How Are Various Surrogate Indicators Consistent with Mechanical Reliability of Water Distribution Systems: From a Perspective of Many-Objective Optimization

Qi Wang [1] , **Wen Huang [1]**, **Libing Wang [1]**, **Yongshi Zhan [2]**, **Zhihong Wang [1]**, **Xu Wang [3,*]**, **Xuyi Zhan [4] and Shuming Liu [4,*]**

[1]   School of Civil and Transportation Engineering, Guangdong University of Technology, Guangzhou 510006, China
[2]   International Department, the Affiliated High School of South China Normal University, Guangzhou 510000, China
[3]   State Key Laboratory of Simulation and Regulation of Water Cycle in River Basin, China Institute of Water Resources and Hydropower Research, Beijing 100038, China
[4]   School of Environment, Tsinghua University, Beijing 100084, China
*   Correspondence: wangxu_04@126.com (X.W.); shumingliu@tsinghua.edu.cn (S.L.); Tel.: +86-10-6878-5503 (X.W.); +86-10-6278-7964 (S.L.)

Abstract: Exploring the trade-off between cost and system reliability of water distribution systems (WDSs) has been focused for two decades. Due to the intensive computation associated with the reliability analysis, it is popular in the research community to replace this procedure by using surrogate indicators. However, the discussion on the correlation among different types of such indicators is generally lacking, which implies that a deeper understanding of this aspect is needed. This paper proposes a novel methodology of investigating the relationships among many commonly used surrogate indicators for measuring the mechanical reliability of WDSs. In particular, the optimal design of WDSs is formulated as a many-objective optimization problem, using cost and each surrogate indicator as an individual goal. Two benchmark design problems of different scales and complexities are considered for verifying the proposed method. The well-known multi-objective evolutionary algorithm (MOEA), namely Borg that is suitable for coping with problems involving many objectives, is used to obtain the best approximation to the Pareto-optimal fronts for both cases. Afterward, the one-pipe burst testing is conducted to quantify the correlation between mechanical reliability and surrogate indicators. Results suggest that investigating the correlation of surrogate indicators from the perspective of many-objective optimization provides a direct and efficient way of distinguishing better indicators from worse ones. Resilience-based surrogate indicators and the Redundancy indicator that only depends on nodal pressures are highly related to the mechanical reliability of WDSs. In contrast, entropy-based indicators exhibit poor performance in reflecting the mechanical reliability. These insights contribute to the selection of more appropriate surrogate indicators for the optimal design of WDSs for researchers and practitioners.

Keywords: water distribution system; mechanical reliability; resilience; surrogate indicators; many-objective optimization

1. Introduction

Water Distribution Systems (WDSs), which are regarded as "lifeline engineering" along with other utilities (e.g., gas, electricity), may be the most critical infrastructures of modern cities. Proper

operation of WDSs means supplying adequate water to various types of users at the adequate quantity and good quality. However, water utilities are vulnerable to a wide range of human-caused and natural disasters, such as pipe breakage due to improper operation, firefighting, and earthquake [1]. As such, it is imperative for water utility managers and decision-makers to take the reliability of the network into consideration during the design stage.

Reliability has been the most prevailing indicator for measuring the performance of WDSs. It is generally defined as the probability of a given system to continuously meet the demand of its customers at the predefined service level (mostly by nodal pressure) under varying conditions [2]. However, the incorporation of this indicator directly into the optimal design model, either as a constraint or an individual objective, would inevitably be computationally intensive [3]. Therefore, a series of surrogate indicators, which are simple yet effective to represent system reliability, have been proposed in the literature during the past two decades. Table 1 provides an incomplete list of surrogate indicators which have been frequently used in WDSs. These surrogate indicators can be generally categorized into three groups: (1) resilience-based; (2) entropy-based; and (3) hybrid.

Table 1. Surrogate indicators for measuring WDS reliability in the literature.

Year	Authors	Surrogate Indicators	Type	Ref.
2000	Tanyimboh & Templeman	flow entropy (FE)	(2)	[4]
2000	Todini	resilience index (RI)	(1)	[5]
2004	Prasad & Park	network resilience (NRI)	(1)	[6]
2008	Jayaram & Srinivasan	modified resilience index (MRI)	(1)	[7]
2010	Raad, et al.	mixed reliability surrogate (FE + RI)	(3)	[8]
2011	Wu, et al.	surplus power factor	(1)	[9]
2011	Tanyimboh, et al.	statistical entropy	(2)	[10]
2014	Jung, et al.	robustness index	-	[11]
2014	Liu, et al.	diameter sensitive flow entropy (DSFE)	(2)	[12]
2014	Liu, et al.	Redundancy	-	[13]
2016	Creaco, et al.	RI + loop diameter uniformity	(3)	[14]
2016	Liu, et al.	available power index (API)pipe hydraulic RI (PHRI)	(1)	[15]
2018	Bin Mahmoud & Piratla	probabilistic resilience index (PRI)	(1)	[16]

There also exist limited comparative studies on the correlation between various surrogate indicators and system reliability of WDSs [3,8,10,16–18]. However, the conclusions of these works are somehow contradictory to each other, which may confuse the research community. For instance, Raad, et al., made a regression analysis on the Pareto-optimal sets obtained from the multi-objective design of WDS benchmarks. They concluded that the resilience-based surrogate measures (specifically the RI and NRI) are more practical reliability surrogate indicators than FE, because they have better performance on reflecting the system reliability under demand uncertainty and pipe failure conditions [8]. Later, Tanyimboh, et al., compared four surrogate indicators, including statistical entropy, NRI, RI, and MRI. By using the prototype software that could produce more realistic results for pressure-deficient WDSs, they claimed that statistical entropy outperformed RI by demonstrating a stronger correlation with failure tolerance [10].

Recently, some improved surrogate indicators have been continuously proposed in the literature. Liu, et al., proposed a new surrogate indicator, called diameter-sensitive flow entropy (DSFE), to overcome the disadvantages of typical FE [12]. DSFE was found to have a stronger correlation with reliability measures.

Later, Liu, et al., developed two new energy-related indices, termed available power index (API) and pipe hydraulic resilience index (PHRI) and found that energy-based indices generally outperform DSFE under demand and pipe failure uncertainties [15]. Bin Mahmoud and Piratla developed a new resilience-based index, termed probabilistic resilience index (PRI), and found that PRI has the best performance in the low to moderate cost range for all the benchmarks considered in their study [16].

With the gradually improved insights into system performance, it is envisaged that more suitable surrogate indicators may be proposed to account for the reliability and resilience of the network system under various types of conditions. Therefore, it is crucial to correctly understand the correlation between different indicators and adequately use them to guide the design and operation of WDSs. However, existing methods mainly rely on the reliability assessment conducted after optimization to compare the suitability of those surrogate indicators. This would require intensive post-optimization runs, thus impractical to be extended to larger networks in the real world. In other words, a more efficient methodology for evaluating these indicators may be appealing.

This paper aims to provide an innovative way of understanding the correlation of various surrogate indicators existing in the literature. In particular, how the many-objective optimization technique can be used to reveal the complex relationships among cost, surrogate indicators, and system reliability is demonstrated. On top of that, recommendations are made on how to facilitate the optimal design of WDSs in future works by selecting appropriate surrogate indicators.

2. Methodology

The proposed methodology includes six steps, as shown in Figure 1. In the first step, the many-objective optimization model is established, in which the cost objective is minimized, and all the surrogate indicators for measuring WDS reliability are maximized. This formulation enables to identify the correlation of various surrogate indicators in a single optimization run. Next, the well-known Borg algorithm [19] is applied to the above many-objective optimization model, and carry out multiple independent runs for each case study to eliminate the impact of randomness involved in the initialization of population. In the third step, the best approximation to the Pareto-optimal front is generated by comparing all the non-dominated solutions from multiple runs. In particular, these solutions are first aggregated as a whole dataset, and the fast non-dominated sorting procedure [20] is then used to screen the most approximated Pareto front. Afterward, the single-pipe burst testing is implemented on each solution in the Pareto set at the fourth step. This is aimed at evaluating the impact of pipe burst on a specific WDS configuration. Then, the mechanical reliability of each candidate WDS design solution is assessed by quantifying to what extent the system can maintain the normal water supply (i.e., meeting the minimum pressure head requirements at demanding nodes). The mechanical reliability is expressed as a weighted score, as shown in the following subsection. At the last step, the correlation among cost, mechanical reliability of WDSs, and various surrogate indicators is illustrated. To this end, a novel, intuitive metric for evaluating the consistency of a dataset with multiple attributes is developed.

