

Article

Robust Design of a Real-Life Water Distribution Network under Different Demand Scenarios

Ina Vertommen ^{1,*}, Karel van Laarhoven ¹ and Maria da Conceição Cunha ²

¹ KWR Water Research Institute, 3430 BB Nieuwegein, The Netherlands; karel.van.laarhoven@kwrwater.nl

² Department of Civil Engineering, INESC Coimbra—Institute for Systems Engineering and Computers at Coimbra, University of Coimbra, 3004-531 Coimbra, Portugal; mccunha@dec.uc.pt

* Correspondence: ina.vertommen@kwrwater.nl; Tel.: +31-611-599-514

Abstract: In this paper a scenario-based robust optimization approach is proposed to take demand uncertainty into account in the design of water distribution networks. This results in insight in the trade-off between costs and performance of different designs. Within the proposed approach the designer is able to choose the desired degree of risk aversion, and the performance of the design can be assessed based on the water demand effectively supplied under different scenarios. Both future water demand scenarios and scenarios based on historical records are considered. The approach is applied to the design of a real-life water distribution network supplying part of a city in the Netherlands. From the results the relation between costs and performance for different scenarios becomes evident: a more robust design requires higher design costs. Moreover, it is proven that numerical optimization helps finding better design solutions when compared to manual approaches. The developed approach allows water utilities to make informed choices about how much to invest in their infrastructure and how to design it in order to achieve a certain level of robustness.

Keywords: water distribution networks; design; optimization; robustness; uncertainty; water demand

Citation: Vertommen, I.; van Laarhoven, K.; Cunha, M. C. Robust Design of a Real Life Water Distribution Network under Different Demand Scenarios. *Water* **2021**, *13*, 753. <https://doi.org/10.3390/w13060753>

Academic Editor: Stefano Alvisi

Received: 20 January 2021

Accepted: 3 March 2021

Published: 10 March 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 (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The problem of water distribution network (WDN) design consists in the definition of improvement decisions that can optimize the system given a certain objective, or objectives. These decisions can be made under three assumptions: certainty, risk and uncertainty [1,2].

When assuming certainty, the input parameters, such as water demand, are considered to be deterministic and well-known. Practice shows that water utilities often take into account a peak factor (according to equations available in literature), or, for instance, the maximum (hourly) water demand of the past ten years (as done in the Netherlands) for the design of WDN. Sometimes, safety factors are considered to take uncertainty into account. This deterministic approach leads to a design that performs well for the specific water demand considered but may underperform if the water demand turns out differently.

Decision making under risk means that the input parameters are recognized to be uncertain and are assumed to follow certain probability distributions. The problems that consider risk are known as stochastic optimization problems. In academic research, uncertainty in water demand is often taken into account by means of stochastic approaches, which are computationally very demanding, and therefore less attractive for application to optimization of real-life WDN.

Scenario-based robust optimization can be seen as a more viable approach in this context and leads us to decision making under uncertainty. In these problems no

information about the probability distributions is considered. Instead, uncertainty is taken into account by means of a limited number of scenarios, lessening the computational burden, and, not less important, is easy to understand by water utilities and is more in line with current practice. Both stochastic and robust optimization problems are aimed at finding solutions that perform well under any possible realization of the random input variables.

1.1. Water Demand

Water demand is one of the most important and uncertain aspects in the context of drinking water distribution systems. In order to properly answer questions about the design and management of these systems, it is therefore essential to understand and take into account the inherent variability and uncertainty regarding present and future water demand. Water demand varies with the behavior and habits of people, as well as the surrounding environment. A first step in modelling water demand therefore requires the consideration of aspects such as the type of consumer (residential, commercial and industrial uses all have different specific demand needs and patterns), type of day (working or weekend days, holidays or vacation periods lead to different consumption patterns) and weather (both the total daily demand, peak and demand pattern vary with temperature and precipitation). Specifically for residential use, aspects such as the type of house (for instance with or without garden), household composition, and living area (households in rural, urban or suburban areas have different demand patterns) are also important [3].

The literature is rich in approaches for modelling water demand. Some examples consider regression analysis and black box models, including artificial neural networks, random forests, support vector regression, among others [4]; end-use models, such as SIMDEUM for modelling both residential and non-residential demand [5]; future scenario studies [6]; short-term demand forecasting (48 h) for operational uses, based on historical data and measurements [7], or auto-regressive integrated moving average and machine learning, in order to construct scenarios with similar statistics of the historical measurements [8]; linear demand growth models considering different phases in a planning horizon [9,10]; machine learning techniques, to determine extreme values for water demand in the future based on climate scenario's and vacation behaviour [11]; statistical approaches and analysis of demand time series [12–14], also including scaling laws [15,16] and joint probabilities [17].

It becomes clear that depending on the purpose for which the water demand is modelled, models can focus on the present water demand or on the future water demand, and different scales can be considered: for a single consumer or for an entire area, on an annual basis, on a daily basis, or on demand patterns or peak demands. It is therefore important to choose and apply the right water demand models, with the right resolution and on the right scale (in space and time) to the different problems concerning water distribution systems. For instance, short-term variability (within one day) is important for operational management, e.g., for the optimal control of pumping stations, daily patterns with a short time resolution are important for water quality modelling, total daily demand and daily peak demand factors are important for determining the production capacity needed to supply a supply area and, hour or instantaneous peak water demand at the different nodes in a network model is important for the design of a WDN.

1.2. Optimization Problems

For many years the objective of the optimal design of WDN was to satisfy demands at least cost. The traditional deterministic optimization problem is formulated as follows: the minimization of costs as the objective function (considered as a function of the pipe diameters in the network) and pipe diameters as the decision variables. The constraints are the satisfaction of demands (given by the continuity equation) and the minimum requirements on the pressure heads at each node of the system. The input parameters,

such as water demands, were considered as being deterministic. This deterministic approach is a major limitation. In fact, in real-life, the demands are not foreseen with certainty and as pointed out in the literature, deterministic approaches can lead to under designed networks, increasing the risk of failure due to demands that exceed the design values. From practice, we see that the opposite may also be true, as the risk averse nature of water utilities may lead to grossly (costly) over designed networks in order to be able to withstand any eventuality, which often lead to water quality issues. Nevertheless, the certainty in input parameters was usually assumed in the design process, due to lack of reliability measures and knowledge about computational feasibility [18].

In the last decades and thanks to the increase of computational capacity, researchers began to address the optimal design of WDS under uncertainty. Several approaches have been developed to this end, from stochastic approaches, deterministic equivalents and surrogate approaches, to scenario-based approaches and to flexible or phased design of WDN. Table 1 provides an overview of these approaches and some examples of literature references.

Table 1. Concise overview of approaches for the design of WDN that include uncertainty in the decision process.

Approach	Concept	Options	Examples of Literature References
Stochastic approach	Reliability is usually expressed in terms of the probability that the system does not fail, where failure is seen as not satisfying required nodal pressure. Failure probability is computed through sampling-based techniques, such as Monte Carlo simulation.	Single-objective: chance constrained optimization problems wherein costs are minimized and the desired level of reliability is defined through a constraint.	[19,20]
		Multi-objective: minimization of costs and maximization of reliability.	[21–27]
Deterministic equivalent	Safety margins are added (redundancy) to the constraints or to the uncertain variables resulting in a deterministic equivalent formulation of the uncertain problem. An extra reliability assessment through Monte Carlo simulation is sometimes performed a posteriori. Optimization problems can be single- or multi-objective.	FOSM	[28,29]
		FORM	[30,31]
		Integration method and Sampling method	[20]
		θ -method	[32]
Robust-counterpart	[33–36]		
Surrogate approach	A surrogate for reliability is used, such as the resilience index, which defines head surplus as a measure of reliability of the system. Several resilience indicators exist. The optimization problem is often formulated as a multi-objective problem minimizing costs and maximizing the resilience index.		[37–43]
Fuzzy logic	The uncertainty is represented using fuzzy theory with membership functions describing the uncertainty in demands. The design problems is usually formulated as a multi-objective problem considering minimization of costs and maximization of reliability.		[44]
Scenario-based		Model and solution robustness	[45–47]

	In scenario-based robust optimization approaches the optimization problem is solved for a limited number of scenarios with given probabilities of occurrence. The solution best performing over all scenarios is chosen. The performance can be expressed by model and solutions robustness or regret functions.	Regret models	[48–51]
	The design and a long-term planning of network interventions are considered side by side as the best way to deal with uncertain future conditions imposed to a WDN. This results in a phased design plan for the entire planning horizon of a network.	Flexible design	[52–58]
Phased and flexible design approaches		Multi-criteria decision analysis	[9]

Stochastic, deterministic equivalent, surrogate, and fuzzy logic approaches deal with uncertainty by generating random demands according to a probability distribution with defined mean and variance. In a stochastic approach, for instance, a system is designed to meet a certain level of reliability, e.g., 99%. This means that a 1% probability of failure does exist and how the system performs in this situation is unknown. Some authors argue that it is better to design systems that perform “well enough” under all possible circumstances [59]. This is where the concept of robustness emerges: robustness is understood as the ability of the system to continue to function under different conditions. In this context, uncertainty is also dealt with in a different way, by explicitly considering different possible realizations of the uncertain parameters, this is, different scenarios, and look for a solution which is feasible and as close as possible to the optimum for all of them, instead of defining parameters through probability distributions. Scenarios differ from predictions or forecasts in the sense that they represent a range of plausible futures rather than a single favorable outcome. It can also be said that risk is central in this approach: the probability and effects of scenarios are explicitly taken into account. Although scenario-based planning techniques exist for some time [60], robust optimization is a relatively new approach to handle optimization problems affected by uncertainty, and has only recently gained notoriety in different fields of science. In [61] an overview is provided of used methodologies and indicators for robustness in different research areas regarding water distribution systems, namely, design and planning, operation and management. The authors argue more research is needed to properly understand the relation between robustness and the other resilience components, such as redundancy, rapidity and resourcefulness. Little attention is given to the uncertainty of input variables. In scenario-based robust optimization models, input variables are often described by scenarios with a given probability of occurrence. The optimization model then takes all scenarios into account in order to arrive at a solution that is “robust.” But how can robustness be quantified? In [47] two types of variables are defined: design variables and control decision variables. The first refer to the variables whose optimal values do not depend on the uncertain parameters, while the latter refer to the variables whose values depend on the uncertain parameters and on the optimal value of the design variables. A generic robust optimization model is then proposed, consisting in minimizing an objective function, bound to structural constraints (subject to the design variables) and control

constraints (subject to both design and control decision variables). Following, a set of scenarios is introduced, each with an associated probability. By considering the scenarios, the control constraints might become restrictive, and even lead to a declaration of infeasibility, hindering the model to find solutions. In order to avoid this, the model is allowed to consider nearly feasible solutions, or feasible solutions under most (but not all) scenarios. This aspect is what leads to a particular characteristic of robust optimization: allowing some constraints to be violated, by considering a specific objective function. This objective function consists of two terms: a first term quantifying optimality (or solution) robustness, and a second term quantifying model robustness, in the form of a feasibility penalty function. This function is used to penalize failures in satisfying the control constraints under some scenarios and is what mainly distinguishes the robust optimization approach from other models dealing with uncertainty. It allows the model to find solutions, even if they are not feasible for all scenarios.

