



## Article

# Optimal Number of Pressure Sensors for Real-Time Monitoring of Distribution Networks by Using the Hypervolume Indicator

Bruno Ferreira <sup>1,\*</sup> , Nelson Carriço <sup>1</sup>  and Dídia Covas <sup>2</sup> 

<sup>1</sup> INCITE, Barreiro School of Technology, Polytechnic Institute of Setúbal, 2910-761 Setúbal, Portugal; nelson.carriço@estbarreiro.ips.pt

<sup>2</sup> CERIS, Instituto Superior Técnico, Universidade de Lisboa, 1049-001 Lisbon, Portugal; didia.covas@tecnico.ulisboa.pt

\* Correspondence: bruno.s.ferreira@estbarreiro.ips.pt

**Abstract:** This article proposes a novel methodology to determine the optimal number of pressure sensors for the real-time monitoring of water distribution networks based on a quality hypervolume indicator. The proposed methodology solves the optimization problem for different numbers of pressure sensors, assesses the gain of installing each set of sensors by means of the hypervolume indicator and determines the optimal number of sensors by the variation of the hypervolume indicator. The methodology was applied to a real case study. Several robustness analyses were carried out. The results demonstrate that the methodology is hardly influenced by the method parameters and that a reasonable estimation of the optimal number of sensors can be easily achieved.

**Keywords:** real-time monitoring; pressure sensors; optimization; multi-objective; hypervolume



**Citation:** Ferreira, B.; Carriço, N.; Covas, D. Optimal Number of Pressure Sensors for Real-Time Monitoring of Distribution Networks by Using the Hypervolume Indicator. *Water* **2021**, *13*, 2235. <https://doi.org/10.3390/w13162235>

Academic Editor: Athanasios Loukas

Received: 29 June 2021

Accepted: 15 August 2021

Published: 17 August 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Water distribution networks (WDN) constitute vital infrastructure to ensure drinking water in sufficient quantity and quality for users. Real-time monitoring of these networks is essential to ensure the quality of the service provided. This is achieved by employing on-site sensors to measure certain water quality (e.g., pH, temperature, contaminants) and hydraulic (i.e., pressure and flowrate) parameters [1]. These sensors produce different types of data (e.g., flowrate, pressure, contaminants concentration) that can be used in different activities. Only a limited number of sensors can be installed in a given WDN due to budget constraints. Moreover, these WDN may rapidly increase in complexity (as a function of the area served), resulting in a challenging task of real-time monitoring of large networks when few sensors are installed [2].

Numerous methods have been developed over the recent decades to find the optimal location for sensors to detect contamination events in real time and mitigate the associated effects. In the Battle of the Water Sensor Networks [3], fifteen different approaches of optimal sensor location for contaminant detection were compared according to four different objectives. Aral et al. [4] blended four distinct criteria into one single objective which was solved using genetic algorithms. According to Weickgenannt et al. [5], the risk of contamination is explicitly evaluated as the product of likelihood of not detecting the contaminant intrusion and the corresponding consequence (water consumed). In a different approach, Zhao et al. [6] attempted to minimize the consumption of contaminated water prior to contamination detection. More recently, Naserizade et al. [7] used multi-objective optimization and a multicriteria decision-making technique whilst Ponti et al. [8] proposed a new multi-objective evolutionary algorithm which showed improvements over NSGA-II. However, sensor location techniques aiming at monitoring water quality parameters may not provide the best solutions for hydraulic parameters since they have different goals in the utility activities.

In particular, pressure data are widely used by water utilities in the real-time monitoring of WDN for creating alerts and use in hydraulic modelling calibration and pipe burst detection. Alarm systems can be designed to respond to the occurrence of abnormal pressure values (low or high pressure) at specific locations of the network [9]. The calibration process of hydraulic models requires pressure data from specific locations of the network to calibrate high uncertainty variables such as nodal demands and pipe roughness coefficient [2,10]. Pipe burst detection and location techniques use pressure data in a variety of ways, for instance, to carry out inverse analysis of the pipe burst location (using hydraulic simulation and optimization) [11,12] or by using a data-driven classifier approach [13].

Distinct methods have been proposed in the literature to optimally locate pressure sensors in WDN, most of them focusing on the total gain of sensor placement considering different aims. Cao et al. [14] positioned pressure sensors in such a way as to represent the pressure patterns of the homogeneous area of the WDN. Zhao et al. [15] aimed at maximizing the detection coverage rate of pipe burst events whilst Raei et al. [16] aimed at reducing the detection time of these events. Both Casillas et al. [17] and Steffelbauer et al. [18] focused on maximizing the percentage of leak scenarios correctly identified (according to the introduced criteria). Sarrate et al. [19] focused on maximizing the leak detectability performance. Furthermore, the maximization of nodal pressure sensitivities was considered both by de Schaetzen et al. [20] and Francés-Chust et al. [21]. The maximization of the accuracy of the hydraulic model was also considered by Kapelan et al. [22] and Behzadian et al. [23].

Ideally, the optimal location and number of pressure sensors should be determined simultaneously. Nonetheless, most of the previously presented methods are not able to directly provide the optimal number of pressure sensors to be installed. As a result, the number of pressure sensors to be installed in a WDN is usually defined using different criteria. Sarrate et al. [19] and Sanz et al. [24] defined the number of pressure sensors for pipe burst detection by budget limitations whilst both Sophocleous et al. [11] and Wu et al. [25] used a metric of the number of households per pipe unit length. For the same objective of pipe burst detection, Soldevila et al. [26] used a metric based on the network size, and Quintiliani et al. [27] used engineering good sense. Regarding pressure sensors for model calibration, de Schaetzen et al. [20] used the number of households per unit length to decide the number of sensors whilst Wéber et al. [2] emphasized the minimum percentage of calibration errors to be achieved.

This article presents a novel methodology to determine the optimal number of pressure sensors in a WDN. The major advantage is the possibility to be coupled with the existing optimization methods of sensor location (such as those previously presented), thus providing the optimal number according to the considered objective functions. Furthermore, the methodology does not require previous knowledge on sensor installation costs nor the consideration of costs in the optimization problem itself. The methodology uses the hypervolume indicator [28] to measure the total gain of installing different numbers of pressure sensors (according to the objective functions being considered) with a single value. In order to reduce the number of solved optimization problems, a discrete set of numbers of sensors was established in a distributed and representative way for which the optimization problem would be solved. A trade-off function was then fitted to the hypervolume data, which can be interpreted as a compromise between the total gain (using the hypervolume) and the number of sensors. Finally, the optimal number of pressure sensors was determined by assessing the evolution of the hypervolume indicator as a function of the number of sensors, specifically by identifying the point of maximum curvature of the hypervolume data. From that point onwards, the inclusion of another sensor was rather not worth the increment on the total gain.

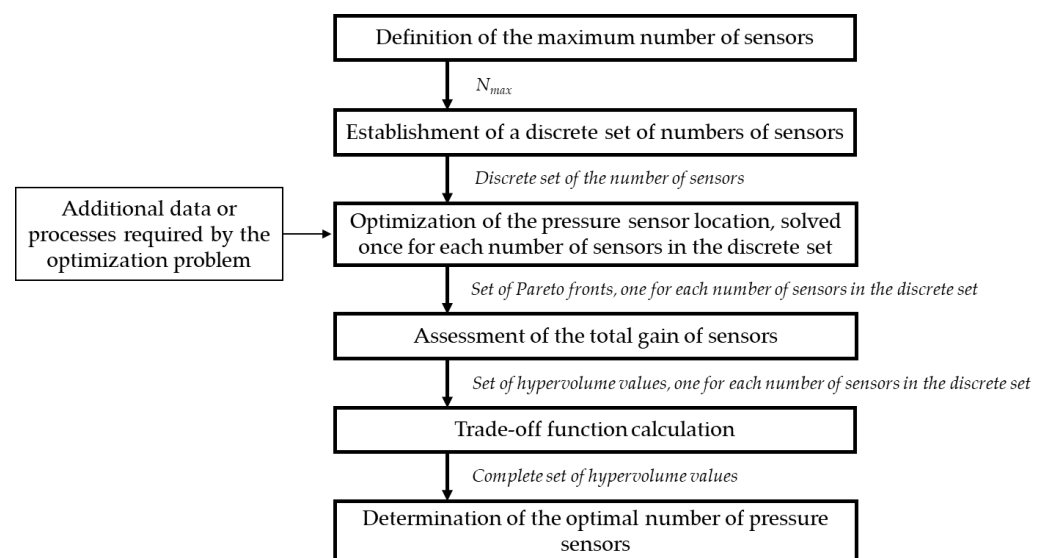
This methodology is demonstrated through application to a real WDN. The pressure sensors were located using an optimization method based on the maximization of nodal sensitivities to both pipe burst events and variations of the pipe roughness coefficient. The obtained results are discussed and the most relevant conclusions are drawn.