2.1. Selected Surrogate Indicators

A total of seven surrogate indicators from Table 1 are chosen for assessing their correlation with capital costs and mechanical reliability of WDSs. The reasons for selecting these indicators are as follows. First, existing surrogate indicators can be generally categorized into two groups: resilience-based and entropy-based. Therefore, four resilience-based indicators and two entropy-based indicators are picked as they are frequently used in previous studies [3,8,15,17]. Second, the Redundancy indicator is also chosen [13], which is quite simple in terms of calculation compared to other ones. Each of these indicators is briefly explained below, including its definition and associated equation(s). Readers are referred to relevant papers for greater detail.

The resilience index (RI), first introduced by Todini, is arguably the first resilience-based surrogate indicator in the domain of WDSs to account for the reliability of a looped network [5].

Herein, the concept of "resilience" refers to the capability of a network system to maintain its functionality under stress or failure conditions. RI can adequately reflect (i.e., strongly related to) the inherent attribute of the system to overcome failures. Furthermore, it avoids the statistical analysis of various types of uncertainty, which is generally considered for calculating reliability.
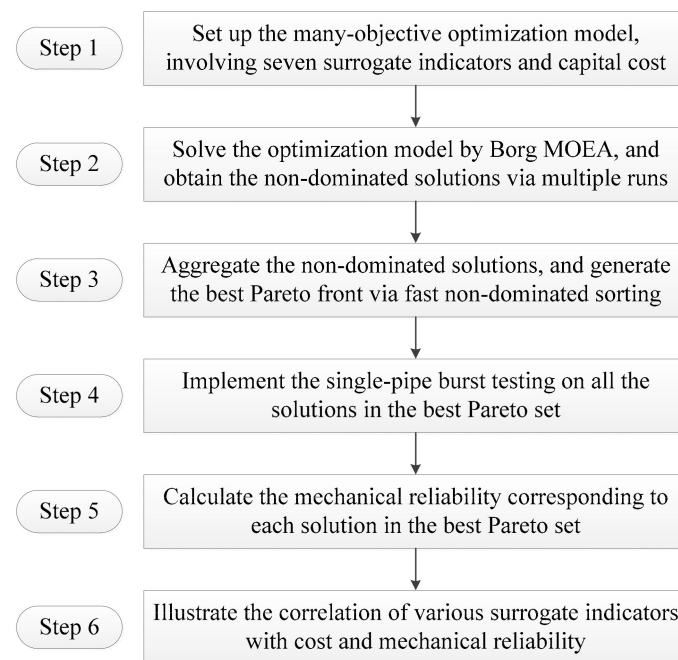
Step 1	Set up the many-objective optimization model, involving seven surrogate indicators and capital cost
Step 2	Solve the optimization model by Borg MOEA, and obtain the non-dominated solutions via multiple runs
Step 3	Aggregate the non-dominated solutions, and generate the best Pareto front via fast non-dominated sorting
Step 4	Implement the single-pipe burst testing on all the solutions in the best Pareto set
Step 5	Calculate the mechanical reliability corresponding to each solution in the best Pareto set
Step 6	Illustrate the correlation of various surrogate indicators with cost and mechanical reliability

Figure 1. Flowchart of the proposed methodology.

The RI indicator is derived from the law of energy conservation, where the total available power injected into a given WDS is dissipated in pipes (due to pipe wall friction resistance) and delivered to demand nodes (users). Equation (1) shows the mathematical expression of RI, which is defined as the ratio of surplus power at demand nodes to the maximum power that would be dissipated internally to overcome pipe resistance. For a given topology of WDS, the increase of RI improves its energetic redundancy, which means a decrease of the internal energy dissipation.

$$RI = \frac{\sum\limits_{i=1}^{nn} q_i(H_i - h_i)}{\sum\limits_{j=1}^{nr} Q_j H_j + \sum\limits_{k=1}^{npu} (P_k/\gamma) - \sum\limits_{i=1}^{nn} q_i h_i} \tag{1}$$

where q_i, H_i, and h_i are the nodal demand, the actual head, and the required head at node i, respectively, Q_j and H_j are the discharge and the head at reservoir j, P_k is the energy provided by pump k, γ is the specific weight of water, and nn, nr, and npu are the number of demand nodes, reservoirs, and pumps in the given network, respectively.

Later, Prasad and Park proposed the network resilience index (NRI) (see Equation (2)) by adding a term named nodal uniformity (see Equation (3)) into the original expression of RI [6]. It can be viewed as a weighted RI by taking the combined effect of both surplus power and pipe size uniformity into account. Using NRI may ensure reliable loops in the design stage as the diameters of pipes connected to each node would not vary widely.

$$NRI = \frac{\sum\limits_{i=1}^{nn} C_i q_i(H_i - h_i)}{\sum\limits_{j=1}^{nr} Q_j H_j + \sum\limits_{k=1}^{npu} \frac{P_k}{\gamma} - \sum\limits_{i=1}^{nn} q_i h_i} \tag{2}$$

$$C_i = \frac{\sum\limits_{j=1}^{np_i} d_j}{np_i \times \max\{d_j\}} \tag{3}$$

where C_i represents the uniformity of node i, np_i is the number of pipes connected to node i, and d_j is the diameter of pipe j connected to node i.

Based on the energy conservation law, Liu, et al., proposed a new energy-related index, called available power index (API), for reliability evaluation (see Equation (4)) [15]. They divide the instantaneous input power, which is derived from reservoirs, tanks, or pumps, into available and unavailable power. The available power refers to the energy delivered at demand nodes. A WDS with higher reliability can be obtained by increasing the available power, implying the additional capacity to resist disturbances imposed by unexpected events (e.g., pipe failure or demand fluctuations). Therefore, API is conceptually analogous to RI.

$$API = \frac{\sum\limits_{i=1}^{nn} Q_i H_i}{\sum\limits_{j=1}^{nr} Q_j H_j + \sum\limits_{k=1}^{npu} \frac{P_k}{\gamma} + \sum\limits_{t=1}^{nt} Q_t H_t} \tag{4}$$

where Q_i and H_i are the actual flow and head at node i, Q_t is equal to the flow into (negative) or out of (positive) tank t, H_t is the elevation of the free surface at tank t, and nt is the number of tanks in the network.

The pipe hydraulic resilience index (PHRI) is another energy-related index introduced by Liu, et al. [15]. It is based on the calculation of areas of triangles along the hydraulic grade line (HGL) within the network system. A series of triangles are formed by the characteristic points along the HGL and the line of the minimum required hydraulic head. No matter how the input power is injected (either from sources with higher elevations or pumps), the HGL always declines along the pipes from sources to users due to friction losses. Therefore, if upstream pipes dissipate less energy, the downstream pipes will have more available heads. As such, PHRI is defined to be the ratio of the accumulated triangular areas associated with the available downstream heads (denoted as A_j in Equation (5)) to that associated with the total upstream head (denoted as $A_j + B_j$ in Equation (5)). It is worth noting that this new indicator also takes the impact of pipe length into account, as each triangular area is derived from the product of hydraulic head and projected pipe length. By maximizing PHRI, a network system can increase its capability of coping with the perturbation caused by uncertainty. For the space limit, readers are referred to [15] for the diagrams of computing PHRI.

$$PHRI = \frac{\sum\limits_{j=1}^{np} A_j}{\sum\limits_{j=1}^{np} (A_j + B_j)} \tag{5}$$

$$A_j = \frac{1}{2}(H_{ds,j} - H_{req})L_{pro,j} \tag{6}$$

$$A_j + B_j = \frac{1}{2}(H_{us,j} - H_{req})L_{pro,j} \tag{7}$$

$$L_{pro,j} = \sqrt{L_j^2 - (Z_{us,j} - Z_{ds,j})^2} \tag{8}$$

where np is the number of pipes in the network, $H_{ds,j}$ and $H_{us,j}$ are heads at downstream and upstream nodes of pipe j, respectively, L_j and $L_{pro,j}$ are the length and projected length of pipe j, and $Z_{ds,j}$ and $Z_{us,j}$ are the elevations at downstream and upstream nodes of pipe j, respectively.