The robust optimization model presented in [47] has been further developed and adapted to suit different applications such as the expansion of a telecommunications network [62], the optimization of chemical reactors and the optimization of a fermentation process [63], the design of biological reactors [64], and many others. Robust optimization has also found different applications on water supply systems. Watkins and McKinney [65] introduced robust optimization in water resources problems as a tool to assess the trade-off between cost, systems performance and reliability; Carr et al. [46,66] addressed the problem for sensor placement in municipal water systems; Cunha and Sousa [45,46] and Marques et al. [67] applied a robust optimization approach to the design of WDN. In [45] the objective function comprises the minimization of the total cost consisting of the sum of two terms: (1) the deviation of the networks construction cost, and (2) a penalty cost for the deviation from the desirable nodal heads. The undelivered demand (due to pressure deficits) is considered as a second level of robustness, added as a penalty term to the objective function, in [46]. The considered scenarios include deterministic peak demands combined with extreme events such as pipe failures or a fire at specific locations in the network. These scenarios are based on expert judgement.

Robust optimization models can also be formulated in terms of the regret of a solution. The total of overpayment (when a larger system is constructed than is necessary as the future plays out, the cost that exceeds the actual requirements is an overpayment) and supplementary (when the implemented design is insufficient to supply actual needs the explicit cost of expanding an undersigned system to meet the requirements is a supplementary expense) costs is called "regret cost." This means that the regret is understood as the difference between the cost of a solution obtained for a set of scenarios, and the cost of the optimal solution for each scenario considered individually. Regret models will not be further described in this paper, but for the interested reader, some relevant references are [1,48–51,68].

More recently, the consideration of deep uncertainty has started to gain attention in water resources optimization. This emerges from the recognition that long-term future conditions should be modelled by considering multiple plausible futures, where it is no longer possible to estimate probabilities of their occurrence, in alternative (or in complementary) of quantifiable (local) uncertainty (through stochastic processes and statistical analysis). Optimization explicitly considering deep uncertainty in its framework is a challenge, due to the computational burden of such approach. Therefore, robustness evaluation in this context is often done post-optimization. In [69] the authors developed a computational efficient optimization approach, by means of a metamodel, for the optimal sequencing of water resources infrastructure under deep uncertainty, wherein robustness is included as an explicit objective during the optimization process. This approach might open the way for more applications in the water sector.

1.3. Motivation and Real-Life Networks

From deterministic, to stochastic, to robust optimization models, researchers have produced significant developments on the subject of optimal design of WDN [70]. Traditional deterministic models are very sensitive to modifications in working conditions, making designs unreliable if reality turns out to be different than planned. Stochastic and robust optimization models are significantly less sensitive to changes in working conditions, and therefore definitely the future for the design cost-effective and safe systems. On the other hand, the stochastic design comes with some challenges. This type of approach requires a significant amount of data, consists of a complex formulation and can quickly become computationally heavy and lengthy. This holds especially true when considering the size of real-life WDN, with hundreds or thousands of nodes and links, where the computation of optimized deterministic solutions is already a challenge by itself. All this contributes to making engineers reluctant to use and apply stochastic models in real-life problems. A scenario-based approach has the advantage of not requiring a probability distribution for the uncertain variables and might appear somewhat more straightforward in application. Even so, the size of scenario-based optimization models is larger and more complex than deterministic models, and increases with the number of considered scenarios, becoming computationally more demanding.

In this contribution we intend to demonstrate how a scenario-based robust optimization approach can be applied to a real-life WDN and what the added value of doing so is.

2. Materials and Methods

2.1. Introduction

It is clear from the literature review that several approaches are possible for taking uncertainty in input parameters into account when designing WDN. In this paper a scenario-based robust optimization approach is proposed, i.e., one in which the uncertainty (in this case with respect to water demand) is described by means of scenarios with corresponding probabilities. This approach has been chosen for the following reasons:

- Although the required computation time of such an approach is larger than deterministic approaches, it is still expected to be manageable for real-life WDN (with hundreds or thousands of nodes and links), as opposed to e.g., stochastic approaches in which a much higher number of Monte Carlo simulations (order of thousands) need to be computed.
- This approach recognizes that, in face of uncertainty, it is not always possible to obtain feasible solutions, i.e., that infeasibilities will inevitably arise. By recognizing this, the approach will generate solutions that present the decision maker with the least number of infeasibilities to be dealt with.
- The approach is applicable to different types of scenarios and can therefore also be used when considering long-term future scenarios in which changes in water demand consumption patterns and/or the addition of new neighborhoods or demand points in the network are taken into account.
- The approach provides insight into how well (or how poorly) a design continues to perform under various scenarios, in contrast to approaches which only look at whether or not the design meets certain boundary conditions.
- The designer can determine for himself how important meeting the constraints is by assigning lower or higher penalty coefficients to the optimization problem.
- The approach fits in well with current practice at, for instance, Dutch water utilities, where the highest measured demand of the past ten years, is considered as the (one) design scenario. This increases the chance of successful implementation.

2.2. Optimization Model

Scenario-based robust optimization models often include two terms: a term for solution robustness and a term for model robustness. The former, measures how close the solution remains to the optimum for any realization of the scenarios (and has to do with “closeness” between the values of the design variables). While the latter measures the feasibility of the solution, i.e., how good the solution performs under the different scenarios, often measured through a feasibility penalty function.

In this contribution we explore the model robustness term, i.e., the performance of the design under the different scenarios. An obvious choice for the model robustness term is the expected value of the function describing the performance. However, the expected value ignores the distribution of the performance values around the mean, and thus also the risk aspect which a decision maker needs to deal with. In cases where decision makers are risk averse, alternative approaches capable of describing and handling risk are more appropriate. One possible approach consists of the so-called mean-variance models. In this type of models, the variance of the outcomes serves as a measure for risk, with a higher variance meaning the outcome (in this case performance) is much in doubt.

The proposed optimization model is thus based on a “mean-variance model” [38,46,47,71] and is aimed at finding a design that minimizes the sum of the costs and the mean and variance of a feasibility function over a set of demand scenarios, i.e.,:

$$\text{Min. } \sum_{j=1}^{NP} L_j \times C_j(D_j) + \mu(f_s) + \lambda \cdot \text{Var}(f_s) \quad (1)$$

where:

$$\mu = \sum_{s=1}^{NS} p_s \cdot f_s \quad (2)$$

$$\text{Var}(f_s) = \sum_{s=1}^{NS} p_s \cdot (f_s - \mu)^2 \quad (3)$$

The first term in the objective function (Equation (1)) corresponds to the cost of the solution to be implemented: cost of pipe j as a function of its length L_j and diameter associated cost $C_j(D_j)$ for all NP pipes in the system. The second and third terms in the objective function are the mean and variance of a function f_s , which describes the performance of the solution under the scenarios: the second term in Equation (1) being the weighted average of the function f_s for the set S of all scenarios, with NS members, see Equation (2), and the third term in Equation (1) being the variance of the function f_s for the set S of all scenarios, see Equation (3). The probability of occurrence of each scenario is given by p_s . The coefficient λ is the so-called variance-factor chosen by the designer and indicates a degree of risk aversion. By taking into account the mean and the variance of the function f_s a distinction can be made between designs with the same mean but different performances under the different scenarios. Designs for which the performance is better, i.e., less penalized, and for which performance deviates less between scenarios, are thus preferred. When only the mean is taken into account, a design that performs well on average but very poorly on a specific scenario can be chosen. It can be said that by taking the variance into account it is easier to control the risk of poor performance.

The function f_s takes into account the performance of the design under the different water demand scenarios and is described by the following feasibility penalty function:

$$f_s = C_{pen} \cdot \left(1 - \frac{\sum_i^{NN} q_{i,s}}{\sum_i^{NN} d_{i,s}}\right) \quad (4)$$

where unsatisfied demands are penalized (being $d_{i,s}$ the demand at each node i for each scenario s and $q_{i,s}$ the actual delivered water to the node in scenario s , both for all NN nodes). This function is based on the “satisfaction rate” [38] and describes the performance of the design based on the degree of satisfaction of water demand under the different water demand scenarios. Although there are various definitions of robustness, according

to [69] robustness metrics based on satisficing criteria are most appropriate, as they align with the way the performance of water resources systems is generally assessed. For a given design (for which costs are determined in the objective function), for each node i and each scenario s the actual water delivered at a single moment in time ($q_{i,s}$) is determined by the hydraulic simulation of the network model. The water supplied at each node depends on the pressure at the node and can be calculated through pressure driven analysis, which can be described by [72]:

$$q_{i,s} = \begin{cases} 0 & \text{if } P_{i,s} < P_{i,0} \\ d_{i,s} \left(\frac{P_{i,s} - P_{i,0}}{P_{i,min} - P_{i,0}} \right)^\gamma & \text{if } P_{i,0} \leq P_{i,s} < P_{i,min} \\ d_{i,s} & \text{if } P_{i,s} \geq P_{i,min} \end{cases} \quad (5)$$

where, $P_{i,s}$ is the actual nodal pressure at node i and scenario s , $P_{i,0}$ is the minimum pressure to allow any flow to the node, and $P_{i,min}$ is the service pressure to fully satisfy nodal demand. The exponent γ is usually set to 0.5.

If the water effectively supplied is lower than the water demand ($d_{i,s}$), a penalty is given. The magnitude of the penalty depends on the penalty coefficient C_{pen} (chosen by the designer) and the extent to which the water demand is not met. Failure to meet the water demand is therefore a disadvantage to the design. It is important to choose a suitable value for the penalty coefficient. This can be done either by (1) calculating the problem with different values and choosing the most appropriate value on the basis of results and interpretation, or (2) on the basis of performance costs which, based on expert knowledge or established business strategy, can be described monetarily.

The objective function is constrained as usual in the optimization models for the design of WDN:

- Hydraulic equilibrium constraints (satisfaction of flows and head loss in pipes)
- The diameters for the pipes can only be chosen from a list of commercially available diameters and only one diameter can be assigned to each pipe
- Minimum pressure requirements

The decision variables are the diameters of the pipes, D_j , in the network model. By solving the optimization model, it is possible to provide a decision maker with the information on how to dimension a network (i.e., which diameters to choose for the pipes) in order to achieve a certain level of robustness.

2.3. Water Demand Scenarios

The proposed optimization model requires the definition of water demand scenarios and probabilities of occurrence.

A water demand scenario is understood as a combination of (peak) water demands which occur simultaneously at the different consumption nodes in the network and can be described by means of a vector:

$$D_s = [d_{1,s}, d_{2,s}, \dots, d_{n,s}], \quad \text{for } i = 1, 2, \dots, n \text{ and } s = 1, 2, \dots, S \quad (6)$$

where D_s is the demand vector for scenario s , and $d_{i,s}$ the (peak) demand at node i for scenario s . The probability of occurrence of the scenario is p_s .

To determine D_s and p_s different approaches are possible. Consulting a panel of experts (an approach often followed in scenario-based optimization) is seen as a good solution but leads to subjective quantification of scenarios and probabilities. Following a more objective approach is desirable, and different methods are available in literature. However, there is no general consensus on which type of approach is best suited to which application, or which information and level of detail should be taken into account. In this contribution two different approaches are followed, one based on historical measurements and one based on exploring alternative future scenarios, both explained in the following sections.