The main novel contributions are as follows: (i) the proposal and application of the hypervolume indicator to quantify the total gain associated with a given number of pressure sensors and (ii) the application of different techniques to detect the point of maximum curvature aiming to determine the optimal number of pressure sensors.

## 2. Methodology

### 2.1. General Approach

A flowchart of the proposed methodology is present in Figure 1 and is composed of six main steps: (1) definition of the maximum number of sensors; (2) establishment of a discrete set of numbers of sensors; (3) optimization of the pressure sensor locations; (4) assessment of the total gain of sensors; (5) trade-off function calculation; (6) determination of the optimal number of pressure sensors. The following sections further detail each step of the proposed methodology.



**Figure 1.** Flowchart of the proposed methodology.

### 2.2. Definition of the Maximum Number of Sensors

The optimal number of pressure sensors is determined by analyzing the variation of the total gain (using the results of optimization problems) with the number of sensors. However, solving optimization problems for an incremental number of sensors (e.g., between 1 and 150) can be computationally expensive due to the exponential increase in the search space. For example, considering 20, 21 and 22 sensors in a total set of 150 possible locations results in  $3.63 \times 10^{24}$ ,  $2.24 \times 10^{25}$  and  $1.31 \times 10^{26}$  possible combinations of sensors, respectively. Therefore, the maximum number of installed sensors,  $N_{max}$ , is firstly defined, aiming at minimizing the computational burden by limiting the number of optimization problems to be solved.

Different approaches can be used for the definition of the maximum number of pressure sensors in a WDN, for instance, by using economic criteria based on the size of the WDN [18] or by determining when the gain of installation of an additional sensor has no significant improvement [15], taking into account the total number of possible locations [16] or the available budget [23,27].

### 2.3. Establishment of the Discrete Set of Numbers of Sensors

A discrete set of numbers of sensors is established between 1 and the maximum number of sensors  $N_{max}$ , for which the optimization problem will be solved. The established set does not have to be a continuous sequence of natural numbers, for instance, it could be 1, 5, 10, 15,  $\dots$ ,  $N_{max}$ . The aim is to reduce the number of optimization problems to

be solved whilst leaving enough results to characterize the total gain as a function of the number of sensors between 1 and  $N_{max}$ . As such, different sets of the number of sensors can be considered, for instance, with an evenly distributed number of sensors between 1 and  $N_{max}$  or with a higher density of observations in the lower number of sensors. The effect of different discrete sets of the numbers of sensors is assessed in this article.

#### 2.4. Optimization of the Pressure Sensor Locations

The total gain with the installation of pressure sensors depends on the method used for locating those sensors. Different pressure sensor location methods can be coupled to the proposed methodology, based on different optimization problem formulations [2,15–23,29].

A method to optimally locate pressure sensors based on the maximization of nodal sensitivities is considered to demonstrate the performance of the proposed methodology. This method is based on the formulation of an unconstrained multi-objective optimization problem, in which the decision variables are the pressure sensor locations (using nodes as possible locations). It requires the prior computation of two pressure sensitivity matrices  $S1$  and  $S2$  (one for the pipe roughness coefficient and the other for the pipe burst size, respectively). The size of the obtained matrices is the number of pipes  $\times$  the number of nodes and the number of nodes  $\times$  the number of nodes, respectively. As such,  $S1_{i,j}$  refers to the variation of the pressure in node  $j$  with respect to the variation of the pipe roughness coefficient of pipe  $i$ . Similarly,  $S2_{i,j}$  refers to the variation of the pressure in node  $j$  with respect to the variation in the pipe burst size in node  $i$ . Further details regarding pressure sensitivity analysis can be found in the works by de Schaetzen et al. [20] and Lansey et al. [30].

Accordingly, two objective functions are formulated for a given sampling design  $X$  of sensor locations. The first function  $f_1$  is defined according to de Schaetzen et al. [20] by using a compromise programming formulation. It aims at maximizing the sensitivity to the pipe roughness coefficient that the sensors cover ( $f_A$ ) whilst ensuring the evenly spread geographical distribution of sensor locations by maximizing the entropy according to Shannon's definition ( $f_B$ ). Functions  $f_A$  and  $f_B$  are defined as follows:

$$\text{Maximize } f_A = \sum_{i=1}^{Np} a_i, \quad \text{where } a_i = \max_{j \in X} (S1_{i,j}) \quad (1)$$

$$\text{Maximize } f_B = - \sum_{i=1}^{Np} p_i \times \log_2(p_i), \quad \text{where } p_i = \frac{a_i}{\sum_{i=1}^{Np} a_i} \quad (2)$$

where  $Np$  is the number of pipes.

Finally, the two functions  $f_A$  and  $f_B$  are combined together into  $f_1$  by taking the weighted sum of the two functions.

The second function  $f_2$  aims at directly maximizing the sensitivity of sensors to pipe burst events and can be computed as follows:

$$\text{Maximize } f_2 = \sum_{i=1}^{Nn} a_i, \quad \text{where } a_i = \max_{j \in X} (S2_{i,j}) \quad (3)$$

where  $Nn$  is the number of nodes.

Further details regarding the formulation of the objective functions can be found in the work by de Schaetzen et al. [20].

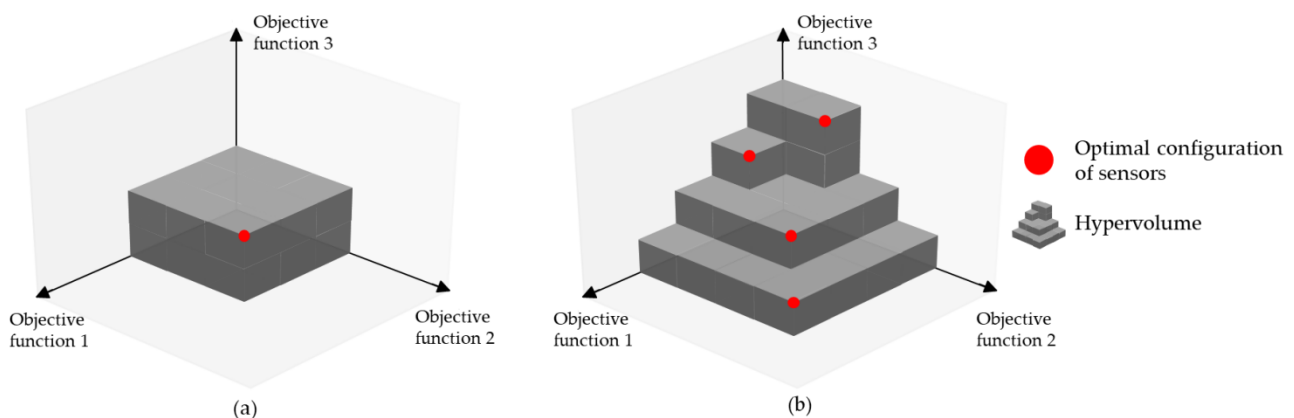
The optimization problem should be solved for each number of sensors in the discrete set. The results of each optimization problem are a single set of pressure sensor locations (for single objective optimization) or a set of optimal combinations of sensors (for multiple objective functions). In the latter case, a Pareto front of solutions is obtained; each solution of a Pareto front refers to a specific combination of sensors for which no other combination exists that presents better results for both objectives. Each optimal combination is characterized by the values associated with each objective function. These values are used to

assess the gain of installing the different number of sensors in the WDN. Note that the same objective functions and optimization method should be used for all numbers of sensors in the discrete set.