Tanyimboh and Templeman held the viewpoint that a WDS which is designed to carry maximum entropy flows is generally reliable to cope with the considerable amount of uncertainty (e.g., demand growth and variations, random bursts and component failures). They proposed the equations [4] to

quantify the flow entropy (FE) of WDSs based on Shannon's entropy form in the context of information theory [21].

$$S = S_0 + \sum_{i=1}^{n} \frac{T_i}{T} S_i \tag{9}$$

$$S_0 = -\sum_{j \in IN} \frac{Q_j}{T} \ln\left(\frac{Q_j}{T}\right) \tag{10}$$

$$S_i = -\frac{Q_j}{T} \ln\left(\frac{Q_j}{T}\right) - \sum_{ij \in ND_i} \frac{q_{ij}}{T_i}\left(\frac{q_{ij}}{T_i}\right) \tag{11}$$

$$\frac{S}{K} = -\sum_{j \in IN} \frac{Q_j}{T} \ln\left(\frac{Q_j}{T}\right) - \frac{1}{T}\sum_{i=1}^{n} T_i\left(\frac{Q_i}{T_i} \ln\left(\frac{Q_i}{T_i}\right) + \sum_{ij \in ND_i} \frac{q_{ij}}{T_i} \ln\left(\frac{q_{ij}}{T_i}\right)\right) \tag{12}$$

where S_0 is the entropy of sources, S_i is the entropy of node i, T_i is the total flow reaching node i, T is the total flow provided by all sources, T_i/T is the ratio of the accumulative flow reaching at node i to the total flow provided by all sources, K is an arbitrary positive constant, usually set to 1, IN is the number of source or input nodes, Q_j is the inflow at source node j, Q_i is the demand at node i, q_{ij} is the pipe flow from node i to node j, and ND_i is the set of pipes in which water leaves from node i.

Liu, et al., proposed the diameter sensitive flow entropy (DSFE) indicator to extend the original definition of FE by considering the impact of pipe diameter on reliability [12]. As shown in Equation (13), the main structure of DSFE is identical to FE except by introducing an additional coefficient (i.e., $\frac{C}{V_{ij}}$) before the last term.

$$\frac{S}{K} = -\sum_{j \in IN} \frac{Q_j}{T} \ln\left(\frac{Q_j}{T}\right) - \frac{1}{T}\sum_{i=1}^{n} T_i\left(\frac{Q_i}{T_i} \ln\left(\frac{Q_i}{T_i}\right) + \sum_{ij \in ND_i} \frac{C}{V_{ij}} \times \frac{q_{ij}}{T_i} \ln\left(\frac{q_{ij}}{T_i}\right)\right) \tag{13}$$

where C is an arbitrary velocity constant, e.g., 1 m/s; V_{ij} is the velocity in pipe ij.

Besides the resilience-based and entropy-based surrogate indicators, some researchers also introduced other simple indicators to reflect the reliability of WDSs, of which the Redundancy indicator is such a metric [13]. Redundancy only depends on the actual and required pressure heads (both minimum and maximum) associated with each node, as shown in Equation (14). The range of Redundancy is between 0 and 1, with a number closer to 1 indicating more surplus head (more reliable) across the whole network.

$$Redundancy = \sum_{i=1}^{nn} \frac{P_i - P_i^{\min}}{P_i^{\max} - P_i^{\min}} / nn \tag{14}$$

where P_i is the pressure at node i, and $P_i^{\min}$ and $P_i^{\max}$ are the minimum and maximum pressure heads that node i is allowed for.

2.2. Many-Objective Optimization Model

As previously mentioned, this paper proposes an innovative way of identifying the correlation among various surrogate indicators of system reliability. This is achieved by formulating the design of WDSs as a many-objective optimization problem, as shown in Equation (15). In particular, the proposed many-objective optimization model is aimed at minimizing the total capital cost and maximizing each surrogate indicator as defined in Equations (1)–(14). There are two types of constraints associated with the optimization model, which are the conservation of mass and energy across the network and the minimum pressure head requirements at demanding nodes, respectively.

The former type of constraints is automatically satisfied by using the EPANET2 software [22], while the accumulated pressure head deficits across the network are calculated and used to judge the feasibility of solutions. The optimization algorithm eventually maintains only feasible solutions

that all demand nodes have sufficient pressure heads. Additionally, the decision variables considered in this paper only involve pipe diameters, which are restricted to available discrete sizes from the relevant markets.

$$\begin{cases} \min imize : Objective1 = Cost = \sum_{i=1}^{np} U(D_j) \times L_j \\ \max imize : Objective2 = RI \\ \max imize : Objective3 = NRI \\ \max imize : Objective4 = API \\ \max imize : Objective5 = PHRI \\ \max imize : Objective6 = \text{Re}dundancy \\ \max imize : Objective7 = FE \\ \max imize : Objective8 = DSFE \\ s.t. H_i > h_i \forall i \in [1,n] and D_j \in [D_1, D_2, \ldots, D_m] \end{cases} \tag{15}$$

where $U(D_j)$ is the unit cost of pipe j depending on its selected size D_j, the relationship between $U(D_j)$ and D_j is usually non-linear and case by case. D_m denotes the maximum discrete pipe size available in the market. It is worth noting that the implicit constraints related to mass balance at nodes and energy conservation in loops are considered by using the hydraulic solver and thus omitted here.

Compared with the traditional multi-objective optimization model, which usually involves two or three goals, the many-objective optimization model (i.e., more than three goals) can identify more complex trade-offs among different objectives in the hyper Pareto space. The direct benefit of using the many-objective optimization model is to obtain the relationships among all eight objectives in a single run. More importantly, this formulation can intuitively reflect the intrinsic consistency of selected surrogate indicators, which is the primary concern of this study.

2.3. Borg MOEA

The number of non-dominated solutions in a many-objective optimization problem tends to increase exponentially as the number of objectives expands. This so-called "dominance resistance" [19] usually challenges the capabilities of MOEAs to cope with problems involving many objectives. Traditional MOEAs, such as the well-known NSGA-II [20], cannot deal with such problems efficiently due to the weakness of the Pareto dominance concept. In particular, as the existence of "dominance resistance", the selection pressure of traditional MOEAs declines rapidly throughout optimization, which results in the premature of population. Thus, advanced algorithms that can effectively overcome "dominance resistance" are required to solve many-objective optimization problems.

Borg MOEA is one of such tools which are suitable to solve many-objective optimization problems [19]. As a powerful and popular optimization tool in the domain of evolutionary computation, Borg MOEA has many advanced features and functionalities. However, elaborating how it works is out of the scope of this paper. Therefore, readers are referred to [19] for greater detail. Only its three key features that are superior to traditional MOEAs are briefly described below.

First, it is based on the static MOEAs framework and uses the ε-dominance concept [23] to maintain the convergence and diversity concurrently throughout the search. This implies that the best solutions are stored and updated in an external archive with higher computational efficiency. Second, the population size changes adaptively according to the volume of the external archive. When potential stagnation is identified, Borg will restart the search according to two types of criteria (i.e., ε-progress triggered restart and population-to-archive ratio triggered restart). Third, it simultaneously employs a variety of recombination operators for the offspring generation, which enhances Borg's performance on a wide range of problem domains. Additionally, Borg is found to have large regions of high-performing parameterizations in terms of so-called sweet spots [24], which is an appealing feature for solving many-objective optimization problems.

