulation

Consider all strain indicators

**Mean stress-strain curves**

Comparable Stress Range

Critical stress

Relative range < 5%

**Equal-cost line**

**Mean Strain**

**Network performance curve**

**Stress-strain curve**

**Node demand pattern**

Set up extended period simulation

Consider all strain indicators

**Mean stress-strain curves**

Comparable Stress Range

Critical stress

Relative range < 5%

**Equal-cost line**

**Mean Strain**

**Network performance curve**

**Stress-strain curve**

**Node demand pattern**

Set up extended period simulation

Consider all strain indicators

**Mean stress-strain curves**

Comparable Stress Range

Critical stress

Relative range < 5%

**Equal-cost line**

**Mean Strain**

**Network performance curve**

**Stress-strain curve**

**Node demand pattern**

Set up extended period simulation

Consider all strain indicators

**Mean stress-strain curves**

Comparable Stress Range

Critical stress

Relative range < 5%

**Equal-cost line**

**Mean Strain**

**Network performance curve**

**Stress-strain curve**

**Node demand pattern**

Set up extended period simulation

Consider all strain indicators

**Mean stress-strain curves**

Comparable Stress Range

Critical stress

Relative range < 5%

**Equal-cost line**

**Mean Strain**

**Network performance curve**

**Stress-strain curve**

**Node demand pattern**

Set up extended period simulation

Consider all strain indicators

**Mean stress-strain curves**

Comparable Stress Range

Critical stress

Relative range < 5%

**Equal-cost line**

**Mean Strain**

**Network performance curve**

**Stress-strain curve**

**Node demand pattern**

Set up extended period simulation

Consider all strain indicators

**Mean stress-strain curves**

Comparable Stress Range

Critical stress

Relative range < 5%

**Equal-cost line**

**Mean Strain**

**Network performance curve**

**Stress-strain curve**

**Node demand pattern**

Set up extended period simulation

Consider all strain indicators

**Mean stress-strain curves**

Comparable Stress Range

Critical stress

Relative range < 5%

**Equal-cost line**

**Mean Strain**

**Network performance curve**

**Stress-strain curve**

**Node demand pattern**

Set up extended period simulation

Consider all strain indicators

**Mean stress-strain curves**

Comparable Stress Range

Critical stress

Relative range < 5%

**Equal-cost line**

**Mean Strain**

**Network performance curve**

**Stress-strain curve**

**Node demand pattern**

Set up extended period simulation

Consider all strain indicators

**Mean stress-strain curves**

Comparable Stress Range

Critical stress

Relative range < 5%
Exploring the Puzzles in Comparing Reliability Surrogate Measures for Water Distribution System Design: From the Dual Perspectives of Equal-Cost and Consistency

Yuehua Huang, Qianqian Zhou, Qi Wang, Zhiwei Zhao

1School of Civil and Transportation Engineering, Guangdong University of Technology, Guangzhou 510006, China
2Cross Research Institute of Ocean Engineering Safety and Sustainable Development, Guangdong University of Technology, Guangzhou 510006, China

Contents of this file

Text S1
Figures S1 to S6
Tables S1 to S4

Additional Supporting Information (Files uploaded separately)

No additional supporting information.

Introduction

Figure S1 are the layouts of two single-source medium-scale networks to compare RSMs in the dual-perspective framework. Figure S2 shows global resilience analysis (GRA) (Diao et al., 2016) for obtaining stress-strain curves. Figure S3 and Figure S4 show the score radar chart of five strain indicators in HAN and FOS, respectively, evaluating RSMs in three costs at minimum stress; at medium stress; at maximum stress of CSR. Within CSR, a set of radar charts containing five strains (excluding TTS) were used to show and compare the network resilience of different RSMs. As the stress increases, the radar area decreases, reflecting the degradation of the network resilience. Here, we presented the snapshots (i.e., minimum, medium, maximum stress within CSR) of resilience degradation processes. Figure S5 and Figure S6 show the Spearman correlation coefficients between the cost and MRS; the cost and NRS in HAN and FOS, respectively.
Table S1 shows some key network meta information for calculating RSMs. Table S2 and Table S3 show ranking of RSMs calculated by MRS and NRS under ECP for HAN and FOS, respectively. Table S4 shows HAN’s and FOS’s rankings of RSMs by MRS and NRS under TCP.

Text S1 explains some strain indicators used in the GRA framework.

Text S1. Strain indicators in the GRA framework.

The strain indicators come from network performance curve, which depicts the absorptive, restorative, and adaptive capability facing external stresses (Shuang, Liu, & Porse, 2019). Meng et al. (2018) have summarized six strain indicators: time to strain (TTS), duration (DUR), failure rate (FR), recovery rate (RR), magnitude (MAG), and severity (SEV) (Figure S2).