### 2.5. Assessment of the Total Gain of Sensors

The result of the previous step is a Pareto front for each number of sensors in the discrete set. Each Pareto front embraces all the possible combinations of sensors for that specific number of sensors.

The assessment of the total gain for a given number of sensors is not a trivial task as the gain can be described by multiple objective functions and, therefore, can have multiple different values; consider the example presented in Figure 2a, where a single optimal configuration of five sensors is depicted with a red dot. The gain of that specific configuration of sensors is characterized by three different values in accordance with some hypothetical objective functions 1, 2 and 3 to maximize (these are deliberately different from  $f_1$  and  $f_2$  since the methodology can be adopted to different optimization problem formulations with different dimensions). Furthermore, multiple optimal configurations might exist for a given number of sensors due to a trade-off between the objectives. This is represented in Figure 2b with a Pareto front of four optimal combinations of five sensors depicted with red dots. Note that the increase of objective function 3 will lead to the decrease of objective functions 1 and 2 as a trade-off between these contradictory objectives.



**Figure 2.** Characterization of the total gain for a given number of sensors using the hypervolume indicator.

In order to assess the total gain of installing a different number of pressure sensors, the quality of each obtained Pareto front is assessed using a quality measure [28,31–36]. This calculation leads to a single value for each Pareto front, characterizing the total gain associated with each specific number of sensors.

The use of the hypervolume indicator [28] is proposed herein to measure the quality of each Pareto front. It is one of the most widely used quality measures, the main advantages whereof being its easy interpretation and properties, such as guaranteeing strict monotonicity regarding Pareto dominance [37]. Nonetheless, this indicator also has disadvantages, such as sensitivity to the presence or absence of extreme points in a Pareto front [38]. In sum, this indicator considers the “volume” (i.e., in as many dimensions as the number of the objective functions) of the region of the objective space dominated by the set of optimal solutions in relation to the global worst point. Note that the hypervolume has no defined units, and its magnitude is related to the objective functions used. The hypervolume is graphically depicted in Figure 2 as a grey volume and considering the global worst point as (0,0,0). The formal mathematical definition of the hypervolume can be found in the work by Guerreiro et al. [37].

Different methods of the hypervolume calculation have been developed [37–42], with different efficiency levels against high dimensionality problems [37]. The efficiency of the hypervolume calculation in this particular methodology is not a major concern because

no high dimensionality is expected and a reduced number of hypervolume calculations is required (specifically, one calculation for each Pareto front).

### 2.6. Trade-Off Function Fitting

The result of the previous step is a set of hypervolume values, with one value associated with each number of sensors in the discrete set previously defined in Section 2.3. As the number of installed pressure sensors increases, the hypervolume also increases since more sensors lead to a higher total gain. This can be interpreted as a trade-off between the total gain of installing pressure sensors (using the hypervolume) and the number of sensors.

Therefore, observations of the hypervolume are used to derive a trade-off function that describes the behavior of the hypervolume (interpreted as the total gain) as a function of the number of sensors. By doing so, it is possible to estimate the total gain associated with the number of sensors for which the optimization problem is not solved.

Different algorithms can be used to fit a function to the hypervolume values by solving the nonlinear least-squares problem. A review on numerical methods for nonlinear least-squares optimization problems can be found in the work by Yuan [43]. The use of the Levenberg–Marquardt method is proposed to optimally fit a function by finding its optimal set of parameters. It is a well-established method with fast convergence, although it may be sensitive to the initial parameters guess.

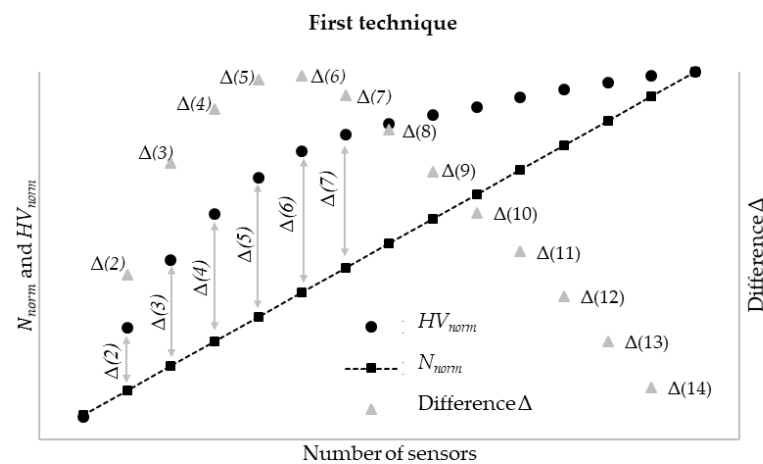
Distinct trade-off functions (with different mathematical formulations and numbers of parameters) can be fitted to the same hypervolume data. In this study, the well-known root-mean-square error (RMSE) is used to assist in deciding which function (after the fitting) best describes the hypervolume data. The smaller the value of the RMSE is, the better the fitted function represents the behavior of the hypervolume.

The selected fitted function is used to derive estimations of the hypervolume for the complete set of numbers of sensors between 1 and  $N_{max}$ .

### 2.7. Determination of the Optimal Number of Pressure Sensors

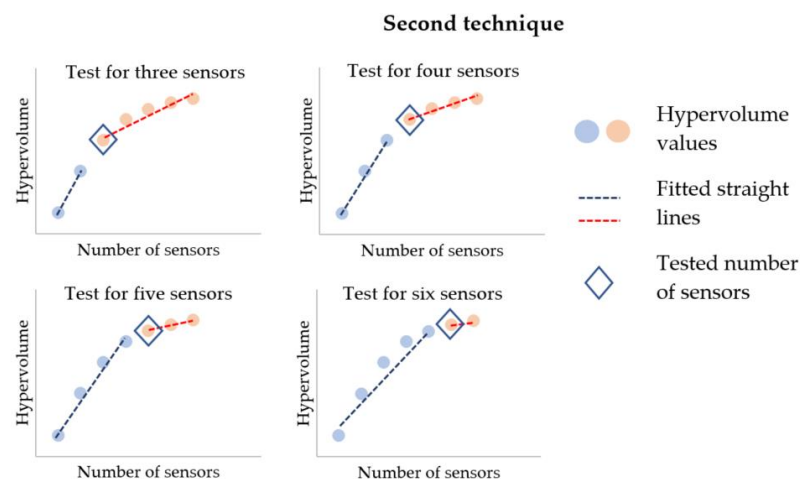
The hypervolume estimations are analyzed in order to determine the optimal number of sensors. This optimal number corresponds to the point where the inclusion of another sensor does not significantly improve the increment on the total gain. Such a point is defined as the point of maximum curvature in the hypervolume curve. The mathematical formulation for the point of maximum curvature can be found in the work by Satopaa et al. [44]. This point is defined herein as the “knee/elbow” of the trade-off function and can be identified using different automatic techniques [44–47]. The two techniques used herein are based on the works of Satopaa et al. [44] and Salvador and Chan [46], respectively. The performance of both techniques is assessed later in the article.

The first technique, the Kneedle method, starts by rescaling the estimated hypervolume data to  $[0,1]$ , leading to the normalized hypervolume values,  $HV_{norm}$ . The number of sensors between 1 and  $N_{max}$  should also be rescaled to  $[0,1]$ , leading to the normalized number of sensors values,  $N_{norm}$ . These values are represented in Figure 3 as black circles and squares, respectively, considering  $N_{max} = 15$  in this example. Finally, the differences  $\Delta$  between  $HV_{norm}$  and  $N_{norm}$  are calculated, as exemplified with grey arrows and triangular markers in Figure 3. The optimal number of sensors,  $N_{opt}$ , is selected as the number of sensors that presents the maximum difference.



**Figure 3.** Determination of the optimal number of sensors based on the evolution of the hypervolume by using the Kneedle method.