Borg has been successfully applied to dealing with challenging, many-objective, real-world problems in the domain of water resources, of which a detailed review can be found in [25].

2.4. Post Analysis of System Reliability due to Mechanical Failures

There are generally two types of failure, namely, mechanical and hydraulic failure [3]. In this paper, only the impact of mechanical failure was taken into account. Specifically, the single-pipe burst testing was carried out to evaluate the consequences of pipe breakage throughout the network. Coincident pipe bursts were not considered as their possibilities would be much lower than that of the single-pipe burst event. For a single source network, each pipe was enumerated but with the one(s) directly connected to the reservoir excluded during the pipe burst testing, since the breakage of those pipe(s) would inevitably result in the out-of-service of the entire system. For a multi-source network, all pipes were enumerated in the pipe burst testing. Each pipe was first closed, followed by the hydraulic simulation and the retrieval of pressure heads at demand nodes during each burst event. This iterative process continued until all candidate pipes were processed. Then, the corresponding mechanical reliability score (MRS) was calculated by Equations (16)–(19) for each non-dominated solution in the Pareto set.

$$MRS = \frac{\sum_{j=1}^{cp} TotalScore_j}{cp} \tag{16}$$

$$TotalScore_j = \sum_{i=1}^{nn} \omega_i \times Score_i \tag{17}$$

$$Score_i \begin{cases} 1 & p_i \geq P_{\min} \\ \frac{p_i}{p_{\min}} & 0 < p_i < p_{\min} \\ 0 & p_i \leq 0 \end{cases} \tag{18}$$

$$\omega_i = \frac{q_i}{\sum_{i}^{nn} q_i} \tag{19}$$

where cp is the number of candidate pipes in the single-pipe burst testing. $TotalScore_j$ is the total score achieved at a pipe burst event j, which is calculated as the weighted sum of *Score* at each demand node. $Score_i$ is the score evaluated at demand node i, which depends on the relationship between the actual nodal pressure P_i and the minimum pressure head requirement $P_{\min}$ (as shown in Equation (19)). ω_i is the weighting coefficient of demand node i that is equal to the ratio of water demand at node i to the total demand of the network.

Illustrating the Pareto front of a many-objective optimization problem is not as easy as that when dealing with the problem with two or three objectives. As such, it is fashionable to use the parallel coordinate plot (PCP) to demonstrate the complex variations among many objectives (usually conflicting with each other) [26]. Recall that the critical question of this study is to investigate to what extent various surrogate indicators can reflect the reliability of WDSs to account for failure scenarios. Therefore, it is essential to quantify the consistency among reliability, cost, and surrogate indicators. To this end, an intuitive way of computing the level of intrinsic consistency of any two or more metrics is proposed.

Figure 2a shows the procedure for calculating the consistency value in pseudo-code format; while Figure 2b illustrates how the variable "*accordance*" is accumulated depending on the two types of sorting used in that procedure. Briefly speaking, every two solutions in the Pareto set are selected to form a temporary matrix *M*. The number of columns in *M* is equal to the number of metrics considered in the calculation of their consistency value. *M* is sorted based on two criteria, one is the ascending order of the first column (denoted as m_1), and the other is the ascending order of each column.
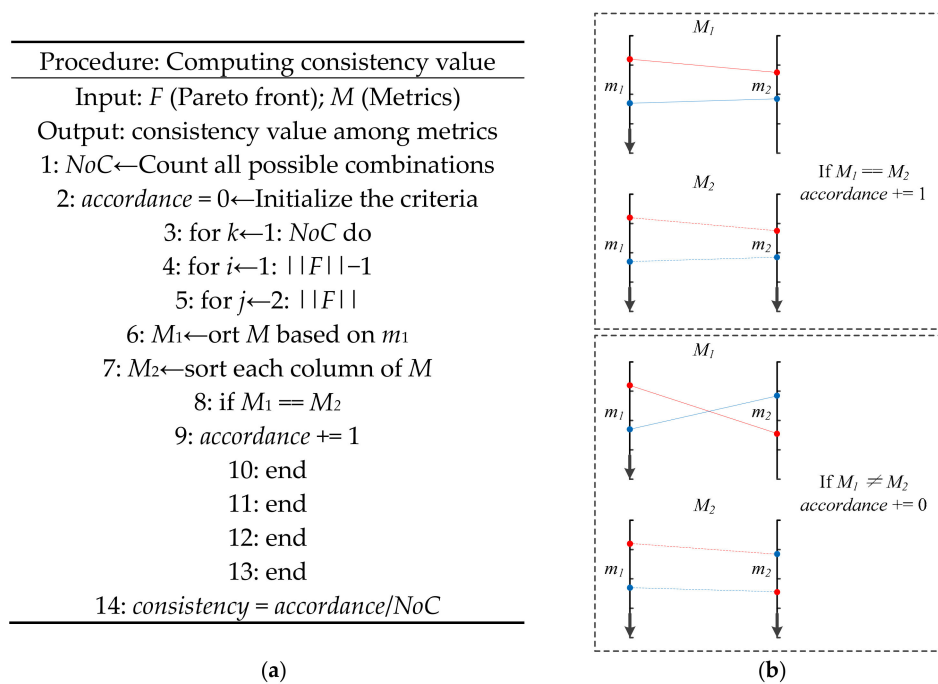
Procedure: Computing consistency value
Input: F (Pareto front); M (Metrics)
Output: consistency value among metrics
1: $NoC \leftarrow$ Count all possible combinations
2: *accordance* = 0 $\leftarrow$ Initialize the criteria
3: for $k \leftarrow 1$: NoC do
4: for $i \leftarrow 1$: $||F||-1$
5: for $j \leftarrow 2$: $||F||$
6: $M_1 \leftarrow$ ort M based on m_1
7: $M_2 \leftarrow$ sort each column of M
8: if $M_1 == M_2$
9: *accordance* += 1
10: end
11: end
12: end
13: end
14: *consistency* = *accordance*/*NoC*

(**a**)   (**b**)

Figure 2. Schematic of calculating the consistency value. (**a**) Pseudo-code for computing consistency; (**b**) Judging of relationship between data columns.

As such, two new matrices M_1 and M_2 are yielded, respectively. If M_1 is equal to M_2, the variable *'accordance'* is added by 1, which means that the metrics considered vary accordingly between two solutions (i.e., consistent with each other).

On the contrary, if M_1 is different from M_2, the variable "*accordance*" remains invariant because at least two metrics vary in different trends (as interlaced in Figure 2b). In short, the consistency value denotes the ratio of the number of non-crossed lines to the total number of twin-lines. Hence, a larger percentage of non-crossed lines implies a higher level of consistency among different metrics (i.e., better correlation).

3. Experimental Setup

3.1. Case Studies

Two benchmark WDSs from the literature were selected for the demonstration purposes, of which each network layout is illustrated in Figure 3. The first WDS (Figure 3a) is the well-known Hanoi network (HAN) [27], which resembles a water distribution system in Hanoi, the capital of Vietnam. The HAN consists of thirty-four pipes organized in three loops, thirty-one demand nodes and one reservoir with a fixed head of 100.0 m. The Hazen-Williams roughness coefficient for all pipes is 130. The minimum head above the ground elevation of each node is 30.0 m. There are six commercially available pipe sizes, ranging from 304.8 mm (12.0 in.) to 1016.0 mm (40.0 in.). Therefore, the search space is equal to $6^{34} \approx 2.87 \times 10^{26}$ discrete combinations. Due to a very limited range of pipe sizes, the HAN has a vast region of infeasible solutions in the landscape of decision variables, thus increasing the level of difficulty to identify the non-dominated solutions.