2.3.1. Demand Scenarios Based on Historical Records

The first approach proposed for generating water demand scenarios is a top-down approach based on historical records measured at pumping stations. One can say that in this approach uncertainty is characterized as being statistical or probabilistic [73]. In this approach, the demand pattern (and peak factor) at the pumping station is distributed equally among all nodes (and users) in the area supplied by the pumping station. This means that it is assumed that all users have the same demand pattern and that the peak demand occurs at all nodes in the network at the same time. This is of course a simplification of reality. This approach allows to estimate the probability distribution for peak demands, based on long-term measurements (time series), assuming that this distribution is constant over time. The advantage is that the statistics of water demand are based on measurements, so the “real” variability is taken into account. Moreover, this approach is followed by Dutch water utilities to determine the design demand: the highest consumption of the last 10 years measured at pumping stations supplying a supply area. In this way, the proposed approach is in fact an extension of current practice, but instead of determining one single peak demand, it determines several peak demands, D_s , with different probabilities p_s . The disadvantages of the approach are that (1) it requires long-term measurements (not often available), (2) it assumes the same demand pattern for all users in a network and (3) it assumes that the probability distribution of past measurements is representative of future behaviour. The following steps are proposed to generate water demand scenarios from flow measurements at a pumping station, or inlet of a supply area:

1. Data collection and preparation: these steps include the collection of long-term time series measured at the pumping station or inlet of a supply area, the identification of gaps and erroneous measurements and, consequent correction of the time series.
2. Statistical analysis: includes the estimation of the time series of peak demand factors, i.e., the maximum demand occurring each day (at a minute or hourly basis, depending on the available data) divided by the average demand over the entire measured period, and estimation of the corresponding cumulative probabilities.
3. Scenarios: this step includes the choice of the desired number of scenarios to consider, the choice of the same number of peak demand factors, the estimation of scenario probabilities (cumulative probability of the scenario minus the cumulative probability of the scenario with lower peak demand factor) and the assignment of these peak demand factors to the nodes in the network model (if necessary, updating the average water demand in the network model).

2.3.2. Future Water Demand Scenarios

The approached previously described results in historical water demand scenarios. This implies the assumption that the future will be as the present. However, the future may differ due to e.g., changes in population (including average household and family composition), buildings, activities, water using devices, people’s behaviour (e.g., more environmentally conscious or comfort oriented behaviour) and circumstances (e.g., climate change). Drinking water systems, designed on the basis of historical peak demands, will not work optimally when water demand changes dramatically, e.g., by installing comfort rain showers instead of water-saving showers. An alternative would be the consideration of multiple plausible scenarios [73]. The steps for the top-down approach for generating historical water demand scenarios can be adapted and used to generate future demands scenarios. In this case, the starting point is not a time series of measurements but requires the generation of a time series of future water demand values. Such a time series can be generated on the basis of the approach developed in [6]. In the aforementioned study 13 scenarios were developed, one being the current average water demand and 12 scenarios are based on changes in demographics, policy and technology. Table 2 contains a description of the scenarios.

Table 2. Future water demand scenarios described in [6].

Scenario	Name	Description
F1	Now	Baseline: current situation
F2	RC	Regional Communities: per capita demand declines because the economic downfall results in (water) saving behavior, coupled with decreasing population. The average age of the population increases.
F3	SE	Strong Europe: despite low economic growth, mobility increases due to open borders. Personal hygiene habits have changes with an increase in shower frequency. Water pricing based on real costs drives alternative water resources to be adapted on a larger scale, e.g., rain water tanks for watering the garden.
F4	TM	Transatlantic Market: population growth causes increase in drinking water demand. Innovation is aimed at luxury and wellness products.
F5	GE	Global Economy: economic growth causes an increase in consumption. Innovations are aimed at luxury and wellness. People shower longer and water their garden more frequently to cope with the effects of climate change.
F6	Dual	Toilet, laundry machine and outside tap are not supplied by the drinking water distribution systems.
F7	Eco_RC	Based on RC but with innovative sanitation concepts. 100% adaption of 1 L flushing toilets.
F8	Lux.	Luxury, based on current situation with 100% adaptation of luxurious showers.
F9	GE+	Based on GE but with a shower frequency of one shower per day.
F10	Leak	Based on current situation but with leakage equal to 20%.
F11	Lux. + dual	Based on current situation with 100% adaptation of luxurious shower, with dual system for toilet, laundry and outside tap.
F12	Eco+	Adoption of innovative sanitation concepts plus water use efficient showers, washing machines and dishwashers.
F13	DP	Diminishing population: 30% reduction of the population in the area due to emigration (empty houses).

For all these 13 future scenarios the water demand for an average day, and for a given supply area, can be simulated with SIMDEUM [74]. For the approach proposed in this paper, it is necessary to consider, not the demand on an average day, but different peak demand scenarios and corresponding probabilities. This means that to use the scenarios developed in [6], it is necessary to multiply these by a peak factor and assign probabilities of occurrence. To do so, we assumed the peak factors based on historical measurements, and assigned an equal probability (due to unpredictability) to each scenario. As with the approach described in the previous sub-section, the average and peak demand factors are then assigned to each node in the network model.

2.4. Optimization Tool

To solve the proposed optimization problem the generic optimization tool for drinking water networks, Gondwana [75,76] was used. Gondwana has been recently updated to perform hydraulic simulations in EPANET 2.2 [77], and uses the Inspyred library [78] for metaheuristic optimization methods, in particular (modified) genetic algorithms are used.

Applying optimization techniques to real-life WDN is not without challenges [79], one of them being the computational effort involved in exploring (very) large solution spaces and being able to converge to optimal solutions. In order to deal with this, specific variators to tune a genetic algorithm to the optimization of a least cost design were built in Gondwana, namely the heuristic “flatiron” and the “list proximity” mutators [76]. Classic mutation can cause a larger diameter pipe to be surrounded by smaller diameter pipes, which is hydraulically insensible. The flatiron mutator speeds up convergence by

detecting and “smoothing out” these artifacts. In this way, the flatiron mutator guides the search through the solution space and helps reducing the number of iterations needed to achieve convergence. The list proximity mutator enhances convergence by using system specific knowledge to generate solutions highly likely to be viable, specifically by limiting the possible outcomes of a mutation to diameters close to the original value. This does not guide the search but avoids spending time evaluating unfeasible solutions. Besides, the current design values of the WDN (installed pipe diameters) are used in the initial population. These designs are often relatively good solutions, and thus a good starting point for the search.

2.5. Case Study and Workflow

The proposed methodology was applied to the network model representing part of the WDN serving city S in the Netherlands (chosen to be kept anonymous by the water utility). The network model consists of 497 nodes, 474 pipes and one reservoir, see Figure 1. For real-life networks, this is relatively small. It is therefore considered as a good (first) case study to test the feasibility of applying a scenario-based robust optimization model to real-life WDN. The available diameters are summarized in Table 3. A meter price of 1 euro per mm diameter can be assumed to determine the construction costs.

The flowchart in Figure 2 illustrates the work process followed, starting from the generation of demand scenarios, serving as inputs for the optimization models, the computation of results with the use of Gondwana, and finalizing with the post-evaluation of these results under the water demand scenarios not considered during the optimization process.

Table 3. Available pipe diameters for the re-design of the WDN.

Pipe Diameters (mm)					
34	58.2	101.6	147.6	184.6	230.8

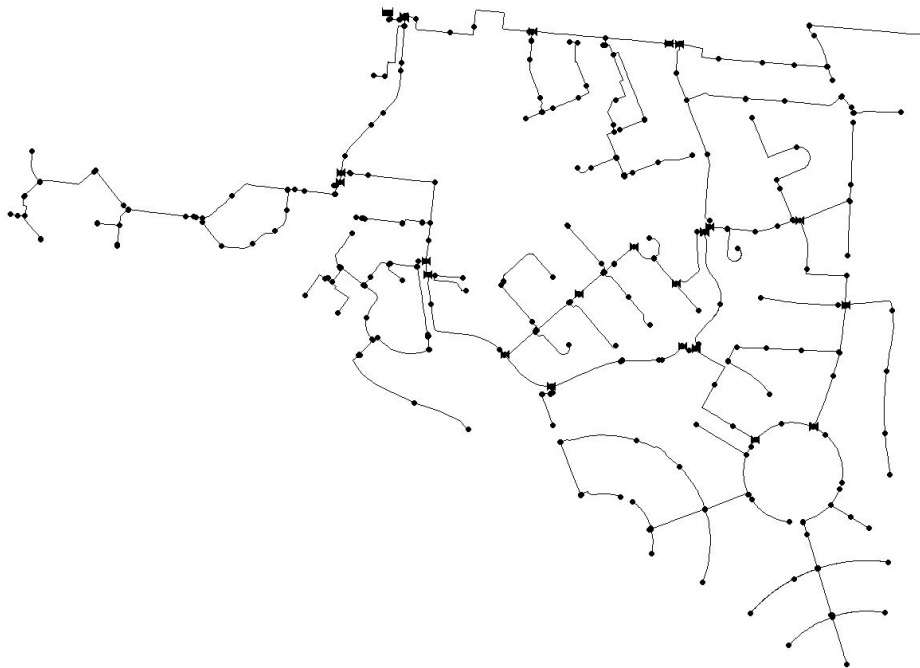


Figure 1. Network model for the WDN serving part of city S in the Netherlands. This network model consists of 497 nodes, 474 pipes and one reservoir. Valves are also depicted in the figure.

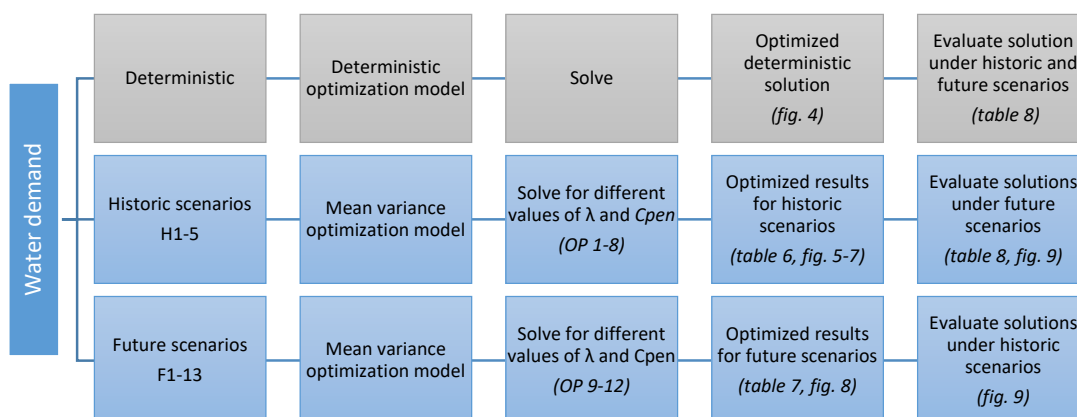


Figure 2. Flowchart illustrating the work process, from the generation of the water demand scenarios as inputs to the optimization models, the different considered optimization models, and lastly the computation and evaluation of different design solutions.

3. Results

3.1. Water Demand Scenarios

3.1.1. Historical Scenarios

The Dutch water utility Dunea provided time series corresponding to 14 years of measurements of the total consumption of the supply area of Wassenaar (ca. 27,000 inhabitants), in the Netherlands, with a 5-min time resolution. Together with the water utility, the time series was checked for gaps and erroneous data and corrected where necessary. The steps described in Section 2.3 were considered to generate water demand scenarios. In Table 4 an example with 5 scenarios is provided. For each scenario the peak factor, cumulative probability and scenario probability are given. As can be seen from the results, the peak factor does not vary very much, being between 2.32 and 2.82 for the chosen scenarios.