The second technique, the L-method, determines the “knee/elbow” as the point that best divides the hypervolume data into two straight lines. Pairs of lines are fitted to the hypervolume data using a linear regression procedure, and their junction point is the looked-for “knee/elbow”. Each line must contain at least two points and start at either end of the hypervolume data. Figure 4 shows all the possible pairs of straight lines for an example of the hypervolume data considering  $N_{max} = 7$ . The seven hypervolume values are presented both as blue or orange circles (either they are on the left or right side of the junction point). The corresponding fitted straight lines are represented with blue or red dashed lines, respectively. The number of sensors tested for the optimal number is in a blue diamond. The optimal number of sensors is the junction point that minimizes the weighted RMSE for the two linear parts of the hypervolume data.



**Figure 4.** Determination of the optimal number of sensors based on the evolution of the hypervolume using the second technique.

The final optimal locations can be found in the Pareto front related to the optimal number of sensors  $N_{opt}$ . Nonetheless, the Pareto front might not have yet been obtained (i.e., when  $N_{opt}$  was not considered in the discrete set). In these cases, the optimization problem should be solved for the number of sensors equal to  $N_{opt}$ .

### 3. Case Study

This methodology is demonstrated in a real WDN with a total network extension of ca. 70 km, 2200 service connections and an average inlet flowrate of about 60 m<sup>3</sup>/h. The

supplied area consists mainly of single-family houses with irrigated gardens, most of them with a swimming pool. The WDN also supplies a few golf courses. The hydraulic simulation model was developed in EPANET [48] and included four storage tanks, 4474 pipes and 4429 nodes, of which 2200 were consumers with an individual hourly demand pattern. Figure 5 presents the layout of the hydraulic network.

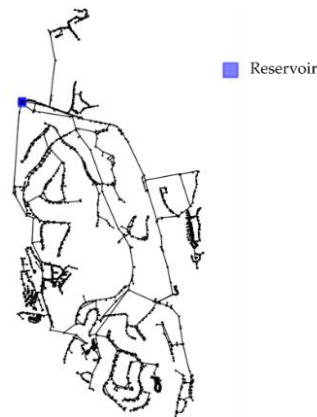


Figure 5. Layout of the case study.

The maximum number of installed sensors,  $N_{max}$ , was determined based on the intrinsic network characteristics, namely, on network length. A maximum of one sensor per kilometer of pipe network was adopted according to several Portuguese expert opinions, leading to  $N_{max} = 70$  sensors (i.e., one sensor per km).

Five discrete sets of numbers of sensors were defined. The objective was to assess if major differences in the optimal number of sensors were found when considering distinct sets of numbers of sensors. These five sets presented distinct characteristics and were defined as follows:

$$Set 1 = [1, 10, 20, 30, 40, 50, 60, 70] \quad (4)$$

$$Set 2 = [1, 10, 30, 50, 70] \quad (5)$$

$$Set 3 = [1, 10, 20, 30, 70] \quad (6)$$

$$Set 4 = [1, 5, 10, 15, 20, 25, 30, 70] \quad (7)$$

$$Set 5 = [1, 2, 3, \dots, 24, 25, 70] \quad (8)$$

Note that both *Set 1* and *Set 2* attempted to describe evenly distributed observations of the hypervolume between 1 and  $N_{max}$  whilst *Set 3*, *Set 4* and *Set 5* incorporated a higher density in the lower number of sensors. Furthermore, the number of observations highly varied between the sets: *Set 2* and *Set 3* contained only five numbers of sensors (leading to five optimization problems to be solved) whilst *Set 5* contained 26 numbers of sensors (leading to a total of 26 optimization problems).

The method described in Section 2.4 was used to optimally locate the pressure sensors. Note that other methods can be used with distinct objective functions (both in number and formulation). Two pressure sensitivity matrices were computed prior to the optimization problem itself. A variation of Hazen-Williams's pipe roughness coefficient of 10 was considered for the computation of the first matrix. The second matrix was obtained by generating a burst of fixed size for every node of the hydraulic model with a single emitter coefficient of 0.25 (leading to an average burst of 5 L/s). The hydraulic simulations were carried in EPANET [48].

The optimal sensor location for a given number of sensors was formulated as an unconstrained multi-objective optimization problem. The decision variables were the nodes where pressure sensors could be potentially installed, and all the nodes were considered as possible locations. The two objective functions presented in Section 2.4 were considered,

aiming at the maximization of nodal pressure sensitivities to both pipe roughness coefficient variations ( $f_1$ ) and pipe burst events ( $f_2$ ).

The problem was solved once for each number of sensors in each of the five discrete sets. When the number of sensors appeared in more than one set (e.g., 70), the problem was only solved once. The NSGA-II algorithm was applied in the Python environment using the Pymoo package [49] to solve several optimization problems. Discrete variables and integer encoding were used for problem formulation; each discrete location (i.e., node) was translated by an integer value. A population member is a set of locations for the pressure sensors and each variable of a population member represents a possible pressure sensor location.

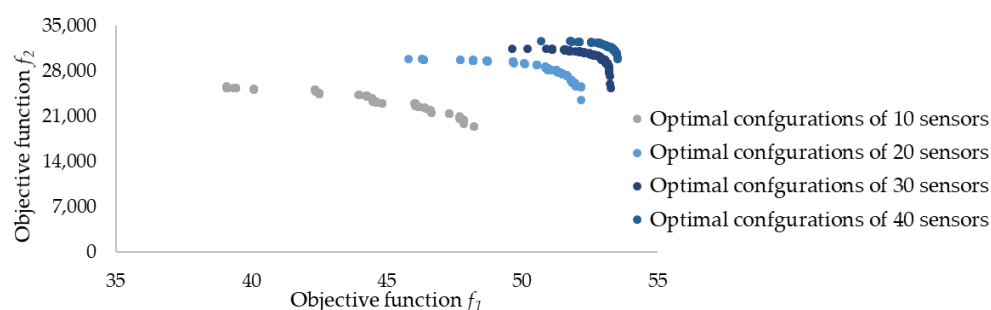
The following NSGA-II parameters were used: random integer sampling and selection operators; integer polynomial mutation with the probability  $p_m = 0.05$  and index parameter  $n_m = 20$ ; simulated binary crossover with probability  $p_c = 0.95$  and index parameter  $n_c = 20$ .

A population size of 100 was considered and all NSGA-II runs were carried out for 500 generations. As such, each NSGA-II run (i.e., solving an optimization problem for a given number of sensors in a given set) resulted in  $500 \times 100 = 50,000$  objective function evaluations. The optimization problems were solved using an Intel Core i5-8250U processor of 1.80 GHz and 8 GB of memory, with a total running time of around 1 h per optimization problem. Table 1 presents the number of objective function evaluations associated with each of the five discrete sets; compare the computation effort (i.e., the number of objective function evaluations) in Set 2 and Set 3 (both with five optimization problems to be solved) with that of Set 5 (with 26 optimization problems to be solved).

**Table 1.** Number of objective function evaluations for each discrete set.

| Sets of Numbers of Sensors              | Number of Observations | Objective Function Evaluations |
|---|------------------------|--------------------------------|
| Set 1 = [1, 10, 20, 30, 40, 50, 60, 70] | 8                      | 400,000                        |
| Set 2 = [1, 10, 30, 50, 70]             | 5                      | 250,000                        |
| Set 3 = [1, 10, 20, 30, 70]             | 5                      | 250,000                        |
| Set 4 = [1, 5, 10, 15, 20, 25, 30, 70]  | 8                      | 400,000                        |
| Set 5 = [1, 2, 3, ..., 24, 25, 70]      | 26                     | 1,300,000                      |

The result of each optimization problem was a Pareto front of optimal pressure sensor locations, according to the objective functions  $f_1$  and  $f_2$ . For the sake of simplicity, the obtained Pareto fronts for 10, 20, 30 and 40 sensors are depicted in Figure 6. Each circle represents an optimal configuration of either 10, 20, 30 or 40 sensors.



**Figure 6.** Example of obtained Pareto fronts for 10, 20, 30 and 40 sensors.

The hypervolume indicator was calculated for each Pareto front of each set of numbers of sensors, leading to a total of five sets of Hypervolume values (i.e., one for each set of numbers of sensors). The global worst point was considered equal to (0,0) since both functions aimed at the maximization of the objective function. The calculation method for

the hypervolume was based on variant 3 of the algorithm proposed by Fonseca et al. [42]. The five sets of hypervolume values are depicted in Figure 7 as black triangles.