Unlike the HAN, the second WDS (Figure 3b) is known as the Balerma irrigation network (BIN) [28]. It represents an adaptation of the existing irrigation system in the Sol-Poniente irrigation district located in Balerma, province of Almería, Spain. The irrigation system is fed by four reservoirs with fixed heads between 112.0 m and 127.0 m and includes 454 relatively short pipes and 443 demand nodes (hydrants). A feature of BIN is that all nodes have the same demand of 5.55 l/s across the network. The material of pipes is polyvinyl chloride (PVC), and a universal Darcy-Weisbach roughness

coefficient of 0.0025 mm is applied to all pipes. The minimum pressure head above the ground elevation is 20.0 m for all demand nodes.
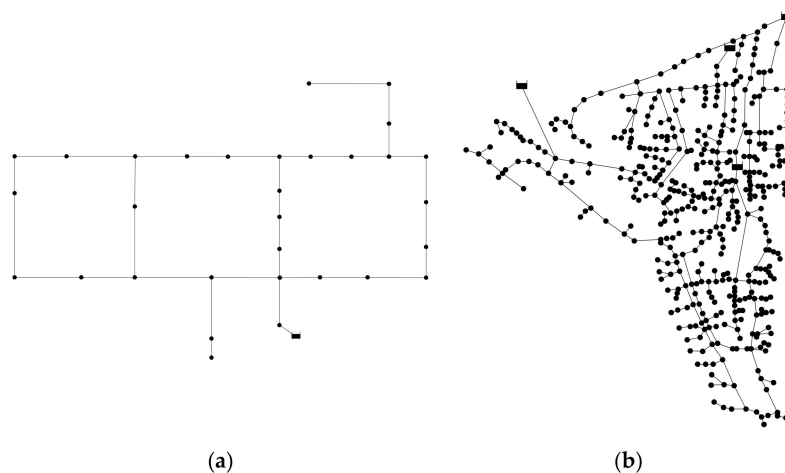

(a)   (b)

Figure 3. Layout of each benchmark network. (**a**) HAN; (**b**) BIN.

There are a total of ten commercially available sizes, ranging from 113.0 mm to 581.8 mm. Therefore, the search space is equal to 10^{454} discrete combinations, which serves as a more complex, real-world design problems.

3.2. Parameter Settings of Many-Objective Optimization

As previously described, Borg has several advanced features that ensure its performance without the burden associated with the parameterization issue. Therefore, no special attention was paid to fine-tuning many parameters involved in Borg and the default ones for most parameters (see Table 2) were used. For both cases, the epsilon precision settings for all the objectives were equal to 0.01, except for Cost and entropy-based indicators which were set to 0.1 instead (the epsilon value of DSFE for BIN was set to 1 due to a broader range). To ensure sufficient convergence of Borg on each design problem, preliminary optimization runs were conducted to determine the appropriate computational budget in terms of the number of function evaluations (NFEs). A total of 100,000 and 250,000 NFEs were used to solve the HAN and BIN problems, respectively. These were approximately equivalent to 1000 and 2500 generations as the initial population size was set to 100 in Borg. Also, for each benchmark design problem, Borg was implemented ten times (i.e., ten independent runs) to overcome the impact of randomness introduced by population initialization.

Table 2. Parameter settings of Borg MOEA.

Parameters	Values	Parameters	Values
Initial population size	100	Crossover rate of DE	0.6
Population to archive ratio (γ)	4	Differential weight of DE	0.6
Selection ratio (τ)	0.02	Eta of PCX	0.1
Tournament size	$Max\,[2, \lfloor \tau(\gamma A) \rfloor]$	Zeta of PCX	0.1
Probability of PM	$1/L$	Zeta of UNDX	0.5
Distribution index of PM	1.0	Eta of UNDX	0.35
Probability of SBX	0.9	Expansion rate of SPX	3.3
Distribution index of SBX	1.0	Probability of UM	$1/L$

Note: A = the archive size in Borg; L = the number of decision variables; $\lfloor a \rfloor$ returns the nearest integer of a in the direction of negative infinity; $Max[a,b]$ returns the larger value between two numbers a and b. PM, SBX, DE, PCX, UNDX, SPX, and UM are the search operators used in Borg, which stand for Polynomial Mutation, Simulated Binary Crossover, Differential Evolution, Parent-Centric Crossover, Unimodal Normal Distribution Crossover, Simplex Crossover, and Uniform Mutation, respectively.

A post-processing procedure was implemented to generate the best approximation to the Pareto-optimal set over multiple runs. A temporary archive of non-dominated solutions was established by aggregating all the final population returned by Borg via ten independent runs. The volume of this archive was checked and reduced by eliminating duplicated solutions. Then, the fast non-dominated sorting procedure [20] was applied to this archive to identify the overall best solutions and remove the dominated ones. The best approximation to the Pareto-optimal set of each design problem was subsequently used to reveal the relationships among reliability, cost, and various surrogate indicators.

4. Results and Discussion

The relationships among eight objectives and system reliability were interpreted at two different levels. At the top level, the PCP was used to illustrate the hyper Pareto front obtained from the many-objective optimization model. Each PCP was produced as follows: (1) the whole dataset was rearranged and drawn according to the ascending order of capital cost, and (2) the polylines were painted with rainbow colors with red representing higher costs. Therefore, the solutions with lower costs (drawn earlier) were partly covered by those with higher costs. The MRS values, which were obtained from the single-pipe burst testing, were demonstrated on the PCP as well (the leftmost axis). This is aimed at identifying the intrinsic consistency among reliability, cost, and surrogate indicators. At the bottom level, the main concern is to quantify the correlation between surrogate indicators and also their relationship with cost and reliability, based on the consistency indicator described in the Section of Methodology. Therefore, the selected indicators can be categorized into two groups: highly related and poorly related, which probably benefit the research community by providing useful guidelines.

The overall consistency among reliability, cost, and seven surrogate indicators is not satisfactory from an engineering perspective; and it is evident that a larger amount of investment does not necessarily lead to a more reliable design under mechanical failure scenarios (see the lower PCPs in Figure 4c,d with selected solutions). One will expect that the network system is more reliable under possible failure conditions with more investment (i.e., larger pipe diameters leading to more surplus power at demand nodes). For the HAN problem, the range of capital costs within the best approximation to the Pareto-optimal set is between $6.4M and $11.0M.

A higher investment of $10.3M leads to an MRS of 0.93, which is 5% less than the MRS of 0.98 achieved by an investment of $9.78M. In other words, a more reliable solution can be achieved with 5% less budget. Surprisingly, a nearly identical investment of $9.78M can also lead to a quite low level of MRS (i.e., 0.81). In contrast, the lowest investment of $6.49M can achieve an MRS of 0.79, saving more than 30% expenditure with a very similar level of MRS. A comparable pattern can also be found for the BIN design problem.