Table 4. Example of 5 water demand scenarios and corresponding probabilities estimated based on the measurements at the supply area of Wassenaar, The Netherlands (Dunea).

Scenario	Peak Factor	Scenario Probability
H1	2.32	0.53
H2	2.45	0.24
H3	2.56	0.14
H4	2.64	0.05
H5	2.82	0.04

3.1.2. Future Scenarios

With regard to the future water demand scenarios, in [6] the proposed methodology was used to generate the total average water demand for the network model serving part of city S in the Netherlands, see results in Table 3. It was then decided to take the peak factor with an exceeding probability equal to 1% from Wassenaar’s historical data into account to determine the corresponding peak water demand for the 13 future scenarios. This means that it is assumed that the average water demand changes according to demographic and technological developments described in the future scenarios, but that the peak factor remains equal to the historical peak factor. In the case study, this means that for each future scenario the average water demand of the scenario is multiplied by 2.82 (see Table 5). Of course, this is only an assumption to illustrate the case study. In a real application, this aspect certainly merits more attention. Based on the description of

the scenarios it is to be expected that peak factors for the various future scenarios differ. For example, in scenario GE the peak factor is expected to be higher than in scenario SE, in which rainwater is used to water the garden. With respect to the probability of occurrence of each scenario it was decided to assign an equal probability (due to unpredictability) to each scenario, i.e., equal to 1/13.

Table 5. Average water demand for city S for different future scenarios [6], and assumed values for the peak factor and scenario probabilities.

Scenario	Average Demand (m ³ day ⁻¹)	Peak Factor	Probability
F1-Now	363	2.82	1/13
F2-RC	250	2.82	1/13
F3-SE	255	2.82	1/13
F4-TM	280	2.82	1/13
F5-GE	310	2.82	1/13
F6-Dual	165	2.82	1/13
F7-Eco	265	2.82	1/13
F8-Lux	510	2.82	1/13
F9-GE+	350	2.82	1/13
F10-Leak	440	2.82	1/13
F11-Lux+	320	2.82	1/13
F12-Eco+	140	2.82	1/13
F13-DP	225	2.82	1/13

3.2. Optimization Results

In order to assess the outcomes of the proposed scenario-based robust optimization model, the problem was solved for both the demand scenarios based on historical data and the future demand scenarios, and for different values of the penalty coefficient (C_{pen} in Equation (4)) and the variance-factor (λ in Equation (1)). A minimum pressure equal to 10 m was considered as a constraint, and a service pressure ($P_{i,min}$) equal to 20 m was considered to fully satisfy nodal demand. By considering pressure-driven demand analysis, it is possible to compute the demand that is actually delivered to each node of the network for each considered scenario.

To put the results into context, the pipe diameters of the current infrastructure are depicted below in Figure 3. These pipe diameters lead to a design cost of 792 k€. With this design the lowest pressure in the network during the peak demand in the model (corresponding to a peak factor (PF) equal to 2.77) is equal to 34.15 m. To further contextualize the results, a deterministic optimization problem was solved considering the peak demand factor in the network model (2.77), and a minimum pressure constraint equal to 20 m (i.e., the service pressure to supply 100% of the water demand). This deterministic problem corresponds to a specific case of the optimization model described in Section 2.2, wherein only the first term of Equation (1) is considered (i.e., minimization of costs as a function of the diameter and length of the pipes) and a minimum pressure constraint of 20 m. The obtained design costs are in this case 459 k€, after 1×10^6 function evaluations. The corresponding pipe diameters are depicted in Figure 4. This first result shows the added value of considering optimization techniques for designing real-life WDN: it is possible to achieve a significantly leaner network while at the same time satisfying all deterministic peak demands in the model.

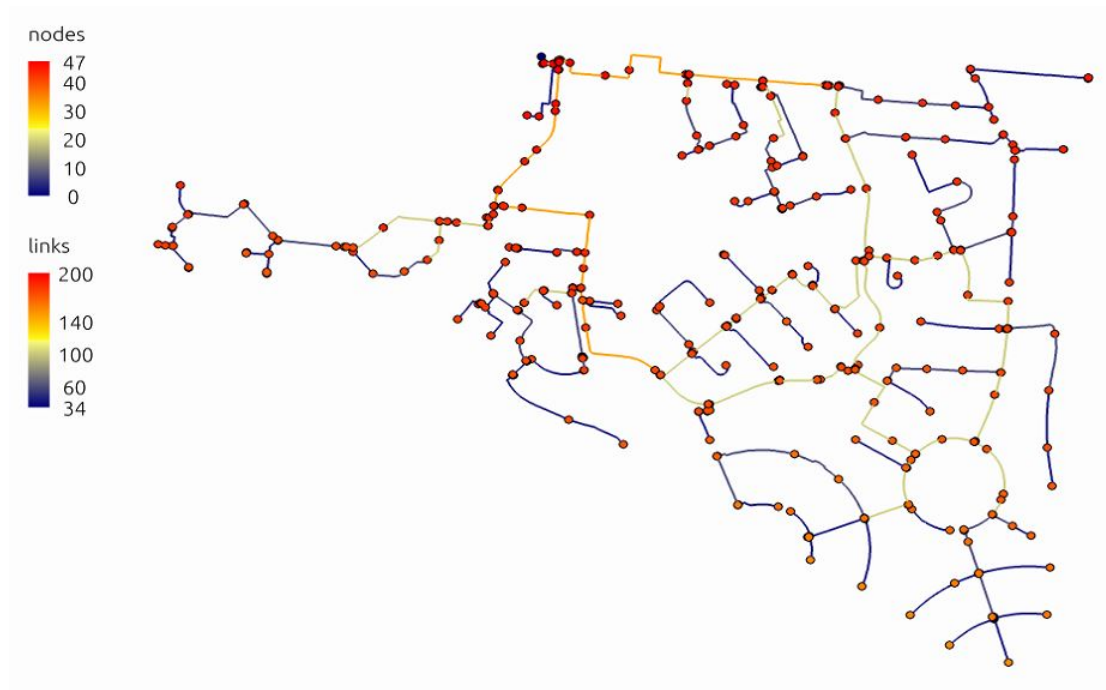


Figure 3. Pipe diameters of the infrastructure currently installed. The different colors indicate the diameter for each pipe (links in mm) and the pressure at the nodes (nodes in m) for the deterministic water demand in the hydraulic model (PF = 2.77). The minimum pressure in the network is 34.15 m.

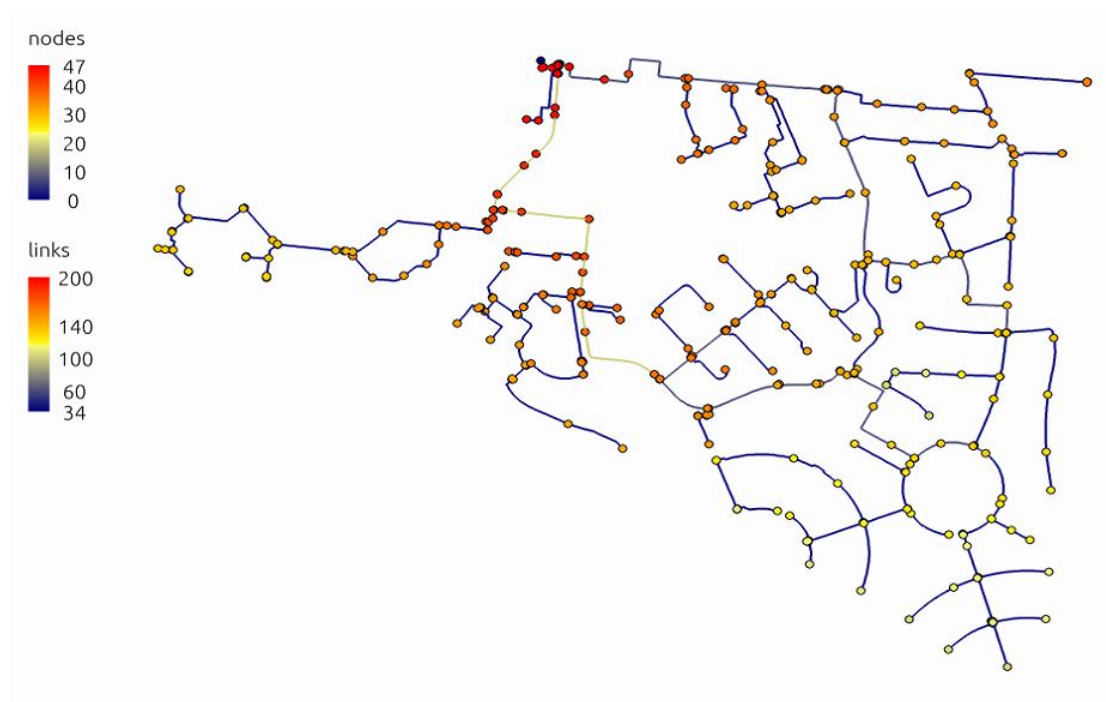


Figure 4. Pipe diameters for the optimized deterministic design of the WDN. The different colors indicate the pipe diameter for each pipe (links in mm) and the pressure at the nodes (nodes in m) for the deterministic water demand in the hydraulic model (PF = 2.77). The minimum pressure requirement of 20 m is met at all nodes in the network. Diameters are significantly decreased when compared to the diameters of the currently installed network (Figure 2).

The obtained results for both historical and future demand scenarios are summarized in the following sections. An evaluation of the performance of the design solutions obtained considering the historic demand scenarios on the future demand scenarios, and vice-versa, is also performed.

3.2.1. Optimization Results for Historical Scenarios

Table 6 summarizes the obtained design results when considering the historic demand scenarios. For each of the solved optimization problems (OP) it is reported: the considered variance-factor (λ in column 2), penalty coefficient (C_{pen} in column 3), the outcome design costs of the chosen solution (column 4) and the corresponding performance under each of the considered historic demand scenarios and the weighted total, in terms of undelivered demand (columns 5–10). Undelivered demand equal to zero means that the demand of the scenario is fully satisfied. The results were obtained after a maximum of 1×10^6 function evaluations. Of course, for each solution the diameters of all pipes in the network were obtained, and this is the information decision makers need when planning their infrastructure. Since the network comprises 474 pipes, the corresponding diameters are not extensively reported in the results.

Table 6. Summary of the results obtained for the different optimization problems considered.

OP	λ	C_{pen}	Costs (k€)	Undelivered Demand (m ³) – Historic Scenarios					
				H1	H2	H3	H4	H5	Total
1	0.1	1×10^0	433.4	3.8×10^{-1}	7.3×10^{-1}	1.1×10^0	1.4×10^0	2.2×10^0	6.9×10^{-1}
2		1×10^2	435.1	2.7×10^{-1}	5.8×10^{-1}	9.2×10^{-1}	1.2×10^0	1.9×10^0	5.5×10^{-1}
3		1×10^4	437.7	0.0×10^0	0.0×10^0	0.0×10^0	6.6×10^{-3}	4.5×10^{-2}	2.1×10^{-3}
4		1×10^6	481.0	0.0×10^0	0.0×10^0	0.0×10^0	0.0×10^0	1.8×10^{-3}	7.1×10^{-5}
5	1	1×10^0	434.8	2.9×10^{-1}	5.9×10^{-1}	9.1×10^{-1}	1.2×10^0	2.0×10^0	5.6×10^{-1}
6		1×10^2	443.8	1.4×10^{-1}	3.0×10^{-1}	5.0×10^{-1}	6.8×10^{-1}	1.2×10^0	3.0×10^{-1}
7		1×10^4	444.4	0.0×10^0	0.0×10^0	7.6×10^{-6}	2.0×10^{-5}	1.8×10^{-4}	9.1×10^{-6}
8		1×10^6	509.2	0.0×10^0	0.0×10^0	0.0×10^0	0.0×10^0	0.0×10^0	0.0×10^0

In the optimization problems numbered 1–8 the mean-variance robust model is solved taking into consideration different values for the variance-factor and penalty coefficients for not satisfying water demand. In particular, the penalty coefficients vary between 1 and 1×10^6 , and the variance-factors are equal to 0.1 and 1. This gives an idea of how these parameters influence the results and/or push the optimization process in a certain direction.

