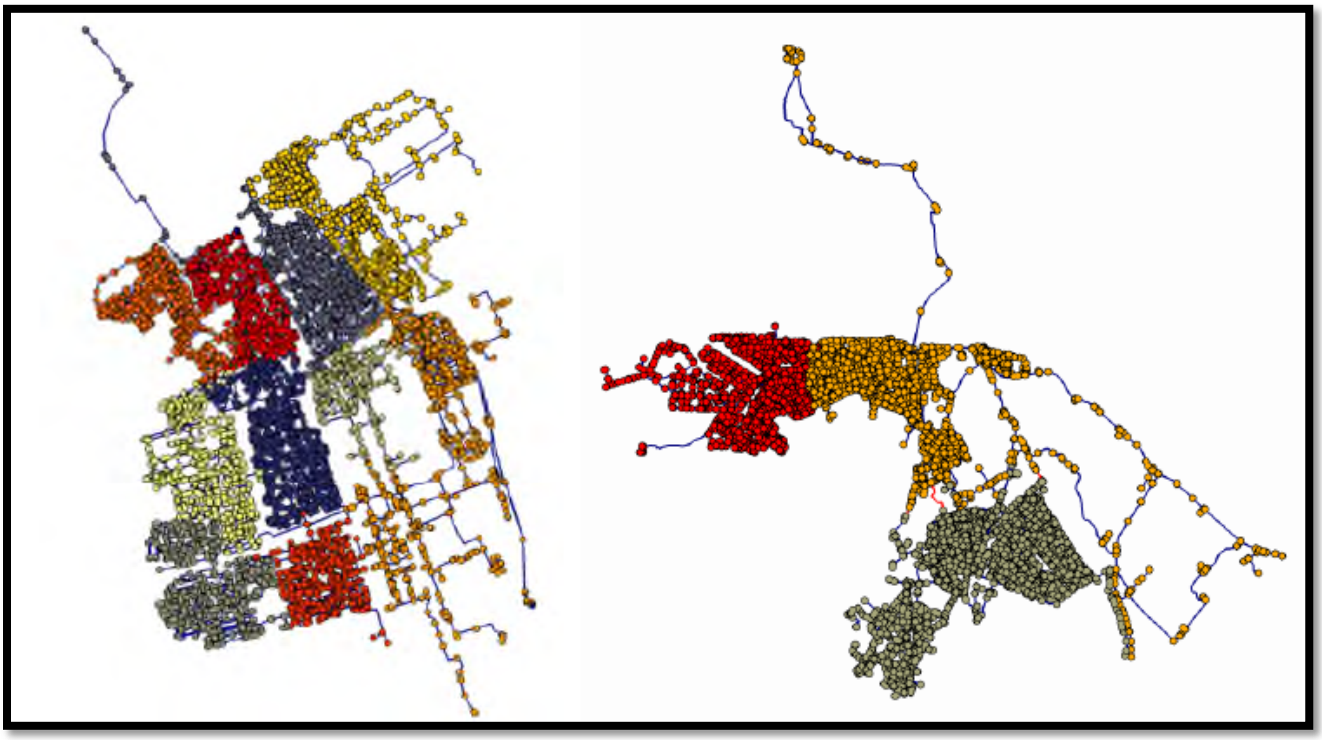
Analysis and classification of water distribution networks

A case study Den Haag

Looking for a proper characterization of the WDN to estimate DMAs design

Laura Viviana Garzon Useche

MSc Thesis Identifier 1071564
April 202



Analysis and classification of water distribution networks

Looking for a proper characterization of the WDN to estimate DMAs design

Master of Science Thesis
by

Laura Viviana Garzon Useche

Supervisors
Professor Dimitri P. Solomatine

Mentors
Dr. Gerald A. Corzo
Dra. Mirjam Blokker
Dr. Leonardo A. Segura

Examination Committee
Professor Dragan Savic
Dr. German R. Santos

This research is done for the partial fulfilment of requirements for the Master of Science degree at the IHE Delft Institute for Water Education, Delft, the Netherlands.

Delft
19/04/2022

Abstract

For decades, establishing and maintaining proper management in a WDS water distribution system has been a challenge for water professionals. In response, concepts such as the District Metering Area (DMA) have emerged. Its application has been one of the best methods to improve the performance of water distribution systems. However, there are still significant limitations to implement it in actual cases. The aim of this research is to associate more than 600 model networks with more than 35 characteristics to find relationships between the main characteristics and the DMA designs. The proposed approach consists of 4 steps—first, the creation of 600 model networks through EPANET and the optimisation tool, GONDWANA. Second, the development of a comprehensive characterisation of the network in terms of hydraulic, water quality, water safety, economic and topological metrics, then these are extracted from the previously created model networks. Third, the main components and the performance of characteristics patterns are analysed and identified. Fourth, DMA designs from 60 networks taken from the step 1 by using GODWANA. Fifth, integration of the step 3 with the step fourth to find the relation between the characteristics and the DMAs design.