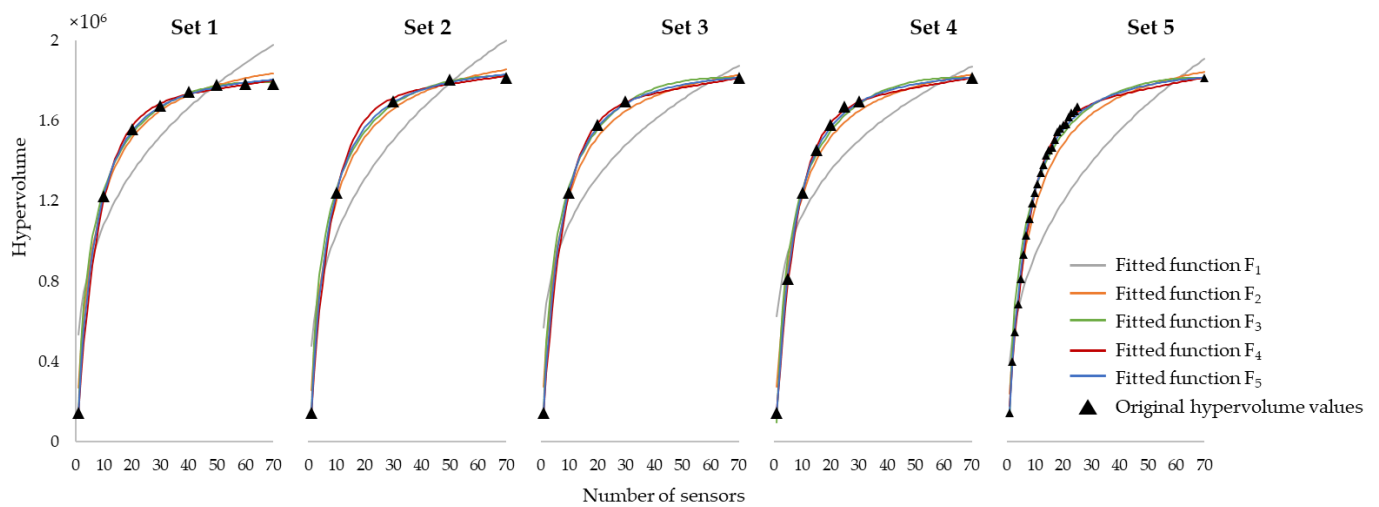


Figure 7. Estimated hypervolume values using different fitted functions  $F_i$  and different discrete sets of numbers of sensors.

Five different functions,  $F_i$ , were considered to describe the hypervolume data as a function of the number of sensors,  $N$ :

$$F_1(N) = aN^b \tag{9}$$

$$F_2(N) = (aN)/(b + N) \tag{10}$$

$$F_3(N) = a + b \log_{10}(N) - cN \tag{11}$$

$$F_4(N) = a e^{bN} + c e^{dN} \tag{12}$$

$$F_5(N) = a(N + b)^c + d \tag{13}$$

where  $a$ ,  $b$ ,  $c$  and  $d$  are the function parameters. These parameters, once optimized, can represent the curve of the hypervolume as a function of the number of sensors. Note that different functions can be considered to describe the increasing behavior one can see in Figure 7 in black triangles.

The Levenberg–Marquardt method was used to fit each function  $F_i$  to each set of hypervolume values by finding the optimal set of parameters. This was done by using MATLAB’s Curve Fitting Toolbox. A reasonable initial parameter guess was found by coarsely gridding the parameter space. This was carried out to cope method sensitivity to the initial parameters. Figure 7 presents in colored lines the estimated hypervolume values obtained by using each model  $F_i$ . The original hypervolume values are represented as triangular black markers.

Table 2 presents the five functions’ specific parameters found using the Levenberg–Marquardt method for Set 1 of hypervolume values.

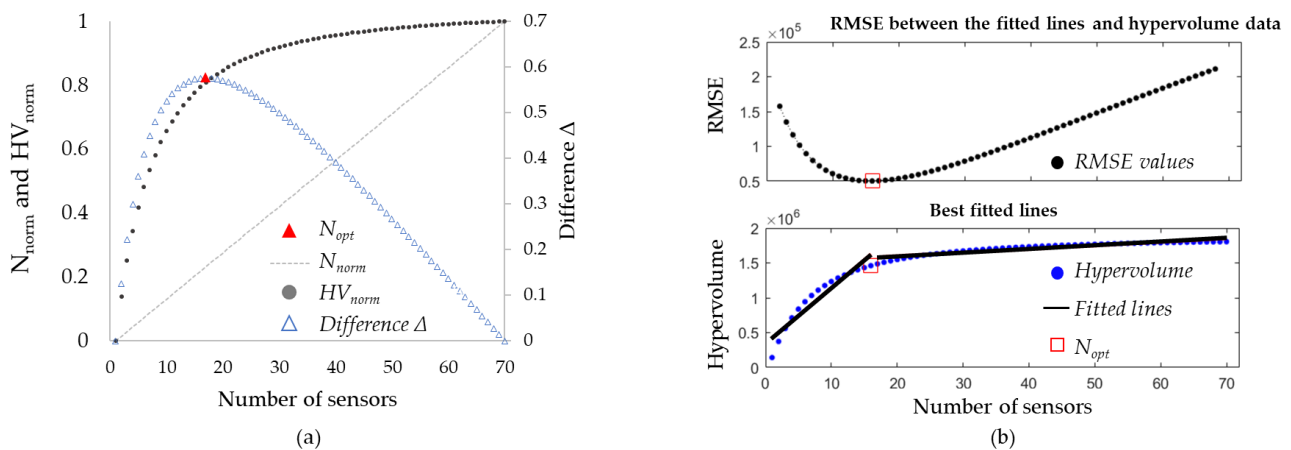
Table 2. Fitted functions’ parameters for Set 1 of hypervolume values.

| Fitting Function                   | $a$          | $b$       | $c$        | $d$       |
|------------------------------------|--------------|-----------|------------|-----------|
| $F_1(N) = aN^b$                    | 532,809      | 0.308     | -          | -         |
| $F_2(N) = (aN)/(b + N)$            | 2,007,063    | 6.465     | -          | -         |
| $F_3(N) = a + b \log_{10}(N) - cN$ | 144,141      | 1,190,045 | 7835       | -         |
| $F_4(N) = a e^{bN} + c e^{dN}$     | 1,665,878    | 0.001     | -1,733,051 | -0.129    |
| $F_5(N) = a(N + b)^c + d$          | -136,128,665 | 10.779    | -1.772     | 1,861,283 |

The RMSE was calculated for each function  $F_i$  of each set of numbers of sensors in order to assist in deciding which function  $F_i$  best describes the hypervolume data. The obtained results are presented in Table 3, with the best fitted function for each set of numbers of sensors presented in bold and shaded areas. Finally, the optimal number of sensors  $N_{opt}$  (considered as the point of maximum curvature) was determined using an automatic detection technique. The two distinct techniques presented in Section 2.7 were used for each function  $F_i$  of each set of numbers of sensors, with the obtained results presented in Table 3. Figure 8 presents the application of both Kneedle method and L-method techniques to the estimated hypervolume for the  $F_5$  function of Set 1.