Furthermore, surrogate indicators are expected to correspond to system reliability consistently. However, this is not the case, as many interlaced polylines can be observed in Figure 4 for both cases. The underlying reason for these counterintuitive phenomena is due to the inclusion of improper surrogate indicators, such as FE and DSFE. As demonstrated in Figure 4, the polylines are generally parallel with each other except for many legible interlaced ones spotted at the coordinates of FE and DSFE, which indicates that FE and DSFE do not coincide with the other five indicators. A further experiment, which conducted the optimization procedure in the same way but excluded FE and DSFE from the objectives, confirms that the correlation of surrogate indicators with cost and mechanical reliability significantly improves, as shown in Figure 5. This finding suggests that maximization of FE and DSFE may not yield more reliable system configurations.
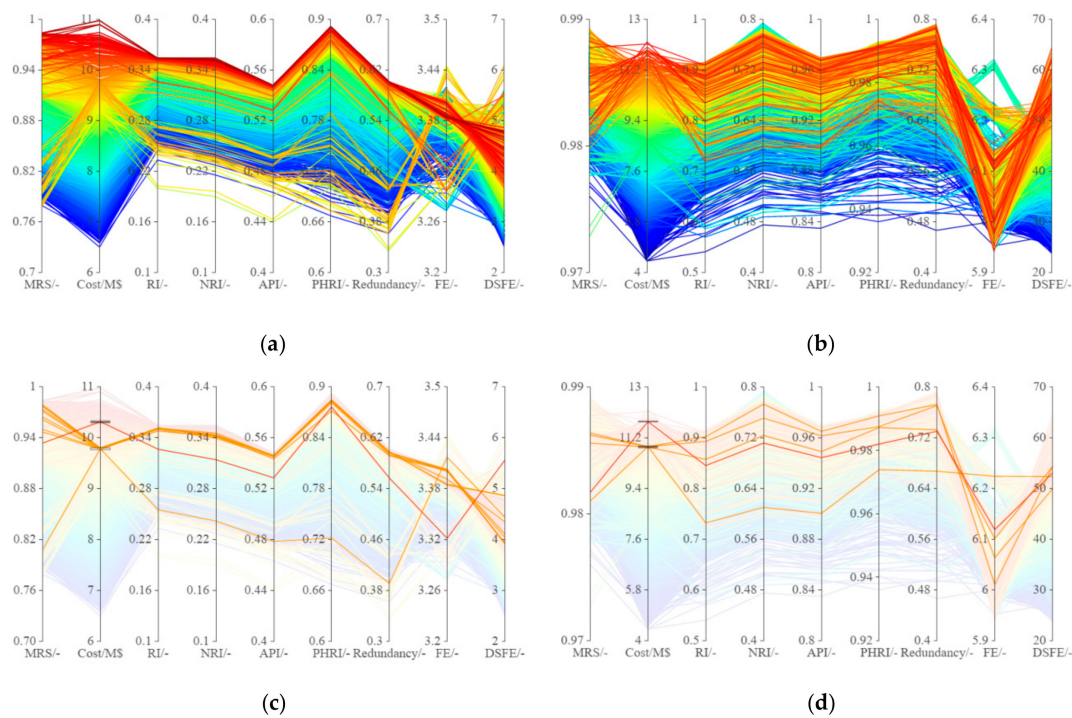
Figure 4. Parallel coordinate plots for the HAN and BIN design problems. (**a**) PCPs with all objectives and MRS for HAN; (**b**) PCPs with all objectives and MRS for HAN; (**c**) Representative solutions for HAN; (**d**) Representative solutions for BIN.
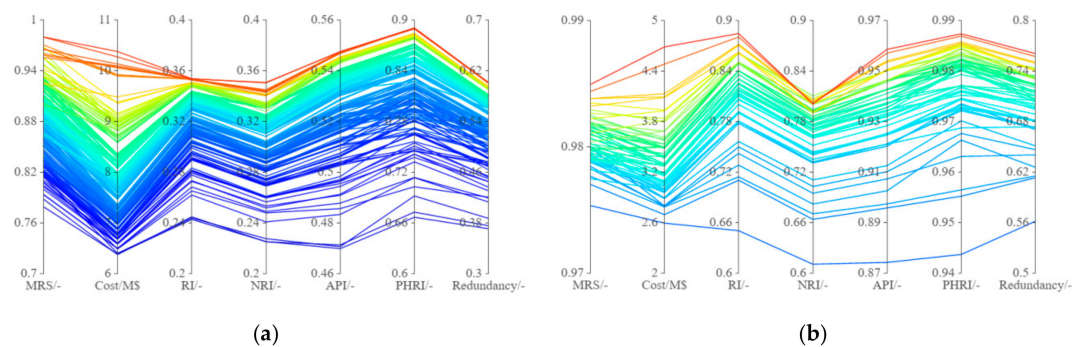

**Figure 5.** New PCPs for the HAN and BIN design problems. (**a**) HAN; (**b**) BIN.

Four resilience-based surrogate indicators (i.e., RI, NRI, API, and PHRI) and Redundancy are well consistent with mechanical reliability under single-pipe burst event. In contrast, two entropy-based indicators (i.e., FE and DSFE) are poorly related to reliability for both cases. This observation suggests that the resilience-based surrogate indicators and Redundancy can adequately reflect the mechanical reliability of WDSs for both cases; on the contrary, the entropy-based indicators are not suitable to represent mechanical reliability considered in this study. For the HAN problem, it is evident from Table 3 that RI, NRI, API, PHRI, and Redundancy are highly related to MRS, with the consistency values all above 0.9. While the consistency values between FE/DSFE and MRS are less than 0.5/0.7, respectively. For the BIN problem, the level of consistency between the non-entropy-based surrogate indicators and MRS is lower than that of the HAN problem but still higher than 0.75. This might be attributed to the differences between the layouts of two networks (i.e., single-source vs. multi-source). Similar to the pattern identified in HAN, the consistency values between FE/DSFE and MRS are both lower than 0.5.

Table 3. Consistency values between cost/reliability and surrogate indicators.

Problem	Consistency	RI	NRI	API	PHRI	Redundancy	FE	DSFE
HAN	MRS	0.930	0.936	0.930	0.907	0.932	0.481	0.684
	Cost	0.894	0.890	0.894	0.888	0.889	0.485	0.756
BIN	MRS	0.775	0.799	0.777	0.757	0.760	0.404	0.489
	Cost	0.722	0.645	0.721	0.721	0.726	0.505	0.840

Additionally, for both cases, the consistency values between non-entropy-based surrogate indicators and capital costs are lower than those between surrogate indicators and MRS. This implies that a higher level of surrogate indicators (generally more reliable) does not always require a larger amount of investment. This phenomenon is probably attributed to the combinatorial nature of WDS design problems.

Figure 6, which demonstrates the consistency value of each pair of axes in the PCPs, agrees well with the observations from Figure 4, and further confirms the relationships among surrogate indicators, costs, and mechanical reliability. That is, the four resilience-based surrogate indicators correspond well (with minor variations) with themselves and Redundancy (shown as the light gray cells in the middle). These five surrogate indicators are highly related to MRS, with the consistency values greater than 0.90 and 0.75 for the HAN and BIN problems, respectively. Their relationships with Cost are in good consistency, despite somewhat lower than those with MRS. FE and DSFE are poorly related to MRS, Cost, and the other surrogate indicators, with the consistency values less than 0.5 and 0.7 for the HAN and BIN problems, respectively. Surprisingly, FE and DSFE are inconsistent with each other even if they are quite similar in theory, especially for the HAN problem.

(a) (b)
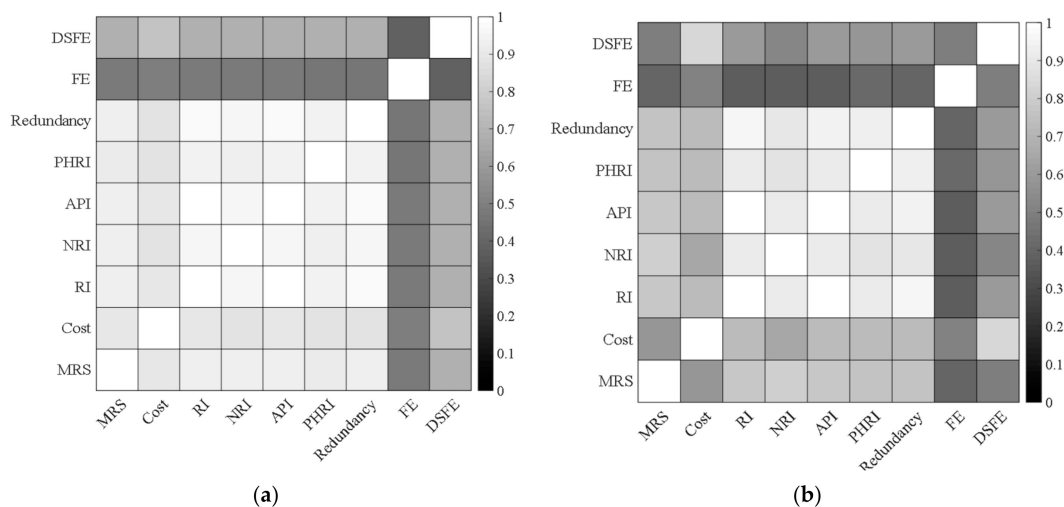
Figure 6. Consistency value of each pair of surrogate indicators. (**a**) HAN; (**b**) BIN.