From the results summarized in the table it can be seen that higher penalty coefficient leads to higher design costs and a better performance under the different scenarios: the total weighted undelivered demand decreases. The performance under each individual scenario also becomes clear: the undelivered demand increases from the historical scenarios H1 to H5, as these scenarios become increasingly more demanding (peak demand increases from H1 to H5). When increasing the penalty coefficient, the undelivered demand in each scenario decreases, and for the higher penalty coefficients, the demand is fully delivered for some (or all) scenarios. This behavior is enhanced when considering the higher variance-factor: for the same penalty coefficients, the costs are higher and the total undelivered demand under the scenarios is lower. Figures 5 and 6 provide further insight into these relationships.

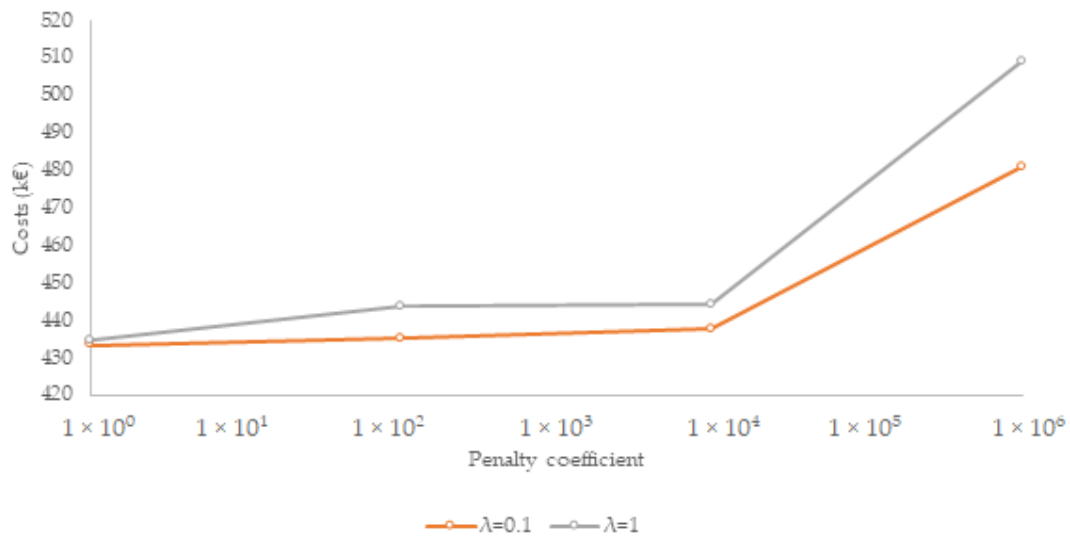


Figure 5. Relation between the design costs and the considered variance-factor (λ) and penalty coefficient (C_{pen}) in the optimization model (considering historic demand scenarios). The different colors indicate the considered variance-factor in the optimization model. The costs are higher for the higher variance-factor ($\lambda = 1$).

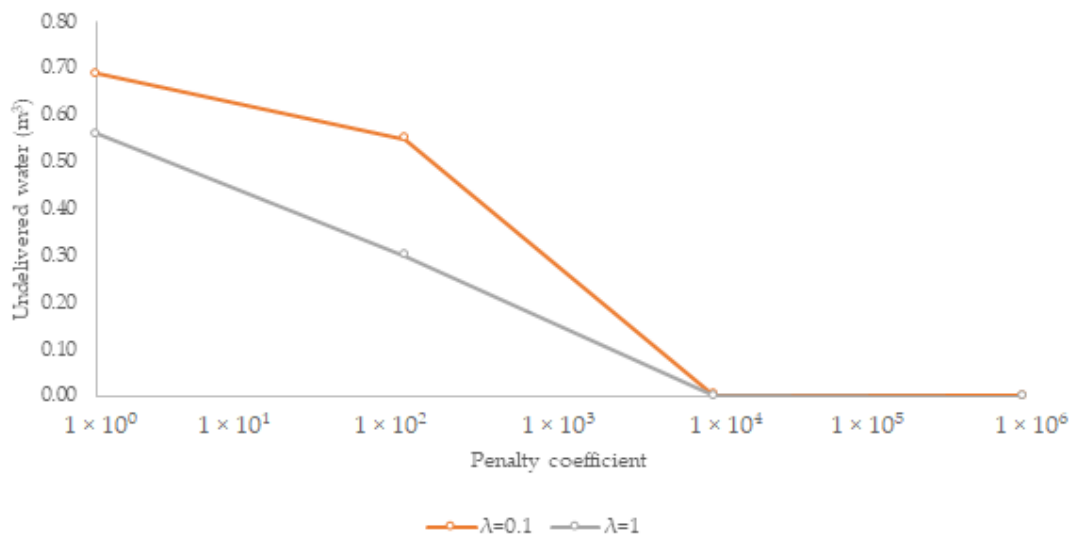


Figure 6. Relation between the total undelivered demand (m³ during peak hour) and the considered variance-factor (λ) and penalty coefficient (C_{pen}) in the optimization model (considering historic demand scenarios). The different colors indicate the considered variance-factor in the optimization model. The total undelivered demand under the historic scenarios is lower for the higher variance-factor ($\lambda = 1$).

As expected, Figures 5 and 6 show that cheaper designs are obtained when lower penalty coefficients are considered, but these designs do not fully meet the water demand under the considered scenarios. As the penalty coefficients increase the designs become more expensive but improve in performance, i.e., the network's capacity to actually deliver water demand under all different demand scenarios.

In terms of the design costs the influence of the variance-factor is clear: for a variance-factor equal to 1 the design costs are higher than the design costs obtained for a variance-factor equal to 0.1, and this difference increases for higher penalty coefficients. This is expected, since higher variance-factors 'push' design costs to be more expensive, in order to reduce risk.

In terms of the performance, by taking a higher variance-factor into consideration, solutions have lower undelivered demand and converge faster to solutions with zero or close to zero undelivered demand.

It has to be noted that the undelivered demand in the case study is always relatively low, so it is valid to question if the water utility would be willing to invest more in order to reduce further an (already) very low demand deficit.

Figure 7 provides further insight into the influence of the variance-factor on the results. From this it can be seen that the difference in variance between performances (in terms of undelivered water) under the various historic scenarios, for a variance-factor equal to 0.1 (orange line) and 1 (grey line) is significant. The variance between the undelivered demand under different scenarios of the designs obtained by taking a higher variance-factor is lower, and thus these designs are more robust, although there is of course also a difference in terms of design costs.

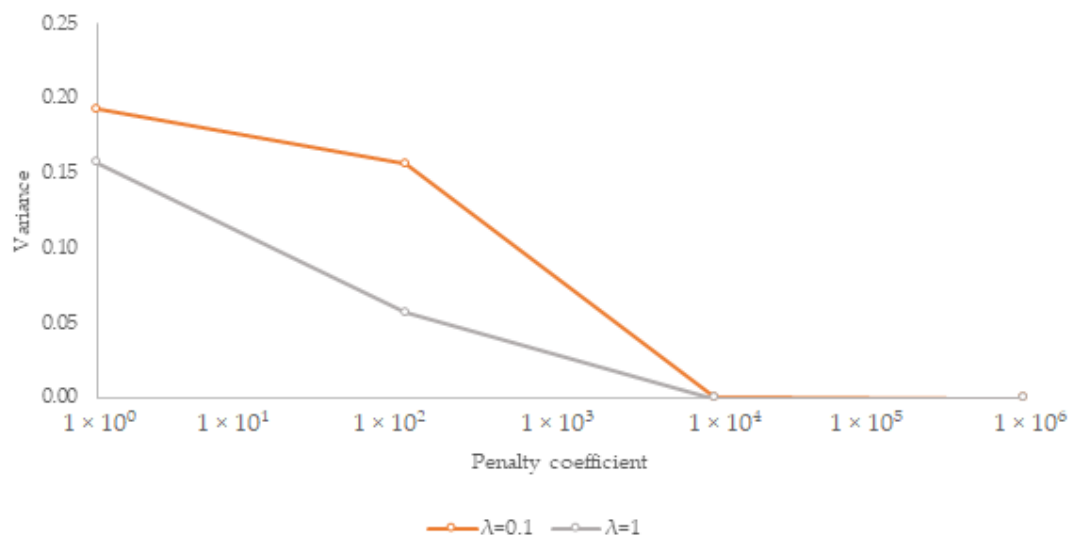


Figure 7. Variance between the performance of the solutions under different historic scenarios (i.e., the undelivered demand), for different values of the variance-factor ($\lambda = 0.1$ in orange and $\lambda = 1$ in grey) and penalty coefficients (C_{pen}). The variance between the undelivered demand under different scenarios decreases with the penalty coefficient and is kept lower for a higher variance. A higher variance-factor increases the robustness of solutions.

3.2.2. Optimization Results for Future Scenarios

In the optimization problems numbered 9–12 in Table 7, the mean-variance robust model is solved for the 13 future water demand scenarios, taking into consideration different values for the penalty coefficient and a variance-factor equal to 1. For each of the solved optimization problems (OP) it is reported: the penalty coefficient (column 2), the outcome design costs of the chosen solution (column 3) and the corresponding performance under each of the considered future demand scenarios and the weighted total, in terms of undelivered demand (columns 4–17). Undelivered demand equal to zero means that the demand of the scenario is fully satisfied. The results were obtained after a maximum of 1×10^6 function evaluations. Note that for the future demand scenarios, the obtained design solutions for the different penalty coefficients, always fully satisfy the demands for scenarios F1–7, F9 and F11–13. Only for scenarios F8 (Lux.) and F10 (Leak) the demand is not always satisfied.

Table 7. Summary of the results obtained for the different optimization problems considering the future demand scenarios.

OP	C_{pen}	Costs (k€)	Undelivered Demand (m ³)—Future Scenarios													Total	
			F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13		
9	1×10^2	480.8	0	0	0	0	0	0	0	0	1.7×10^0	0	2.6×10^{-1}	0	0	0	1.5×10^{-1}
10	1×10^3	482.7	0	0	0	0	0	0	0	0	1.9×10^{-4}	0	6.3×10^{-8}	0	0	0	1.5×10^{-5}
11	1×10^4	486.9	0	0	0	0	0	0	0	0	4.6×10^{-6}	0	0.0×10^0	0	0	0	3.5×10^{-7}
12	1×10^6	501.7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Figure 8 illustrates the relation between costs and undelivered demand for the different obtained design solutions.