**Table 3.** RMSE and the optimal number of sensors  $N_{opt}$  for the different fitted functions and discrete sets of sensors. The results for the best-fitted function of each discrete set are presented in bold and shaded area.

| Number of Sensors' Set                  |                            | $F_1(N)$          | $F_2(N)$          | $F_3(N)$                            | $F_4(N)$                            | $F_5(N)$                            |
|---|----------------------------|-------------------|-------------------|-------------------------------------|-------------------------------------|-------------------------------------|
| Set 1 = [1, 10, 20, 30, 40, 50, 60, 70] | RMSE                       | $2.2 \times 10^5$ | $6.0 \times 10^4$ | $1.9 \times 10^4$                   | $2.0 \times 10^4$                   | <b><math>1.5 \times 10^4</math></b> |
|   | $N_{opt}$ (Kneedle method) | 20                | 17                | 16                                  | 18                                  | <b>17</b>                           |
|   | $N_{opt}$ (L-method)       | 20                | 16                | 16                                  | 17                                  | <b>16</b>                           |
| Set 2 = [1, 10, 30, 50, 70]             | RMSE                       | $2.7 \times 10^5$ | $7.6 \times 10^4$ | <b><math>2.2 \times 10^4</math></b> | $3.0 \times 10^4$                   | $2.6 \times 10^4$                   |
|   | $N_{opt}$ (Kneedle method) | 20                | 18                | <b>16</b>                           | 18                                  | 17                                  |
|   | $N_{opt}$ (L-method)       | 20                | 17                | <b>16</b>                           | 16                                  | 16                                  |
| Set 3 = [1, 10, 20, 30, 70]             | RMSE                       | $3.3 \times 10^5$ | $8.9 \times 10^4$ | $2.9 \times 10^4$                   | <b><math>1.0 \times 10^4</math></b> | $3.3 \times 10^4$                   |
|   | $N_{opt}$ (Kneedle method) | 19                | 17                | 16                                  | <b>17</b>                           | 17                                  |
|   | $N_{opt}$ (L-method)       | 20                | 16                | 16                                  | <b>16</b>                           | 16                                  |
| Set 4 = [1, 5, 10, 15, 20, 25, 30, 70]  | RMSE                       | $2.8 \times 10^5$ | $7.9 \times 10^4$ | $6.1 \times 10^4$                   | <b><math>1.4 \times 10^4</math></b> | $2.2 \times 10^4$                   |
|   | $N_{opt}$ (Kneedle method) | 19                | 17                | 16                                  | <b>17</b>                           | 17                                  |
|   | $N_{opt}$ (L-method)       | 19                | 16                | 16                                  | <b>16</b>                           | 16                                  |
| Set 5 = [1, 2, 3, ..., 24, 25, 70]      | RMSE                       | $4.1 \times 10^5$ | $1.3 \times 10^5$ | $6.4 \times 10^4$                   | <b><math>1.5 \times 10^4</math></b> | $2.4 \times 10^4$                   |
|   | $N_{opt}$ (Kneedle method) | 20                | 18                | 16                                  | <b>17</b>                           | 17                                  |
|   | $N_{opt}$ (L-method)       | 21                | 17                | 16                                  | <b>16</b>                           | 16                                  |



**Figure 8.** Optimal number of sensors for the  $F_5$  function of Set 1 using (a) the Kneedle method and (b) the L-method.

#### 4. Discussion

The increase in the number of installed sensors led to a higher total gain. Nonetheless, the marginal gain decreased when considering additional sensors. This is visible in Figure 7 in black triangles. This figure shows that regardless of the considered number of sensors in the set, the function  $F_1$  (grey line) clearly fails to mathematically describe the trend of the hypervolume as a function of the number of sensors, whereas the remaining functions can approximately describe the hypervolume variation.

Table 3 presents better insights regarding the fitting capabilities of the different functions  $F_i$  for each discrete set of numbers of sensors. The best fit corresponds to the smallest values of the RMSE. Complex models  $F_4$  and  $F_5$  (i.e., both with a larger number of parame-

ters) presented the best fitting results (in bold) for most sets of numbers of sensors, except for *Set 3*, for which  $F_3$  led to the lowest RMSE values.

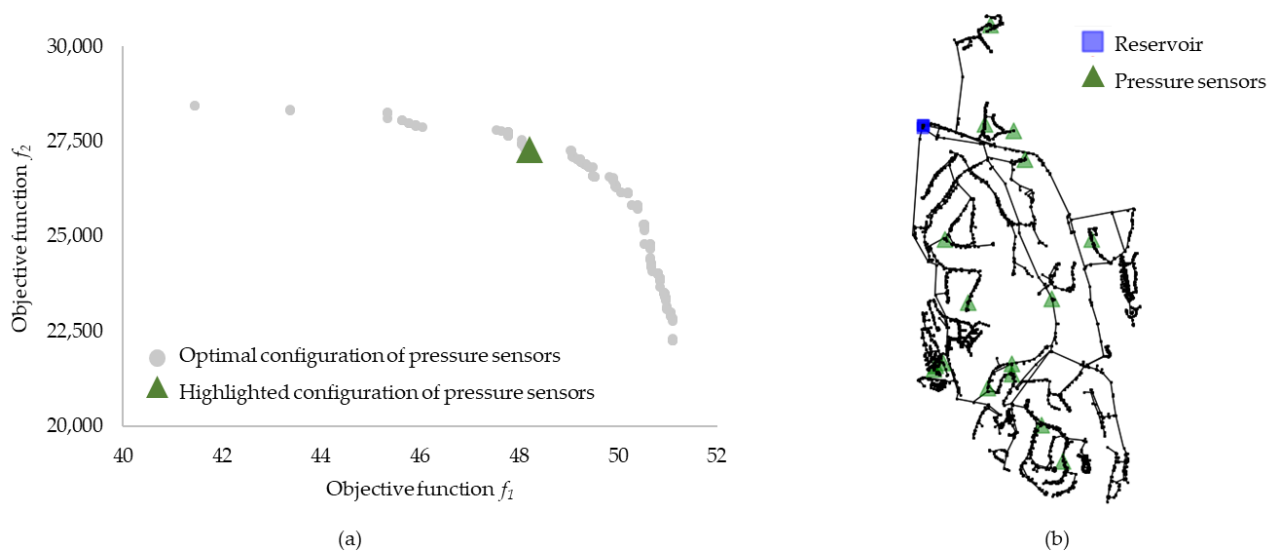
The first robustness analysis was carried out to determine the effect of the different functions  $F$  on the optimal number of sensors. For each technique (Kneedle method and L-method) of each set of numbers of sensors, it was possible to verify that the optimal number of sensors did not significantly vary with the fitted functions  $F_i$  to estimate the hypervolume value (note that results from  $F_1$  were excluded from analysis). Technique 1 presented higher variability in the optimal number of sensors, ranging from 16 to 18. Technique 2 presented smaller variability, between 16 and 17. Therefore, and as long as selected functions  $F_i$  (and respective parameters) can properly represent the hypervolume as a function of the number of sensors, a good estimation of  $N_{opt}$  can be obtained (considering either the Kneedle method or the L-method).

The second robustness analysis was carried out to analyze the effect of different discrete sets of numbers of sensors on the optimal number of sensors. For each technique and function  $F_i$ , it was possible to verify that the optimal number was relatively stable for the different sets, and no differences greater than one were found. Thus, and as long as the selected discrete set of numbers of sensors can properly represent the hypervolume as a function of the number of sensors, a good estimation of  $N_{opt}$  can be obtained. This allows for a significant reduction in the number of optimization problems to be solved as similar results are obtained when both five (in *Set 2* and *Set 3*) or 26 (in *Set 5*) optimization problems are solved.

The third robustness analysis was carried out to assess the effect of the two different “knee/elbow” detection techniques (Kneedle method and L-method) on the optimal number of sensors. No differences greater than two in the optimal number of sensors were found by comparing the pairs of results for both techniques in each discrete set of numbers of sensors and function  $F_i$  as the optimal number varied between 16 and 18 sensors. Furthermore, when the best function  $F_i$  was considered for each set (results in bold in Table 3), the optimal number of sensors was equal to 16 or 17 (depending on the “knee/elbow” detection technique) regardless of the discrete set of numbers of sensors. Based on these results, the Kneedle method was recommended given the straightforward implementation and easier concept understanding.

Overall, the proposed methodology was robust regarding different discrete sets of numbers of sensors, fitting functions and “knee/elbow” automatic detection techniques. The optimal number of sensors for the best fitting function was equal to 16 or 17 (depending on the “knee/elbow” detection method) regardless of the chosen discrete set of numbers of sensors. This allowed for a great reduction in the computational burden associated with the overall analysis as similar results were obtained when considering both *Set 2* and *Set 3* (with five optimization problems to be solved) and *Set 5* (26 optimization problems to be solved). Furthermore, a good approximation between 16 and 18 sensors was found considering any combination of function  $F$  (with the exception of  $F_1$ ), any discrete set of numbers of sensors and any “knee/elbow” detection technique.