Table of Contents

List of Figures

List of Tables

Abbreviations

WDS = Water distribution system
NM= Network model
WNTR= Water network tool for resilience
SK= Skeletonization
LD= Link density
CH= Characterisation
PCA = Principal component analysis
DMA= District metered areas

Chapter 1 Introduction

1.1 Background

A water distribution system (WDS) is a complex set of hydraulic controlled elements connected to transport quantities of water from sources to consumers (Ostfeld, 2015).These systems are typically subject to design and operation that involves optimisation, which can be very complex due to the high number of decision variables, constraints and water planning issues, which makes it difficult to understand and analyse.
(Yazdani & Jeffrey, 2011)(Yazdani & Jeffrey, 2011)(Yazdani & Jeffrey, 2011)(Yazdani & Jeffrey, 2011)The high number of dimensions of the WDS makes it challenging to identify "optimal or near-optimal" solutions. In addition, different sizes of the WDS increase or decrease the decision space and decision variables (Fu et al., 2012).Optimal design of large water distribution networks is a problem that involves making decisions about the layout and size of pipes (length and diameter) while trying to minimise the cost of network design, construction and operation (Yazdani & Jeffrey, 2011).Water utilities must create plans to design and operate their WDS. However, these plans are challenging because they involve multiple uncertainties, such as future consumption, current and future system characteristics, and unknown variables such as the current state of the valves, network configuration, pipe roughness and others (Kapelan et al., 2005).
(Awe et al., 2019; Jayaram & Srinivasan, 2008; Jolly et al., 2014; Sg, 2016)(Awe et al., 2019; Jayaram & Srinivasan, 2008; Jolly et al., 2014; Sg, 2016)(Awe et al., 2019; Jayaram & Srinivasan, 2008; Jolly et al., 2014; Sg, 2016)(Awe et al., 2019; Jayaram & Srinivasan, 2008; Jolly et al., 2014; Sg, 2016)In prior research (Awe et al., 2019; Jayaram & Srinivasan, 2008; Jolly et al., 2014; Sg, 2016)proposed an important number of possible solutions to address with these types of little understandable scenarios. This due to each scenario stems from different characteristics and therefore different operation, constraints among others. Literature shows that most of the proposes are based on the topology and hydraulic variables (Zarghami et al., 2019) in which are related with three main domains. First, pipe characteristics as the robustness, length, etc. Second, hydraulic characteristics as flow, pressure, water quality, etc. Third, topological characteristics as connectivity, robustness, centrality, etc. This points to a better insight into the behaviour of hydraulic parameters and therefore to the development of water plans with less uncertainty.
(Vargas & Saldarriaga, 2019)(Vargas & Saldarriaga, 2019)(Price & Ostfeld, 2015)(Price & Ostfeld, 2015; Tzatchkov et al., 2007)(Bui et al., 2020, 2021; de Paola et al., 2014; Tsitsifli & Kanakoudis, 2021)(Alvisi & Franchini, 2014)(Price & Ostfeld, 2015)(Price & Ostfeld, 2015; Tzatchkov et al., 2007)(Bui et al., 2020, 2021; de Paola et al., 2014; Tsitsifli & Kanakoudis, 2021)(Alvisi & Franchini, 2014)(Vargas & Saldarriaga, 2019)(Vargas & Saldarriaga, 2019)(Price & Ostfeld, 2015)(Price & Ostfeld, 2015)(Price & Ostfeld, 2015; Tzatchkov et al., 2007)(Bui et al., 2020, 2021; de Paola et al., 2014; Tsitsifli & Kanakoudis, 2021)(Alvisi & Franchini, 2014)(Price & Ostfeld, 2015; Tzatchkov et al., 2007)(Bui et al., 2020, 2021; de Paola et al., 2014; Tsitsifli & Kanakoudis, 2021)(Alvisi & Franchini, 2014)In addition to the aforementioned domains, overtime has emerged several concepts to make water systems

increasingly sustainable. One of them is the concept of District Metered Area DMA in which aims to split a network in small areas delimited by hydraulic fixtures in order to get a better monitoring (Alvisi & Franchini, 2014). DMAs have a wide applicability and also different methods to establish a DMA according a given problem (Bui et al., 2020, 2021; de Paola et al., 2014; Tsitsifli & Kanakoudis, 2021). Whitin its applicability is the water quality management, reduction of the non-revenue, water management simplification, leakage identification, among others (Bui et al., 2020). Also, a high variety of methods based on Graph, theory(Price & Ostfeld, 2015; Tzatchkov et al., 2007) , Genetic algorithms (Gupta et al., 1999; Savic· et al., 1999; Wu et al., 2004), water leakage management (Price & Ostfeld, 2015), Multi-agent Metaphor (Mikulecky, 2011), Box covering algorithm (Vargas & Saldarriaga, 2019), and others.

There two aspects identified in the above research. First, most of the solutions are rely on a mid-term solution and not a long-term. Second, regardless the method applied, the core of each study is based on the characteristics of the network. Third, the characterization is limited to analysing either hydraulic properties, topology properties or pipes properties, but not all of them.