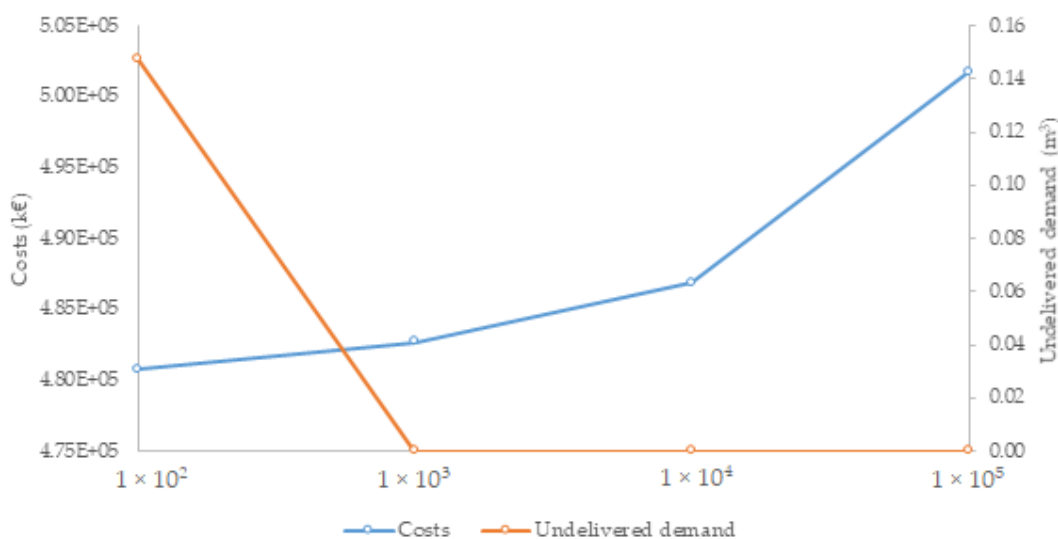


Figure 8. Relation between the design costs (axis y1, blue curve) and the total undelivered demand (axis y2, in m³ during peak hour, orange curve,) for different values of the penalty coefficient (C_{pen}) in the optimization model when taking future demand scenarios into consideration. Fully satisfying demand under all future scenarios requires higher investment costs in the network.

Figure 8 shows that as the penalty coefficient increases, so do the design costs, and the amount of undelivered water decreases. Design costs are higher than in the optimization problems with water demand scenarios based on historical data, but still lower than the current design. The water demand varies more when considering the future scenarios, but it was assumed all scenarios have the same probability of occurrence. Thus, the scenario with the highest water demand weighs as much as the scenario with the lowest water demand. As a result, larger networks are designed. In the optimization problems with scenarios based on historical water demand, the probabilities are different for each scenario: the most demanding scenario has a small chance of occurrence and therefore weighs less in the optimization problem. Assigning different probabilities to the future scenarios would therefore possibly lead to different results.

3.2.3. Design Trends and Performance of Design Solutions

Figure 9a–d illustrates some of the obtained design solutions, namely the pipe diameters obtained considering different values for the variance factor and the penalty coefficients, and historic and future scenarios. The different colors in the figures represent different pipe diameters, usually ranging from 34 to 148 mm (with one exception in Figure 9c, where pipes closest to the reservoir have a diameter of 184.6 mm). The costs of the solutions, and the total undelivered demand for both historic and future scenarios are also

summarized under each figure. It can be seen from the figures that there is a “backbone” for the infrastructure, comprised of pipes with larger diameters, with subsequent pipes with smaller diameters (dark blue, 34 mm). The differences in design are mainly visible along this “backbone”: a first reinforcement of the system happens along the path that has already larger diameters (see changes from Figure 9a to b and d), where more pipes along this path are increased in diameter. An alternative reinforcement (Figure 9c, which is able to satisfy demands in all historic and future scenarios, but at a much higher cost) chooses to increase the pipe diameters along a second path.

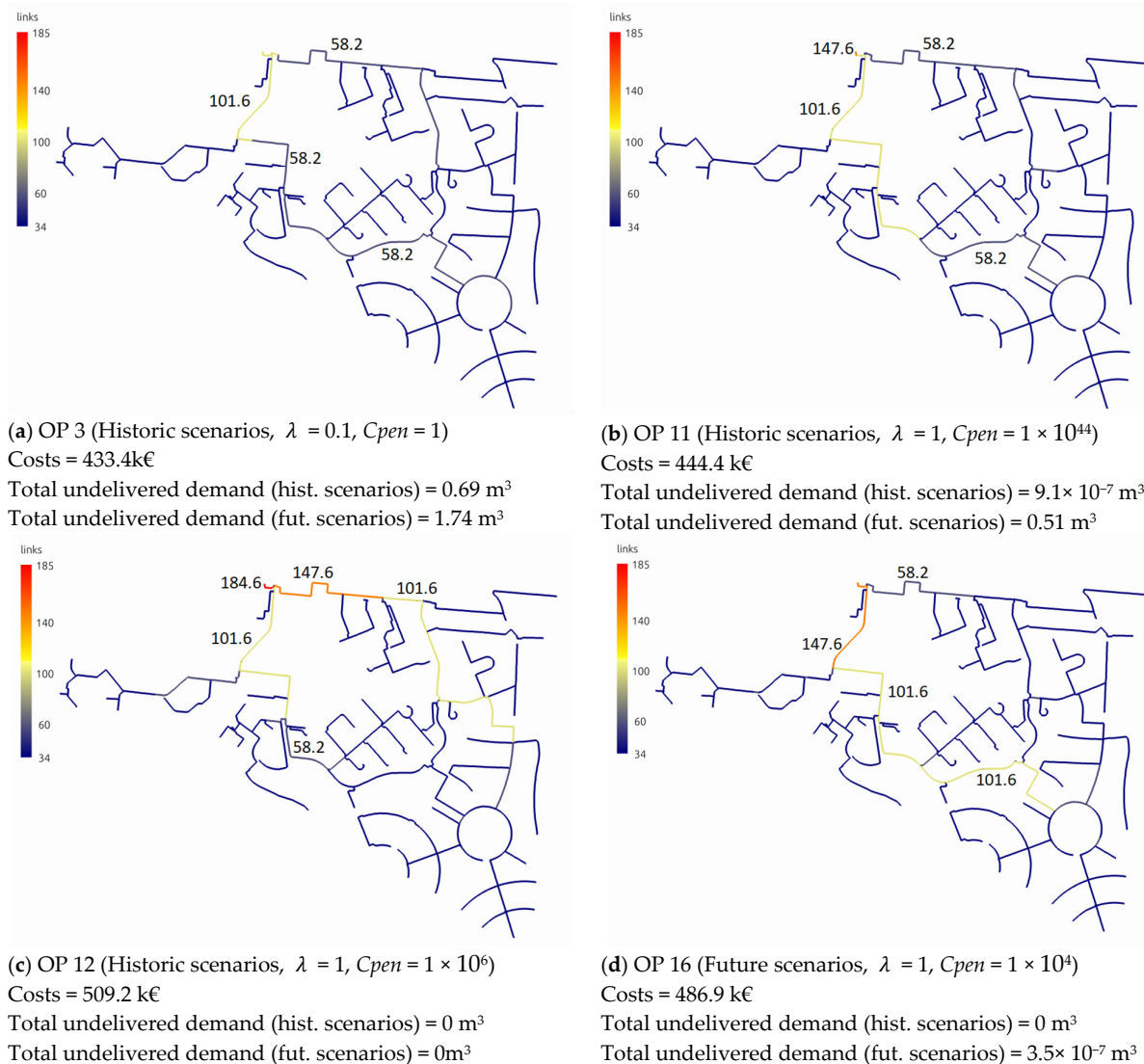


Figure 9. Design solutions obtained for different optimization problems, considering different scenarios (historic or future) and different values for the penalty coefficient (C_{pen}) and variance factors (λ). The different colors indicate the pipe diameters, ranging from 34 to 184.6 mm. The diameters along some of the paths (constituted by several pipes) are indicated with numbers. Designs with lower costs are not able to fully satisfy the demands for both historic and future scenarios (for instance, solutions (a,b)). By increasing more pipe diameters, design costs increase, but it is possible to fully satisfy demand under all scenarios, both historic and future (solution (c)), or for all historic scenarios and almost all future scenarios (solution (d)). This provides the decision maker with the information on how to dimension the network to achieve a certain level of robustness.

The performance of the design solutions obtained considering the historic demand scenarios under the future demand scenarios was also evaluated. Table 8 summarizes

these results. The performance of the deterministic design is also included. The deterministic design solutions are not able to meet the demand of future scenarios 8 and 10, but performs relatively well. Regarding the design solutions obtained considering the historic demand scenarios, these are, in general, not able to meet the demand of different future scenarios. Future demand scenarios 2, 6, and 12, are always fully satisfied. Scenarios 3, 7 and 13 are mostly satisfied. Scenarios 8 and 10, followed by scenario 9, are the ones putting more stress on the system. Only the more robust design solution, obtained for a variance-factor equal to 1 and the highest considered penalty coefficient is able to meet the demand under all possible future scenarios.

Table 8. Performance of the design solutions obtained considering the historic scenarios under the future demand scenarios. The performance is evaluated in terms of undelivered demand (m^3).

OP	λ	C_{pen}	Undelivered Demand (m^3)—Future Scenarios													
			F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13	Total
1		1×10^0	2.20	0.00	0.01	0.15	0.70	0.00	0.03	10.56	1.90	6.08	0.96	0.00	0.01	1.74
2	0.1	1×10^2	1.91	0.00	0.01	0.09	0.56	0.00	0.02	9.94	1.64	5.59	0.79	0.00	0.01	1.58
3		1×10^4	0.05	0.00	0.00	0.00	0.00	0.00	0.00	4.90	0.03	1.85	0.00	0.00	0.00	0.53
4		1×10^6	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.46	0.00	0.37	0.00	0.00	0.00	0.14
5		1×10^0	1.96	0.00	0.01	0.11	0.57	0.00	0.03	10.04	1.67	5.84	0.79	0.00	0.01	1.62
6	1	1×10^2	1.16	0.00	0.00	0.05	0.29	0.00	0.01	7.25	0.98	3.82	0.43	0.00	0.00	1.08
7		1×10^4	0.00	0.00	0.00	0.00	0.00	0.00	0.00	4.65	0.00	1.93	0.00	0.00	0.00	0.51
8		1×10^6	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Det.	-	-	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2.26	0.00	0.44	0.00	0.00	0.00	0.21

The same exercise was performed for the solutions obtained considering the future demand scenarios. In this case, the different design solutions are always able to meet the demand under all historic scenarios. The same happens for the optimized deterministic design solution, which takes a “harder” minimum pressure constraint into consideration.

4. Discussion

This research shows that it is feasible to apply scenario-based robust optimization methods to the design of real-life WDN, and that, by means of the chosen ‘mean-variance’ model, focused on the model robustness term, it is possible to take different water demand scenarios into account during the design process, resulting in more insight into the performance of a design and ultimately in a more robust solution. To our knowledge, although mean-variance models are widely applied in different fields of science, this approach (including modelling both historic and future demand scenarios) has not been applied to the design of a real-life WDN before. By applying this model to a case-study a trade-off between robustness and design costs is quantified. This allows the decision maker to make an informed and substantiated choice to accept some relatively small underperformance of a design in extreme situations in favor of a substantial cost reduction, or to not accept it if that is desirable. The designer controls the degree of risk aversion by adjusting a penalty coefficient for underperformance and a variance-factor to take variance between scenario performances into account. The core and final outcome is then, to provide the decision maker with the optimal diameter for each pipe in the network necessary to install in order to achieve the chosen level of robustness. In this way, a water utility knows how to dimension its network when replacing pipes.