The network performance can be usually represented by the network service availability (also called demand satisfaction rate) (Cassottana, Aydin, & Tang, 2021; Creaco & Franchini, 2012; Jung, Kang, Kim, & Lansey, 2014; Zhao, Chen, & Gong, 2015; Zhuang, Lansey, & Kang, 2013), which is the ratio of the actual system demand to the expected system demand. The normal, threshold, and minimum are three crucial demand levels when pipes burst. The threshold can be defined as 80% of normal water supply (Cassottana et al., 2021; Sirsant & Reddy, 2020), below which the entire network is regarded as breakdown.

As shown in Figure S2, the minimum corresponds to the worst performance moment during the pipe failure period, specifically at T₂. The difference between the threshold and minimum represents the MAG of the failure event. We assume that some pipes are assumed to burst at time T₀. Between T₀ and T₁, the network exhibits a high absorptive capability and withstands disturbances, with performance ranging between the normal and threshold. This time period is referred to as the TTS. The network begins to fail at T₁. From T₁ to T₃, the network performance remains below the threshold, indicating the DUR of network failure. The area, enclosed by the network performance curve and the threshold, represents the SEV of the failure event. The rate at which the network performance decreases from the threshold to the minimum during T₁ to T₂ is known as the FR, while the rate at which it recovers from the minimum to the threshold during T₂ to T₃ is termed the RR. After implementing the GRA within the network solution set, we got three stress-strain curves (e.g., min, max and mean) for each strain indicator.
**Figure S1.** Layout of each benchmark network. (a) HAN; (b) FOS.

**Figure S2.** The process of running GRA.
Figure S3. Score of five strain indicators in HAN for evaluating RSMs in three costs at minimum stress (subplots a, b, c); at medium stress (subplots d, e, f); at maximum stress (subplots g, h, i) of CSR.
Figure S4. Score of five strain indicators in FOS for evaluating RSMs in three costs at minimum stress (subplots a, b, c); at medium stress (subplots d, e, f); at maximum stress (subplots g, h, i) of CSR.
Figure S5. Spearman coefficients between cost and MRS; between cost and NRS in HAN. (a) NRI; (b) DSFE; (c) REDU; (d) PHRI; (e) CERI.
**Figure S6.** Spearman coefficients between cost and MRS; between cost and NRS in FOS. (a) NRI; (b) DSFE; (c) REDU; (d) PHRI; (e) CERI.

<table>
<thead>
<tr>
<th>Table S1. Network meta information for calculating RSMs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Topological Information</td>
</tr>
<tr>
<td>Node Information</td>
</tr>
<tr>
<td>node count</td>
</tr>
<tr>
<td>Required demand, actual demand, head and pressure</td>
</tr>
</tbody>
</table>
### Table S2. HAN’s ranking of RSMs calculated by MRS and NRS under ECP.

<table>
<thead>
<tr>
<th>RSMs</th>
<th>Low Cost</th>
<th>Medium Cost</th>
<th>High Cost</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MRS Rank</td>
<td>RD</td>
<td>MRS Rank</td>
<td>RD</td>
</tr>
<tr>
<td>NRI</td>
<td>0.907</td>
<td>2 0.7%</td>
<td>0.922</td>
<td>2 0.9%</td>
</tr>
<tr>
<td>DSFE</td>
<td>0.885</td>
<td>5 -1.7%</td>
<td>0.886</td>
<td>5 -3.0%</td>
</tr>
<tr>
<td>REDU</td>
<td>0.897</td>
<td>3 -0.3%</td>
<td>0.914</td>
<td>3 0.0%</td>
</tr>
<tr>
<td>PHRI</td>
<td>0.896</td>
<td>4 -0.5%</td>
<td>0.909</td>
<td>4 -0.5%</td>
</tr>
<tr>
<td>CERI</td>
<td>0.916</td>
<td>1 1.8%</td>
<td>0.937</td>
<td>1 2.6%</td>
</tr>
</tbody>
</table>

Note: Relative deviation (RD) reflects the gap between the mean in a group of values and a certain value.