To further explain why a more reliable design does not always require a larger amount of investment, the layouts of two representative solutions (denoted as Solutions H and L, respectively) from the Pareto set of the HAN problem were compared. Table 4 shows the corresponding values of these two solutions. Solution H costs $9.65M with a reliability value of 0.78; while Solution L has a reliability value of 0.89 (14.1% higher) with 16.7% savings of capital costs, which is equal to $8.03M. Figure 7 clearly shows the differences between Solutions H and L in terms of system layout. Specifically, Solution H lays the largest pipe size mainly in the north of the network as well as some carrying water to the most distant user (i.e., Node 12). However, improper connections (from the engineering perspective) within three loops can also be easily identified (larger pipes are connected to smaller pipes at upstream side).

Table 4. Comparison of two representative solutions for the HAN problem.

Solutions	MRS	Cost	RI	NRI	API	PHRI	Redundancy	FE	DSFE
	(-)	(M$)	(-)	(-)	(-)	(-)	(-)	(-)	(-)
Solution H	0.78	9.65	0.24	0.23	0.47	0.67	0.36	3.44	4.55
Solution L	0.89	8.03	0.33	0.31	0.53	0.86	0.56	3.32	3.22
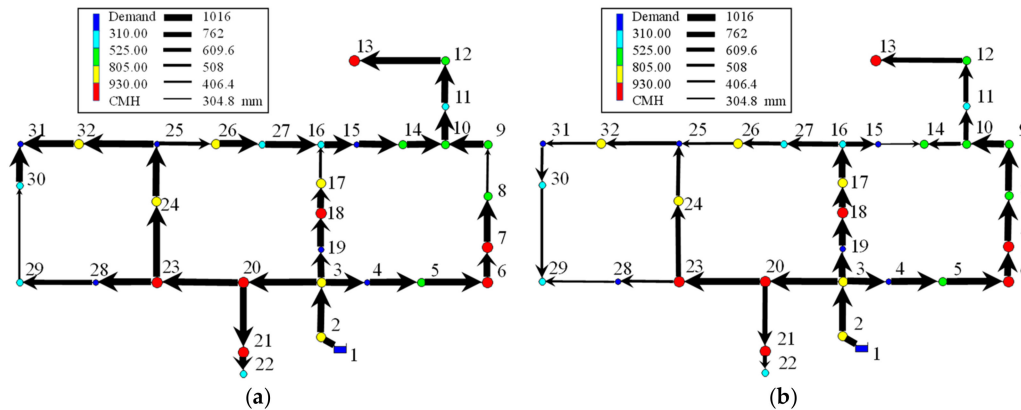
Figure 7. Comparison of two representative solutions found for the HAN problem. (**a**) Solution H with lower reliability; (**b**) Solution L with higher reliability.

In contrast, Solution L generally conveys water by the middle and right loops, with the low demand region (Nodes 28 to 31) connected by much smaller pipes. Also, Node 13 with a high level of demand and some nodes with a middle level of demands (i.e., Nodes 11 and 12) are supplied by smaller branched pipes. As a result, nodes with higher demands (shown as red dots) are more effectively connected to the source with reasonable combinations of pipe sizes from north to south.

Solution L can effectively alleviate the impact of single-pipe burst event as revealed in Figure 8. The y-axis shows the probability of pressure deficit calculated from a collection of 32 single-pipe burst events (except for the two pipes directly connected to the source). The dashed line shows a threshold of 50% of suffering pressure deficiency during those events. Solution H results in nearly one-third of demand nodes (nine in total) bearing the risk of insufficient water supply (or even water cut-off) over 50%. The most vulnerable locations include Nodes 9 to 16, which are concentrated in the north of the Hanoi network. By contrast, Solution L manages to reduce such risks below 35% throughout the system, which indicates a very high level of performance under mechanical failure conditions.
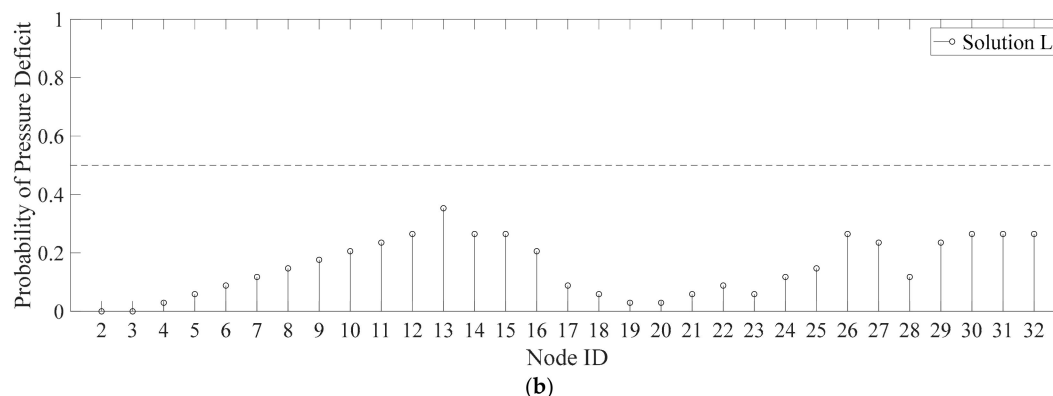

Figure 8. *Cont.*

(b)
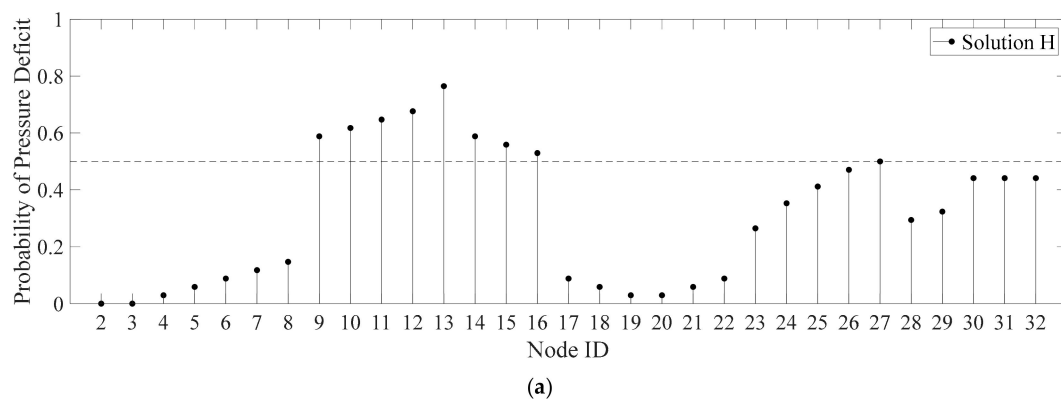
Figure 8. Probabilities of nodal pressure deficit associated with Solutions H and L. (**a**) Solution H; (**b**) Solution L.

5. Conclusions

This paper proposed an innovative way of identifying the relationships among various surrogate indicators for assessing the mechanical reliability of WDSs. The many-objective formulation, which was aimed at minimizing capital cost and maximizing each surrogate indicator, was used to investigate the trade-offs between cost and surrogate indicators in a single run. To further reveal the consistency among reliability, cost, and surrogate indicators, a systematical, post-analysis procedure was implemented based on the single-pipe burst testing. The above methodology was applied to two benchmark design problems from the literature, covering both single-source and multi-source water supply systems fed by gravity.

The main findings of this work are threefold. First, the trade-off between capital costs and surrogate indicators suggest that a larger amount of investment at the design stage does not necessarily lead to a more resilient or reliable design solution for both case studies, which sounds somewhat counterintuitive. This is exactly the case when the two entropy-based indicators are involved in the optimization model. The intrinsic inconsistency between flow entropy and system resilience or reliability hinders the identification of the expected trade-offs between capital costs and surrogate indicators. Another drawback of entropy-based indicators lies in the fact that their values vary substantially from case to case, and are not bound to a fixed range as other surrogate indicators (i.e., between 0 and 1). Therefore, users may have difficulties in interpreting the outcomes from an intuitive way.