With regard to the scenarios itself, identifying the scenarios and assigning probabilities to them is a daunting and difficult task. From the case study in this contribution, it is shown that it is possible to compute substantiated water demand scenarios based on historical records of water consumption. This also makes it possible to assign probabilities to different peak factors, which is a strong advantage. However, the

obtained results also show that this leads to design solutions less able to cope with out of the ordinary changes in water demand. It is therefore also important to explore scenarios outside the historical range. Estimating future scenarios for water demand remains however somewhat more difficult; in particular, assigning probabilities to these type of scenarios remains a subjective step and there is room for improvement on this aspect, with regard to the approach followed in this contribution. For instance, one might think of extending the proposed approach, by estimating future daily and peak demands, in function of climate change, spread of vacation periods and specific characteristics of the supply areas. For example, the model described in [11] can be used as a basis. With this model it is possible to estimate average and daily factors for future water demand. Hour- or instant peak factors, important for the design of WDN, are not currently predicted by these models. This means, therefore, that additional research is needed in order to use the aforementioned approach to generate water demand scenarios for the design of WDN. Another aspect to consider regarding the water demand scenarios, is that, in this paper, a top-down approach is followed, where the peak factor measured at a pumping station is attributed to all nodes of the network model. A bottom-up approach would also be an interesting approach. In such an approach demands are simulated per node in the network model, depending on the characteristics of customers at each node. The choice for a top-down or bottom-up approach should be mirrored to the type of distribution network considered. For example, a top-down method would be more appropriate for larger urban areas, whereas for smaller neighbourhoods a bottom-up approach seems more appropriate.

The case study further shows that, for the deterministic approach, applying numerical optimization techniques results in a significantly smaller (and therefore cheaper) design (the costs of the optimized design are only 58% of the costs of the current existing infrastructure if rehabilitated as is), while still meeting the water demand at all nodes. A leaner design is not only cheaper but is also better for water quality, reducing residence time and increasing flow velocity, which in turn improves customer satisfaction and also reduces the need for flushing pipes. When sizing a network, designers deal with huge solution spaces. Numerical optimization is definitely a valuable tool to explore these in an efficient manner.

With an approach based on the mean-variance model, and focused on the model robustness term, it is possible to know how to dimension a network in order to satisfy demand under different scenarios. As expected, more expensive designs better meet pressure and water demand under different scenarios. Quantifying this is valuable for decision makers. The different obtained solutions show which pipes need to be reinforced (and by how much) in order to cope with the more extreme future scenarios. The variance-factor has an important impact on results: considering a higher variance-factor leads to designs that perform better under the various scenarios than when a lower variance-factor is considered, for the same penalty coefficients. The variance between the performances for the different scenarios is also much smaller for the higher variance-factor. This means that the designer is more certain of how the design performs under different scenarios.

The influence of the considered scenarios should also not be overlooked. The obtained results indicate that considering future demand scenarios, with larger differences amongst them (and in this case assuming the same probabilities of occurrence) leads to more robust (although more expensive) solutions. These solutions perform well also for different peak demand factors derived from the historic data. The same cannot be said about the designs obtained considering the historic demand scenarios; in general, these solutions underperform for some of the future demand scenarios, exception being the solution obtained for the highest variance-factor and penalty coefficient considered.

5. Conclusions

In this paper it is shown that it is possible to consider different water demand scenarios in the optimal design of WDN through a mean-variance model focusing on

model robustness. This provides insight into the trade-off between costs and robustness of design solutions, enabling water utilities to make well-founded choices about how much to invest in their infrastructure when it comes to being prepared for every eventuality. Moreover, it provides the decision maker with information on how to achieve a certain level of robustness in the network, i.e., the diameters that should be installed in order to meet demands under different scenarios. With this information in hands, water utilities know how to prepare their infrastructure for the future.

Different methods to generate water demand scenarios lead to different design results, in terms of costs and performances. Both approaches for generating historic and future demand scenarios have advantages and limitations. While historic demand scenarios are more substantiated, since they enable statistical analyses, they lead to design solutions less able to cope with 'unusual' changes in demand. Historic demand scenarios have thus their limitations, and this highlights the need to also consider the unknown when designing infrastructures that need to perform well on the long term (and thus, in the uncertain future). Considering future demand scenarios, although being a more subjective approach, has the advantage of including out of the ordinary analyses. In the considered case study, the design solutions obtained considering the future demand scenarios are somewhat more expensive (~10%) but are able to perform well for both historic and future scenarios. The relevant questions are then if it is worthwhile to invest a little more in the infrastructure in trade of more certainty in the performance in the future. The generation of future demand scenarios deserves more attention in future research.

From the obtained results it is also possible to conclude that applying optimization techniques to the design of a real-life WDN, leads to a significantly leaner network. In the considered case, the costs of the optimized design are only 58% of the costs of the current existing infrastructure. This design is however not able to fully satisfy demand under all future scenarios, highlighting the drawback of deterministic approaches.

To finalize, it is our belief that thoroughly quantifying water demand uncertainty and including it in optimization problems represents a step forward in the robust design of real-life WDN.

Author Contributions: Conceptualization, I.V.; Data curation, I.V.; Formal analysis, I.V.; Funding acquisition, I.V.; Investigation, I.V.; Methodology, I.V.; Software, K.v.L.; Supervision, M.d.C.C.; Validation, I.V.; Visualization, K.v.L.; Writing—original draft, I.V.; Writing—review & editing, I.V. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Joint Research Programme (BTO) for the drinking water companies, project numbers 402045/008 and 402045/089.

Data Availability Statement: Restrictions apply to the availability of these data. Data was obtained from Dutch Water Utilities and are available from the authors with the permission of the Water Utilities.

Acknowledgments: The authors would like to thank Tjakko Haaijer (Brabant Water), Maurice van de Roer (Dunea) and Anton van Eijden (WML) for their valuable insights and exciting discussions about network design and practical feasibility of theoretical approaches, and for sharing the necessary data for the case-study. The authors would also like to thank Peter van Thienen (KWR) for the constructive discussions, suggestions and guidance through the development of this work.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Snyder, L.V. Facility Location under Uncertainty: A Review. *IEE Trans.* **2006**, *38*, 547–564.
2. Rosenhead, J.; Elton, M.; Gupta, S.K. Robustness and Optimality as Criteria for Strategic Decisions. *J. Oper. Res. Soc.* **1972**, *23*, 413–431.
3. Xenochristou, M.; Blokker, M.; Vertommen. Investigating the influence of weather on water consumption: A dutch case study. In Proceedings of the 1st International WDSA/CCWI 2018 Joint Conference, Kingston, ON, Canada, 23–25 July 2018.

4. Herrera, M.; Torgo, L.; Izquierdo, J.; Pérez-García, R. Predictive models for forecasting hourly urban water demand. *J. Hydrol.* **2010**, *387*, 141–150.
5. Pieterse-Quirijns, I.; van de Roer, M. *Verbruikspatronenbibliotheek*; KWR Watercycle Research Institute: Nieuwegein, The Netherlands, 2013.
6. Agudelo-Vera, C.; Blokker, M.; Vreeburg, J.; Vogelaar, H.; Hillegers, S.; van der Hoek, J.P. Testing the Robustness of Two Water Distribution System Layouts under Changing Drinking Water Demand. *J. Water Resour. Plan. Manag.* **2016**, *142*, 05016003.
7. Bakker, M. *Optimised Control and Pipe Burst Detection by Water Demand Forecasting*; Technische Universiteit Delft: Delft, The Netherlands, 2014; p. 158.
8. Sampathirao, A.K.; Grosso, J.M.; Sopasakis, P.; Ocampo-Martinez, C.; Bemporad, A.; Puig, V. Water demand forecasting for the optimal operation of large-scale drinking water networks: The Barcelona Case Study. *IFAC Proc.* **2014**, *47*, 10457–10462.
9. Marques, J.; Cunha, M.; Savić, D. A Multicriteria Approach for a Phased Design of Water Distribution Networks. *Procedia Eng.* **2015**, *119*, 1231–1240.
10. Creaco, E.; Franchini, M.; Walski, T.M. Accounting for Phasing of Construction within the Design of Water Distribution Networks. *J. Water Resour. Plan. Manag.* **2014**, *140*, 598–606.
11. Vonk, E.; Cirkel, G.; Blokker, E.J.M. Estimating peak daily water demand under different climate change and vacation scenarios. *Water* **2019**, *11*, 1874.
12. Gargano, R.; Tricarico, C.; Marinis, G. Residential Water Demand-Daily Trends. In *Water Distribution Systems Analysis 2010*; ASCE: Reston, VA, USA, 2011.
13. Di Nardo, A.; di Natale, M.; Gargano, R.; Giudicianni, C.; Greco, R.; Santonastaso, G.F. Performance of partitioned water distribution networks under spatial-temporal variability of water demand. *Environ. Model. Softw.* **2018**, *101*, 128–136.
14. Gargano, R.; Tricarico, C.; Granata, F.; Santopietro, S.; de Marinis, G. Probabilistic Models for the Peak Residential Water Demand. *Water* **2017**, *9*, 417, doi:10.3390/w9060417.
15. Vertommen, I.; Magini, R.; Cunha, M.d.C. Scaling Water Consumption Statistics. *J. Water Resour. Plan. Manag.* **2015**, *141*, 04014072.
16. Magini, R.; Pallavicini, I.; Guercio, R. Spatial and Temporal Scaling Properties of Water Demand. *J. Water Resour. Plan. Manag.* **2008**, *134*, 276–284.
17. Ridolfi, E.; Vertommen, I.; Magini, R. Joint probabilities of demands on a water distribution network: A non-parametric approach. *AIP Conf. Proc.* **2013**, *1558*, 1681–1684.
18. Goulter, I.C. Systems Analysis in Water, Distribution Network Design: From Theory to Practice. *J. Water Resour. Plan. Manag.* **1992**, *118*, 238–248.
19. Lansey, K.E.; Duan, N.; Mays, L.W.; Tung, Y. Water distribution system design under demand uncertainty. *J. Water Resour. Plan. Manag.* **1989**, *115*, 630–645.
20. Kapelan, Z.; Babayan, A.V.; Savic, D.A.; Walters, G.A.; Khu, S.T. Two new approaches for the stochastic least cost design of water distribution systems. *Water Sci. Technol. Water Supply* **2004**, *4*, 355–363.
21. Halhal, D.; Walters, G.A.; Ouazar, D.; Savic, D.A. Water Network Rehabilitation with Structured Messy Genetic Algorithm. *J. Water Resour. Plan. Manag.* **1997**, *123*, 137–146.
22. Babayan, A.V.; Savic, D.A.; Walters, G.A. Multiobjective optimization of water distribution systems under uncertain demands and pipe roughnesses. In *Modelling and Control for Participatory Planning and Managing Water Systems*; Elsevier: Venice, Italy, 2004.
23. Babayan, A.V.; Savic, D.A.; Walters, G.A. Multiobjective optimization for the least cost design of water distribution systems under correlated uncertain parameters. In *Impacts of Global Change*; ASCE: Reston, VA, USA, 2005.
24. Kapelan, Z.S.; Savic, D.A.; Walters, G.A. Multiobjective design of water distribution systems under uncertainty. *Water Resour. Res.* **2005**, *41*, doi:10.1029/2004WR003787.
25. Giustolisi, O.; Laucelli, D.; Colombo, A.F. Deterministic versus stochastic design of water distribution networks. *J. Water Resour. Plan. Manag.* **2009**, *135*, 117–127.
26. Perelman, L.; Ostfeld, A.; Salomons, E. Cross Entropy multiobjective optimization for water distribution systems design. *Water Resour. Res.* **2008**, *44*, doi:10.1029/2007WR006248.
27. Tricarico, C.; Gargano, R.; Kapelan, Z.; Savic, D.; de Marinis, G. Economic level of reliability for the rehabilitation of hydraulic networks. *Civ. Eng. Environ. Syst.* **2006**, *23*, doi:10.1080/10286600600789383.
28. Xu, C.; Goulter, I.C. Probabilistic Model for Water Distribution Reliability. *J. Water Resour. Plan. Manag.* **1998**, *124*, 218–228.
29. Jung, D.; Kang, D.; Lansey, K.E. Water Distribution Network Design Using a Robustness Index. In *World Environmental and Water Resources Congress 2012*; ASCE: Reston, VA, USA, 2012.
30. Xu, C.C.; Goulter, I.C. Reliability-based optimal design of water distribution networks. *J. Water Resour. Plan. Manag.* **1999**, *125*, 352–362.
31. Tolson, B.A.; Maier, H.R.; Simpson, A.R.; Lence, B.J. Genetic Algorithms for reliability-based optimization of water distribution systems. *J. Water Resour. Plan. Manag.* **2004**, *130*, doi:10.1061/(ASCE)0733-9496(2004)130:1(63).
32. Sun, S.; Khu, S.-T.; Kapelan, Z.; Djordjević, S. A fast approach for multiobjective design of water distribution networks under demand uncertainty. *J. Hydroinformatics* **2011**, *13*, 143–152.
33. Sela, L.; Housh, M.; Oliker, N.; Ostfeld, A. Non-probabilistic approach for the optimal design of water distribution systems under demand uncertainty. In *14th Water Distribution Systems Analysis Conference 2012, WDSA 2012*; Engineers Australia: Barton, Australia, 2012.