### Table S3. FOS’s ranking of RSMs calculated by MRS and NRS under ECP.

<table>
<thead>
<tr>
<th>RSMs</th>
<th>Low Cost</th>
<th>Medium Cost</th>
<th>High Cost</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MRS Rank</td>
<td>RD</td>
<td>MRS Rank</td>
<td>RD</td>
</tr>
<tr>
<td>NRI</td>
<td>0.673</td>
<td>2 4.8%</td>
<td>0.741</td>
<td>2 4.0%</td>
</tr>
<tr>
<td>DSFE</td>
<td>0.576</td>
<td>5 -10.3%</td>
<td>0.640</td>
<td>5 -10.2%</td>
</tr>
<tr>
<td>REDU</td>
<td>0.632</td>
<td>4 -1.6%</td>
<td>0.718</td>
<td>3 0.9%</td>
</tr>
<tr>
<td>PHRI</td>
<td>0.642</td>
<td>3 -0.1%</td>
<td>0.698</td>
<td>4 -2.1%</td>
</tr>
<tr>
<td>CERI</td>
<td>0.689</td>
<td>1 7.3%</td>
<td>0.765</td>
<td>1 7.4%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>RSMs</th>
<th>Low Cost</th>
<th>Medium Cost</th>
<th>High Cost</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MRS Rank</td>
<td>RD</td>
<td>MRS Rank</td>
<td>RD</td>
</tr>
<tr>
<td>NRI</td>
<td>0.990</td>
<td>1 4.1%</td>
<td>0.999</td>
<td>2 0.7%</td>
</tr>
<tr>
<td>DSFE</td>
<td>0.909</td>
<td>5 -4.4%</td>
<td>0.973</td>
<td>5 -1.9%</td>
</tr>
<tr>
<td>REDU</td>
<td>0.936</td>
<td>4 -1.6%</td>
<td>0.996</td>
<td>3 0.4%</td>
</tr>
<tr>
<td>PHRI</td>
<td>0.938</td>
<td>3 -1.4%</td>
<td>0.999</td>
<td>1 0.7%</td>
</tr>
<tr>
<td>CERI</td>
<td>0.982</td>
<td>2 3.3%</td>
<td>0.994</td>
<td>4 0.2%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>RSMs</th>
<th>Low Cost</th>
<th>Medium Cost</th>
<th>High Cost</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MRS Rank</td>
<td>RD</td>
<td>MRS Rank</td>
<td>RD</td>
</tr>
<tr>
<td>NRI</td>
<td>0.784</td>
<td>1 19.6%</td>
<td>0.902</td>
<td>1 14.4%</td>
</tr>
<tr>
<td>DSFE</td>
<td>0.551</td>
<td>5 -16.0%</td>
<td>0.668</td>
<td>5 -15.3%</td>
</tr>
<tr>
<td>REDU</td>
<td>0.595</td>
<td>3 -9.2%</td>
<td>0.792</td>
<td>3 0.5%</td>
</tr>
<tr>
<td>PHRI</td>
<td>0.589</td>
<td>4 -10.2%</td>
<td>0.800</td>
<td>2 1.5%</td>
</tr>
<tr>
<td>CERI</td>
<td>0.758</td>
<td>2 15.7%</td>
<td>0.780</td>
<td>4 -1.1%</td>
</tr>
</tbody>
</table>
Table S4. HAN’s and FOS’s rankings of RSMs by MRS and NRS under TCP.

<table>
<thead>
<tr>
<th>RSMs</th>
<th>HAN $\rho_{MRS}$</th>
<th>Rank</th>
<th>HAN $\rho_{NRS}$</th>
<th>Rank</th>
<th>FOS $\rho_{MRS}$</th>
<th>Rank</th>
<th>FOS $\rho_{NRS}$</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>NRI</td>
<td>0.996</td>
<td>1</td>
<td>0.995</td>
<td>1</td>
<td>0.986</td>
<td>1</td>
<td>0.971</td>
<td>2</td>
</tr>
<tr>
<td>DSFE</td>
<td>0.833</td>
<td>5</td>
<td>0.957</td>
<td>5</td>
<td>0.829</td>
<td>4</td>
<td>0.832</td>
<td>4</td>
</tr>
<tr>
<td>REDU</td>
<td>0.969</td>
<td>4</td>
<td>0.958</td>
<td>4</td>
<td>0.950</td>
<td>3</td>
<td>0.975</td>
<td>1</td>
</tr>
<tr>
<td>PHRI</td>
<td>0.980</td>
<td>3</td>
<td>0.961</td>
<td>3</td>
<td>0.963</td>
<td>2</td>
<td>0.953</td>
<td>3</td>
</tr>
<tr>
<td>CERI</td>
<td>0.992</td>
<td>2</td>
<td>0.990</td>
<td>2</td>
<td>0.401</td>
<td>5</td>
<td>0.143</td>
<td>5</td>
</tr>
</tbody>
</table>

Note: $\rho_{MRS}$ = Spearman correlation coefficient between cost and MRS; $\rho_{NRS}$ = Spearman correlation coefficient between cost and NRS.