Second, the four resilience-based surrogate indicators and Redundancy are more consistent with the mechanical reliability and capital costs compared to the two entropy-based indicators, implying that the former are more suitable for measuring the performance of WDSs. The consistency values between the mechanical reliability and resilience-based indicators and Redundancy are higher than 0.9 and 0.75 for the HAN and the BIN design problems, respectively. In contrast, the consistency values between the mechanical reliability and entropy-based indicators are less than 0.7 and 0.5 for the above two cases, respectively. These observations correspond well to other researchers' investigations [8,15,29].

Third, due to the combinatorial nature of the design of WDSs, a proper formulation can effectively steer the search towards the near-optimal region. As shown in the current study, using resilience-based surrogate indicators to substitute system reliability analysis is recommended in the optimization model. In particular, Redundancy is highly recommended due to its simplicity in both concept and computation.

Although a practical approach to simultaneously comparing many surrogate indicators was developed, there are still some limitations in the present study. First, the selected design problems do not include pumps and are limited to snapshot simulations only. Therefore, it is still unknown whether the findings derived from this paper would change significantly when pumps and extended period simulations are taken into account. Second, the correlation between surrogate indicators and

the hydraulic reliability, which is related to the impact imposed by demand fluctuations, is not covered in this study. Third, how would the consistency among reliability (both hydraulic and mechanical), costs, and surrogate indicators vary on more complex, real-world networks is still in question. All these limitations mentioned above will be considered in future work.

References

1. Klise, K.A.; Bynum, M.; Moriarty, D.; Murray, R. A software framework for assessing the resilience of drinking water systems to disasters with an example earthquake case study. *Environ. Model. Softw.* **2017**, *95*, 420–431. [CrossRef] [PubMed]

2. Goulter, I. Analytical and Simulation Models for Reliability Analysis in Water Distribution Systems. In *Improving Efficiency and Reliability in Water Distribution Systems. Water Science and Technology Library*; Cabrera, E., Vela, A.F., Eds.; Springer: Dordrecht, The Netherlands, 1995; Volume 14, pp. 235–266.

3. Atkinson, S.; Farmani, R.; Memon, F.A.; Butler, D. Reliability Indicators for Water Distribution System Design: Comparison. *J. Water Resour. Plan. Manag. ASCE* **2014**, *140*, 160–168. [CrossRef]

4. Tanyimboh, T.T.; Templeman, A.B. A quantified assessment of the relationship between the reliability and entropy of water distribution systems. *Eng. Optimiz.* **2000**, *33*, 179–199. [CrossRef]

5. Todini, E. Looped water distribution networks design using a resilience index based heuristic approach. *Urban Water* **2000**, *2*, 115–122. [CrossRef]

6. Prasad, T.D.; Nam-Sik Park, M.A. Multiobjective Genetic Algorithms for Design of Water Distribution Network. *J. Water Resour. Plan. Manage. ASCE* **2004**, *130*, 73–82. [CrossRef]

7. Jayaram, N.; Srinivasan, K. Performance-based optimal design and rehabilitation of water distribution networks using life cycle costing. *Water Resour. Res.* **2008**, *44*, W01417. [CrossRef]

8. Raad, D.N.; Sinske, A.N.; van Vuuren, J.H. Comparison of four reliability surrogate measures for water distribution systems design. *Water Resour. Res.* **2010**, *46*, W05524. [CrossRef]

9. Wu, W.; Maier, H.R.; Simpson, A.R. Surplus Power Factor as a Resilience Measure for Assessing Hydraulic Reliability in Water Transmission System Optimization. *J. Water Resour. Plan. Manag. ASCE* **2011**, *137*, 542–546. [CrossRef]

10. Tanyimboh, T.T.; Tietavainen, M.T.; Saleh, S. Reliability assessment of water distribution systems with statistical entropy and other surrogate measures. *Water Sci. Technol. Water Supply* **2011**, *11*, 437–443. [CrossRef]

11. Jung, D.; Kang, D.; Kim, J.H.; Lansey, K. Robustness-Based Design of Water Distribution Systems. *J. Water Resour. Plan. Manag. ASCE* **2013**, *140*, 04014033. [CrossRef]

12. Liu, H.; Savic, D.; Kapelan, Z.; Zhao, M.; Yuan, Y.; Zhao, H. A diameter-sensitive flow entropy method for reliability consideration in water distribution system design. *Water Resour. Res.* **2014**, *50*, 5597–5610. [CrossRef]

13. Liu, S.; Wang, H.; Xin, K. Multi-objective Optimal Design of Large-scale Urban Water Distribution Network. *China Water Wastewater* **2014**, *30*, 52–55, 60. (In Chinese)

14. Creaco, E.; Franchini, M.; Todini, E. The combined use of resilience and loop diameter uniformity as a good indirect measure of network reliability. *Urban Water J.* **2016**, *13*, 167–181. [CrossRef]

15. Liu, H.; Savic, D.; Kapelan, Z.; Creaco, E.; Yuan, Y. Reliability Surrogate Measures for Water Distribution System Design: Comparative Analysis. *J. Water Resour. Plan. Manag. ASCE* **2017**, *143*, 04016072. [CrossRef]

16. Bin Mahmoud, A.A.; Piratla, K.R. Comparative evaluation of resilience metrics for water distribution systems using a pressure driven demand-based reliability approach. *J. Water Supply Res Technol. Aqua.* **2018**, *67*, 517–530. [CrossRef]

17. Cimorelli, L.; Morlando, F.; Cozzolino, L.; D'Aniello, A.; Pianese, D. Comparison Among Resilience and Entropy Index in the Optimal Rehabilitation of Water Distribution Networks Under Limited-Budgets. *Water Resour. Manag.* **2018**, *32*, 3997–4011. [CrossRef]

18. Tanyimboh, T.T.; Siew, C.; Saleh, S.; Czajkowska, A. Comparison of Surrogate Measures for the Reliability and Redundancy of Water Distribution Systems. *Water Resour. Manag.* **2016**, *30*, 3535–3552. [CrossRef]

19. Hadka, D.; Reed, P. Borg: An Auto-Adaptive Many-Objective Evolutionary Computing Framework. *Evol. Comput.* **2013**, *21*, 231–259. [CrossRef]

20. Deb, K.; Pratap, A.; Agarwal, S.; Meyarivan, T. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Trans. Evol. Comput.* **2002**, *6*, 182–197. [CrossRef]

21. Shannon, C.E. A Mathematical Theory of Communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423. [CrossRef]

22. Rossman, L.A. *EPANET Users Manual*; U.S. Environmental Protection Agency: Washington, DC, USA, 2000.

23. Laumanns, M.; Thiele, L.; Deb, K.; Zitzler, E. Combining Convergence and Diversity in Evolutionary Multiobjective Optimization. *Evol. Comput.* **2002**, *10*, 263–282. [CrossRef]

24. Purshouse, R.C.; Fleming, P.J. On the Evolutionary Optimization of Many Conflicting Objectives. *IEEE Trans. Evol. Comput.* **2007**, *11*, 770–784. [CrossRef]

25. Reed, P.M.; Hadka, D.; Herman, J.D.; Kasprzyk, J.R.; Kollat, J.B. Evolutionary multiobjective optimization in water resources: The past, present, and future. *Adv. Water Resour.* **2013**, *51*, 438–456. [CrossRef]

26. Smith, R.; Kasprzyk, J.; Zagona, E. Many-Objective Analysis to Optimize Pumping and Releases in Multireservoir Water Supply Network. *J. Water Resour. Plan. Manag. ASCE* **2016**, *142*, 04015049. [CrossRef]

27. Fujiwara, O.; Khang, D. A two-phase decomposition method for optimal design of looped water distribution networks. *Water Resour. Res.* **1990**, *26*, 539–549. [CrossRef]

28. Reca, J.; Martínez, J. Genetic algorithms for the design of looped irrigation water distribution networks. *Water Resour. Res.* **2006**, *42*, W05416. [CrossRef]

29. Monsef, H.; Naghashzadegan, M.; Farmani, R.; Jamali, A. Deficiency of Reliability Indicators in Water Distribution Networks. *J. Water Resour. Plan. Manag. ASCE* **2019**, *145*, 04019026. [CrossRef]