34. Perelman, L.; Housh, M.; Ostfeld, A. Robust optimization for water distribution systems least cost design. *Water Resour. Res.* **2013**, *49*, 6795–6809.
35. Perelman, L.; Housh, M.; Ostfeld, A. Explicit Demand Uncertainty Formulation for Robust Design of Water Distribution Systems. In *World Environmental and Water Resources Congress 2013*; ASCE: Reston, VA, USA, 2013.
36. Schwartz, R.; Housh, M.; Ostfeld, A. Least-Cost Robust Design Optimization of Water Distribution Systems under Multiple Loading. *J. Water Resour. Plan. Manag.* **2016**, *142*, 04016031.
37. Todini, E. Looped water distribution networks design using a resilience index based heuristic approach. *Urban Water J.* **2000**, *2*, 115–122.
38. Creaco, E.; Fortunato, A.; Franchini, M.; Mazzola, M.R. Water distribution network robust design based on energy surplus index maximization. *Water Sci. Technol. Water Supply* **2015**, *15*, 1253–1258.
39. Farmani, R.; Walters, G.A.; Savic, D.A. Trade-off between total cost and reliability for Anytown water distribution network. *J. Water Resour. Plan. Manag.* **2005**, *131*, 161–171.
40. Creaco, E.; Franchini, M. Fast network multi-objective design algorithm combined with an a posteriori procedure for reliability evaluation under various operational scenarios. *Urban Water J.* **2012**, *9*, 385–399.
41. Creaco, E.; Fortunato, A.; Franchini, M.; Mazzola, M.R. Comparison between Entropy and Resilience as Indirect Measures of Reliability in the Framework of Water Distribution Network Design. *Procedia Eng.* **2014**, *70*, 379–388.
42. Prasad, T.D.; Hong, S.-H.; Park, N. Reliability based design of water distribution networks using multi-objective genetic algorithms. *KSCE J. Civ. Eng.* **2003**, *7*, 351–361.
43. Tanyimboh, T.T.; Templeman, A.B. A quantified assessment of the relationship between the reliability and entropy of water distribution systems. *Eng. Optim.* **2000**, *33*, 179–199.
44. Fu, G.; Kapelan, Z. Fuzzy probabilistic design of water distribution networks. *Water Resour. Res.* **2011**, *47*, doi:10.1029/2010WR009739.
45. Cunha, M.; Sousa, J. Robust design of water distribution networks for a proactive risk management. *J. Water Resour. Plan. Manag.* **2010**, *136*, 227–236.
46. Cunha, M.C.; Sousa, J.J.O. Robust design of water distribution networks: A comparison of two different approaches. In *Integrating Water Systems—Proceedings of the 10th International on Computing and Control for the Water Industry, CCWI 2009*; CRC Press: Boca Raton, FL, USA, 2010.
47. Mulvey, J.M.; Vanderbey, R.J.; Zenios, S.A. Robust optimization of large scale systems. *Oper. Res.* **1995**, *43*, 264–281.
48. Daskin, M.S.; Hesse, S.M.; Reville, C.S. α -Reliable p-minimax regret: A new model for strategic facility location modeling. *Locat. Sci.* **1997**, *5*, 227–246.
49. Snyder, L.V.; Daskin, M.S. Stochastic p-robust location problems. *IIE Trans.* **2006**, *38*, 971–985.
50. Chen, G.; Daskin, M.S.; Shen, Z.M.; Uryasev, S. The α -reliable mean-excess regret model for stochastic facility location modeling. *Nav. Res. Logist.* **2006**, *53*, 617–626.
51. Zeferino, J.A.; Antunes, A.; Cunha, M.C. Regional Wastewater System Planning under Population Dynamics Uncertainty. *J. Water Resour. Plan. Manag.* **2014**, *140*, 322–331.
52. Basupi, I.; Kapelan, Z. Flexible Water Distribution System Design under Future Demand Uncertainty. *J. Water Resour. Plan. Manag.* **2015**, *141*, 04014067.
53. Basupi, I.; Kapelan, Z. Evaluating Flexibility in Water Distribution System Design under Future Demand Uncertainty. *J. Infrastruct. Syst.* **2015**, *21*, 04014034.
54. Huang, D.; Vairavamoorthy, K.; Tsegaye, S. Flexible design of urban water distribution networks. In *World Environmental and Water Resources Congress 2010: Challenges of Change*; ASCE: Providence, RI, USA, 2010; pp. 4225–4236.
55. Marques, J.; Cunha, M.; Savić, D.A. Multi-objective optimization of water distribution systems based on a real options approach. *Environ. Model. Softw.* **2015**, *63*, 1–13.
56. Marques, J.; Cunha, M.; Savić, D.; Giustolisi, O. Water Network Design Using a Multiobjective Real Options Framework. *J. Optim.* **2017**, *2017*, 13.
57. Kang, D.; Lansey, K.E. Scenario-Based Multistage Construction of Water Supply Infrastructure. In *World Environmental and Water Resources Congress 2012*; ASCE: Reston, VA, USA, 2012.
58. Marques, J.; Cunha, M.; Savić, D.A. Using real options for an eco-friendly design of water distribution systems. *J. Hydroinformatics* **2015**, *17*, 20–35.
59. Brown, C. The End of Reliability. *J. Water Resour. Plan. Manag.* **2010**, *136*, 143–145.
60. Vanston, J.H.; Frisbie, W.P.; Lopreato, S.C.; Boston, D.L., Jr. Alternate scenario planning. *Technol. Forecast. Soc. Chang.* **1977**, *10*, 159–180.
61. Jung, D.; Lee, S.; Kim, J.H. Robustness and Water Distribution System: State-of-the-Art Review. *Water* **2019**, *11*, 974.
62. Laguna, M. Applying Robust Optimization to Capacity Expansion of One Location in Telecommunications with Demand Uncertainty. *Manag. Sci.* **1998**, *44*, S101–S110.
63. Samsatli, N.J.; Papageorgiou, L.G.; Shah, N. Robustness metrics for dynamic optimization models under parameter uncertainty. *Process Syst. Eng.* **1998**, *44*, 1993–2006.
64. Afonso, P.M.; Cunha, M.d.C. Robust Optimal Design of Activated Sludge Bioreactors. *J. Environ. Eng.* **2007**, *133*, 44–52.
65. Watkins, D.W.; McKinney, D.C. Finding Robust Solutions to Water Resources Problems. *J. Water Resour. Plan. Manag.* **1997**, *123*, 49–58.

66. Carr, R.D.; Greenberg, H.J.; Hart, W.E.; Konjevod, G.; Lauer, E.; Lin, H.; Morrison, T.; Phillips, C.A. Robust optimization of contaminant sensor placement for community water systems. *Math. Program.* **2006**, *107*, 337–356.
67. Marques, J.; Cunha, M.C.; Sousa, J.; Savić, D. Robust optimization methodologies for water supply systems design. *Drink. Water Eng. Sci.* **2012**, *5*, 31–37.
68. Gutierrez, G.J.; Kouvelis, P. A robustness approach to international sourcing. *Ann. Oper. Res.* **1995**, *59*, 165–193.
69. Beh, E.H.Y.; Zheng, F.; Dandy, G.C.; Maiera, H.R.; Kapelan, Z. Robust optimization of water infrastructure planning under deep uncertainty using metamodels. *Environ. Model. Softw.* **2017**, *93*, 92–105.
70. Mala-Jetmarova, H.; Sultanova, N.; Savic, D. Lost in Optimisation of Water Distribution Systems? A Literature Review of System Design. *Water* **2018**, *10*, 307.
71. Yamashita, D.S.; Armentano, V.A.; Laguna, M. Robust optimization models for project scheduling with resource availability cost. *J. Sched.* **2007**, *10*, 67–76.
72. Wagner, J.M.; Shamir, U.; Marks, D.H. Water Distribution Reliability: Simulation Methods. *J. Water Resour. Plan. Manag.* **1988**, *114*, 276–294.
73. Maier, H.R.; Guillaume, J.H.A.; van Delden, H.; Riddell, G.A.; Haasnoot, M.; Kwakkel, J.H. An uncertain future, deep uncertainty, scenarios, robustness and adaptation: How do they fit together? *Environ. Model. Softw.* **2016**, *81*, 154–164.
74. Blokker, E.J.M.; Vreeburg, J.H.G. Monte Carlo Simulation of Residential Water Demand: A Stochastic End-Use Model. In *Impacts of Global Climate Change*; ASCE: Reston, VA, USA, 2005.
75. Van Thienen, P.; Vertommen, I. Gondwana: A Generic Optimization Tool for Drinking Water Distribution Systems Design and Operation. *Procedia Eng.* **2015**, *119* (Suppl. C), 1212–1220.
76. Van Laarhoven, K.; Vertommen, I.; van Thienen, P. Technical note: Problem specific variators in a genetic algorithm for the optimization of drinking water networks. *Drink. Water Eng. Sci. Discuss* **2018**, doi:10.5194/dwes-11-101-2018.
77. Rossman, L.A.; Woo, H.; Tryby, M.; Shang, F.; Janke, R.; Haxton, T. *EPANET 2.2 Users Manual*; Environmental Protection Agency: Washington, DC, USA, 2020.
78. Garrett, A. Inspyred 1.0 Documentation. 2012, Python Software Foundation. Available online: <https://aarongarrett.github.io/inspyred> (accessed on 1 January 2015).
79. van Thienen, P.; et al. Nature-based optimization techniques and their application in the water industry. *HydroLink* **2021**, to be published.