Based on these results, the optimal number of sensors was selected as 16. The optimization problem was already solved for 16 sensors (as part of *Set 5*), and the obtained Pareto front is presented in Figure 9a. Each grey point in Figure 9a represents a combination of 16 sensors. Figure 9b depicts a possible optimal location for the 16 sensors in green triangles, associated with the highlighted solution from the Pareto front.



**Figure 9.** (a) Pareto front for 16 sensors; (b) optimal pressure sensor location for 16 sensors.

## 5. Conclusions

This article proposes a six-step methodology to determine the optimal number of pressure sensors. The major advantage of the proposed methodology is that it does not require neither prior knowledge about sensor installation costs nor the consideration of costs in the optimization problem itself to decide on the optimal number of sensors. Furthermore, the methodology can be used with different optimization methods of the pressure sensor location.

In the proposed methodology, the optimal number of pressure sensors is determined by analyzing the relationship between the results of optimization problems and the number of sensors. The total gain of installing a given number of sensors is characterized by using the hypervolume indicator. A trade-off function is derived to describe the gain of installing sensors as a function of the number of sensors. Finally, a technique is used to determine the optimal number of sensors as the point where the inclusion of another sensor is not worth the increment on the total gain.

A real case study was used to demonstrate, assess and discuss the performance of the proposed methodology. Three distinct robustness analyses were carried out and, according to the obtained results, the following key conclusions related to the proposed methodology can be drawn:

- The hypervolume indicator can be used to characterize the total gain of installing pressure sensors whose locations are obtained by solving multi-objective optimization problems.
- A trade-off function can be derived, allowing the characterization of the total gain of installing sensors (i.e., hypervolume) as a function of the number of sensors. This trade-off function fitting process allows for a great reduction in the computational effort associated with the overall analysis, as similar results are obtained when solving both five and 26 multi-objective optimization problems.
- The obtained results are not affected by considering different trade-off functions or different sets of numbers of sensors as long as the selected function and the selected discrete set can properly represent the hypervolume as a function of the number of sensors.
- Both Techniques 1 and 2 lead to similar results. The use of Technique 1 (Kneedle method) is further recommended given the straightforward implementation and easier concept understanding.

In future work, the robustness of the proposed methodology could be further assessed by comparing the obtained results with those obtained by using two distinct methods of

locating those sensors. Furthermore, different quality measures to compare consecutive Pareto fronts and additional techniques to automatically detect the “knee/elbow” could be compared.

**Author Contributions:** Conceptualization, B.F.; methodology, B.F.; investigation, B.F.; writing—original draft preparation, B.F.; writing—review and editing, N.C. and D.C.; visualization, B.F.; supervision, N.C. and D.C.; project administration, N.C.; funding acquisition, N.C. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by Fundação para a Ciência e a Tecnologia through a studentship (reference number SFRH/BD/149392/2019) and the WISDom project (reference number DSAIPA/DS/0089/2018). The APC was funded completely by the WISDom project.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data presented in this study are available upon request from the corresponding author.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

## References

1. Kara, S.; Karadirek, I.E.; Muhammetoglu, A.; Muhammetoglu, H. Real time monitoring and control in water distribution systems for improving operational efficiency. *Desalin. Water Treat.* **2016**, *57*, 11506–11519. [[CrossRef](#)]
2. Weber, R.; Hósz, C. Efficient Technique for Pipe Roughness Calibration and Sensor Placement for Water Distribution Systems. *J. Water Resour. Plan. Manag.* **2020**, *146*, 04019070. [[CrossRef](#)]
3. Ostfeld, A.; Uber, J.G.; Salomons, E.; Berry, J.W.; Hart, W.E.; Phillips, C.A.; Watson, J.-P.; Dorini, G.; Jonkergouw, P.; Kapelan, Z.; et al. The Battle of the Water Sensor Networks (BWSN): A Design Challenge for Engineers and Algorithms. *J. Water Resour. Plan. Manag.* **2008**, *134*, 556–568. [[CrossRef](#)]
4. Aral, M.M.; Guan, J.; Maslia, M.L. Optimal design of sensor placement in water distribution networks. *J. Water Resour. Plan. Manag.* **2010**, *136*, 5–18. [[CrossRef](#)]
5. Weickgenannt, M.; Kapelan, Z.; Blokker, M.; Savic, D.A. Risk-Based Sensor Placement for Contaminant Detection in Water Distribution Systems. *J. Water Resour. Plan. Manag.* **2010**, *136*, 629–636. [[CrossRef](#)]
6. Zhao, Y.; Schwartz, R.; Salomons, E.; Ostfeld, A.; Poor, H.V. New formulation and optimization methods for water sensor placement. *Environ. Model. Softw.* **2016**, *76*, 128–136. [[CrossRef](#)]
7. Naserizade, S.S.; Nikoo, M.R.; Montaseri, H. A risk-based multi-objective model for optimal placement of sensors in water distribution system. *J. Hydrol.* **2018**, *557*, 147–159. [[CrossRef](#)]
8. Ponti, A.; Candelieri, A.; Archetti, F. A New Evolutionary Approach to Optimal Sensor Placement in Water Distribution Networks. *Water* **2021**, *13*, 1625. [[CrossRef](#)]
9. Pérez-Padillo, J.; Morillo, J.G.; Ramirez-Faz, J.; Roldán, M.T.; Montesinos, P. Design and implementation of a pressure monitoring system based on iot for water supply networks. *Sensors* **2020**, *20*, 4247. [[CrossRef](#)]
10. Zanfei, A.; Menapace, A.; Santopietro, S.; Righetti, M. Calibration procedure for water distribution systems: Comparison among hydraulic models. *Water* **2020**, *12*, 1421. [[CrossRef](#)]
11. Sophocleous, S.; Savić, D.; Kapelan, Z. Leak Localization in a Real Water Distribution Network Based on Search-Space Reduction. *J. Water Resour. Plan. Manag.* **2018**, *145*, 04019024. [[CrossRef](#)]
12. Hajibandeh, E.; Nazif, S. Pressure Zoning Approach for Leak Detection in Water Distribution Systems Based on a Multi Objective Ant Colony Optimization. *Water Resour. Manag.* **2018**, *32*, 2287–2300. [[CrossRef](#)]
13. Sun, C.; Parellada, B.; Puig, V.; Cembrano, G. Leak Localization in Water Distribution Networks Using Pressure and Data-Driven Classifier Approach. *Water* **2019**, *12*, 54. [[CrossRef](#)]
14. Cao, H.; Hopfgarten, S.; Ostfeld, A.; Salomons, E.; Li, P. Simultaneous sensor placement and pressure reducing valve localization for pressure control of water distribution systems. *Water* **2019**, *11*, 1352. [[CrossRef](#)]
15. Zhao, M.; Zhang, C.; Liu, H.; Fu, G.; Wang, Y. Optimal sensor placement for pipe burst detection in water distribution systems using cost–benefit analysis. *J. Hydroinform.* **2020**, *22*, 606–618. [[CrossRef](#)]
16. Raei, E.; Shafiee, M.E.; Nikoo, M.R.; Berglund, E. Placing an ensemble of pressure sensors for leak detection in water distribution networks under measurement uncertainty. *J. Hydroinform.* **2019**, *21*, 223–239. [[CrossRef](#)]
17. Casillas, M.; Puig, V.; Garza-Castañón, L.; Rosich, A. Optimal Sensor Placement for Leak Location in Water Distribution Networks Using Genetic Algorithms. *Sensors* **2013**, *13*, 14984–15005. [[CrossRef](#)] [[PubMed](#)]

18. Steffelbauer, D.B.; Fuchs-Hanusch, D. Efficient Sensor Placement for Leak Localization Considering Uncertainties. *Water Resour. Manag.* **2016**, *30*, 5517–5533. [[CrossRef](#)]
19. Sarrate, R.; Blesa, J.; Nejjari, F.; Quevedo, J. Sensor placement for leak detection and location in water distribution networks. *Water Supply* **2014**, *14*, 795–803. [[CrossRef](#)]
20. de Schaetzen, W.B.; Walters, G.; Savic, D. Optimal sampling design for model calibration using shortest path, genetic and entropy algorithms. *Urban Water* **2000**, *2*, 141–152. [[CrossRef](#)]
21. Francés-Chust, J.; Brentan, B.M.; Carpitella, S.; Izquierdo, J.; Montalvo, I. Optimal Placement of Pressure Sensors Using Fuzzy DEMATEL-Based Sensor Influence. *Water* **2020**, *12*, 493. [[CrossRef](#)]
22. Kapelan, Z.S.; Savic, D.A.; Walters, G.A. Optimal Sampling Design Methodologies for Water Distribution Model Calibration. *J. Hydraul. Eng.* **2005**, *131*, 190–200. [[CrossRef](#)]
23. Behzadian, K.; Kapelan, Z.; Savic, D.; Ardeshtir, A. Stochastic sampling design using a multi-objective genetic algorithm and adaptive neural networks. *Environ. Model. Softw.* **2009**, *24*, 530–541. [[CrossRef](#)]
24. Sanz, G.; Pérez, R.; Kapelan, Z.; Savic, D. Leak Detection and Localization through Demand Components Calibration. *J. Water Resour. Plan. Manag.* **2016**, *142*, 04015057. [[CrossRef](#)]
25. Wu, Z.Y.; Sage, P.; Turtle, D. Pressure-Dependent Leak Detection Model and Its Application to a District Water System. *J. Water Resour. Plan. Manag.* **2010**, *136*, 116–128. [[CrossRef](#)]
26. Soldevila, A.; Blesa, J.; Tornil-Sin, S.; Fernandez-Canti, R.M.; Puig, V. Sensor placement for classifier-based leak localization in water distribution networks using hybrid feature selection. *Comput. Chem. Eng.* **2018**, *108*, 152–162. [[CrossRef](#)]
27. Quintiliani, C.; Vertommen, I.; van Laarhoven, K.; van der Vliet, J.; van Thienen, P. Optimal Pressure Sensor Locations for Leak Detection in a Dutch Water Distribution Network. *Environ. Sci. Proc.* **2020**, *2*, 2040. [[CrossRef](#)]
28. Auger, A.; Bader, J.; Brockhoff, D.; Zitzler, E. Theory of the hypervolume indicator: Optimal  $\mu$ -Distributions and the Choice of the Reference Point. In Proceedings of the Tenth ACM SIGEVO workshop on Foundations of Genetic Algorithms—FOGA '09, Orlando, FL, USA, 9–11 January 2009; ACM Press: New York, NY, USA, 2009; pp. 87–102.
29. Blesa, J.; Nejjari, F.; Sarrate, R. Robust sensor placement for leak location: Analysis and design. *J. Hydroinform.* **2016**, *18*, 136–148. [[CrossRef](#)]
30. Lansey, K.E.; El-Shorbagy, W.; Ahmed, I.; Araujo, J.; Haan, C.T. Calibration Assessment and Data Collection for Water Distribution Networks. *J. Hydraul. Eng.* **2001**, *127*, 270–279. [[CrossRef](#)]
31. Laszczyk, M.; Myszkowski, P.B. Survey of quality measures for multi-objective optimization: Construction of complementary set of multi-objective quality measures. *Swarm Evol. Comput.* **2019**, *48*, 109–133. [[CrossRef](#)]
32. Wang, S.; Ali, S.; Yue, T.; Li, Y.; Liaaen, M. A practical guide to select quality indicators for assessing pareto-based search algorithms in search-based software engineering. In Proceedings of the 38th International Conference on Software Engineering, Austin, TX, USA, 14–22 May 2016; ACM: New York, NY, USA, 2016; pp. 631–642.
33. Tan, K.C.; Lee, T.H.; Khor, E.F. Evolutionary algorithms for multi-objective optimization: Performance assessments and comparisons. In Proceedings of the 2001 Congress on Evolutionary Computation (IEEE Cat. No.01TH8546), Seoul, Korea, 27–30 May 2001; IEEE: Piscataway, NJ, USA, 2002; Volume 2, pp. 979–986.
34. Bandyopadhyay, S.; Pal, S.K.; Aruna, B. Multiobjective GAs, Quantitative Indices, and Pattern Classification. *IEEE Trans. Syst. Man Cybern. Part B* **2004**, *34*, 2088–2099. [[CrossRef](#)]
35. Collette, Y.; Siarry, P. Three new metrics to measure the convergence of metaheuristics towards the Pareto frontier and the aesthetic of a set of solutions in biobjective optimization. *Comput. Oper. Res.* **2005**, *32*, 773–792. [[CrossRef](#)]
36. Zitzler, E.; Deb, K.; Thiele, L. Comparison of Multiobjective Evolutionary Algorithms: Empirical Results. *Evol. Comput.* **2000**, *8*, 173–195. [[CrossRef](#)] [[PubMed](#)]
37. Guerreiro, A.P.; Fonseca, C.M.; Paquete, L. The Hypervolume Indicator: Problems and Algorithms. *arXiv* **2020**, arXiv:2005.00515.
38. While, L.; Hingston, P.; Barone, L.; Huband, S. A faster algorithm for calculating hypervolume. *IEEE Trans. Evol. Comput.* **2006**, *10*, 29–38. [[CrossRef](#)]
39. Daulton, S.; Balandat, M.; Bakshy, E. Differentiable Expected Hypervolume Improvement for Parallel Multi-Objective Bayesian Optimization. *arXiv* **2020**, arXiv:2006.05078.
40. Jaskiewicz, A. Improved quick hypervolume algorithm. *Comput. Oper. Res.* **2018**, *90*, 72–83. [[CrossRef](#)]
41. Zhao, G.; Arroyave, R.; Qian, X. Fast Exact Computation of Expected Hypervolume Improvement. *arXiv* **2018**, arXiv:1812.07692.
42. Fonseca, C.M.; Paquete, L.; López-Ibáñez, M. An improved dimension-sweep algorithm for the hypervolume indicator. In Proceedings of the 2006 IEEE International Conference on Evolutionary Computation, Vancouver, BC, Canada, 16–21 July 2006; pp. 1157–1163. [[CrossRef](#)]
43. Yuan, Y.-X. Recent advances in numerical methods for nonlinear equations and nonlinear least squares. *Numer. Algebr. Control Optim.* **2011**, *1*, 15–34. [[CrossRef](#)]
44. Satopaa, V.; Albrecht, J.; Irwin, D.; Raghavan, B. Finding a “Kneedle” in a Haystack: Detecting Knee Points in System Behavior. In Proceedings of the 2011 31st International Conference on Distributed Computing Systems Workshops, Minneapolis, MN, USA, 20–24 June 2011; IEEE: Piscataway, NJ, USA, 2011; pp. 166–171.
45. Zhao, Q.; Hautamaki, V.; Fränti, P. Knee Point Detection in BIC for Detecting the Number of Clusters. In *Advanced Concepts for Intelligent Vision Systems (ACIVS'08), Proceedings of the 10th International Conference, ACIVS 2008, Juan-les-Pins, France, 20–24 October 2008*; Springer: Berlin/Heidelberg, Germany, 2008; Volume 5259, pp. 664–673; ISBN 3540884572.

46. Salvador, S.; Chan, P. Determining the number of clusters/segments in hierarchical clustering/segmentation algorithms. In Proceedings of the 16th IEEE International Conference on Tools with Artificial Intelligence, Boca Raton, FL, USA, 15–17 November 2004; IEEE: Piscataway, NJ, USA, 2004; pp. 576–584.
47. Antunes, M.; Gomes, D.; Aguiar, R.L. Knee/Elbow Estimation Based on First Derivative Threshold. In Proceedings of the 2018 IEEE Fourth International Conference on Big Data Computing Service and Applications (BigDataService), Bamberg, Germany, 26–29 March 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 237–240.
48. Rossman, L. *EPANET 2 User's Manual*; United States Environmental Protection Agency: Washington, DC, USA, 2000.
49. Blank, J.; Deb, K. Pymoo: Multi-Objective Optimization in Python. *IEEE Access* **2020**, *8*, 89497–89509. [[CrossRef](#)]