A universal methodology to perform DMAs under any scenario with long-term vision is still not well understood. A better understanding of the impact of every characteristic under diverse scenarios is needed to perform DMA's for now and the future.

This study aims to propose a method to identify the network characteristics that can design a proper DMA design. This DMA design is performed rely on the characterized networks to classify which scenario(s) (or combination of characteristics) is related to a DMA design. The same way the characteristics describe the behaviour of the assessed network in different scenarios and contribute to a better understanding of possible variables influencing in the design of optimal DMAs. This study does not solve all the problems of the WDS, but contributes to the characterization and design of the optimal area for the networks. In this sense, network characterization requires examining large urban datasets to compare and identify unexpected similarities and differences between urban water systems (Diao et al., 2014).

1.2 Problem statement

Countless methodologies for characterizing WDNs are increasingly being applied. Along their development each methodology shows an important number of characteristics in the WDS analysis, however, most of them have not been tested in complex networks where actual WDS mainly occur. Thereby, a universal definition of how it should characterize WDNs is still not established. As a consequence, there is not a well-known set of main characteristics that depict the influence of a DMA design under any scenario.

This research concentrates specifically on the identification of the most important characteristics that may influence the problem of District Metered Areas (DMA) design, a subdivision of a network into smaller, optimal areas to help assessing, monitoring and analyzing them. Additionally, it is required to characterize the current system to try to predict the future characteristics of the system (Basupi & Kapelan, 2015). These analyses can lead to changes in the current systems to generate sustainable designs in the future.

1.3 Research questions

The main research question is formulated as follows:

1. How can a proper characterization of a WDN help successful DMA design solutions?

Followed by the specific research questions:

2. Which characterization methods are available in the literature that are feasible for this study

3. How can it be generated multiple model networks defining hydraulic limitation from the demands and supply?
4. How does the current optimization practices conceptualise DMA design?
5.

1.4 Research objectives

The above research questions lead to the following objectives:
Main Objective

1. To propose a method to characterize the water distribution networks and test it on a case study to determine a proper characterization of successful DMA design.

Specific objectives

2. To select the methods from the literature that can be used for developing a proper characterization of the network.
3. To explore how to use Gondwana and Epanet to produce multiples network models.
4. To design DMAs based on current optimization practices by using Gondwana tool.
5. To identify the set of characteristics and their relation with the new model networks.

1.5 Innovation

The lack of a firm and clear methodology that defines a universal DMA design leads to address five points; First, integration of 38 characteristics based on the analysis of multiples characterization method. Second, characteristics testing in more than 600 network models to identify typical values and establishment ranges. Third, exploration and applicability of optimization tools as Gondwana to create DMAs. Fourth, examination and implementation of techniques for data management in order to reduce the dimensionality of the characteristics and to extract valuable insights. As a result, a novel methodology that allows the integration of great part of the characterization methods, based-optimization testing, and unification of multiples data sources into versatile one.

1.6 Practical value

This study has value for both water companies and researchers:
Regarding water companies, are currently several methodologies to evaluate and monitor the WDNs. However, the cost of installing monitoring systems, the variability in network characteristics and the completeness of network delimitation still shows a high uncertainty. By evaluating the influence of one or a set of characteristics with respect to a given model network, which could determine a optimal DMA under any scenario. In other words, the identification of the main characteristics, creation of scenarios to change the model network to provide better "optimal" value on the characteristics and plan different changes to the existing network to provide better and stable DMAs.
That is because this research could be a starting point to evaluate the range value of the characteristics and how they determine the optimal DMAs design.
Regarding research, this study could be a reference to provide a efficient characterization of the networks in terms of time and acquired data. This characterization will be possible by using a python script and create a json or excel table with the results of the integration of a number of characteristics. besides, the analysis characteristics of the 600 networks in a box whisker plot to cooperate the actual range values with the new network.
A dataset of 600 model networks generated on this research will be left open for future investigations.

Chapter 2 Literature review

Due to the extensive research about the context of this study, this chapter is summarized in three main sections, characterization, District Metered Areas (DMAs) and data analysis. The first section gives rise to brief description of the main or most used characteristics in a water systems problem. The second section focuses on DMAs method and methodologies. The third section shows how current technologies and new learning methods can address this type of studies in which implies high dimensionality, large dataset handling, and optimization of processes. As a result, the theory is analysed to establish the basis and criterion to carry out this research.

Characteristics

Exists countless criteria to categorize the main characteristics or metrics in water distribution systems (Zarghami, Gunawan, and Schultmann, 2019). So that this subsection classifies them in the categorization most common, network characteristics, topological metrics, hydraulic metrics, water quality metrics and economic metrics (Chang et al., 2021; Diao et al., 2014; Kara et al., 2016; Mohammed et al., 2021; Ratnayaka et al., 2009; Shin et al., 2018; Suribabu, 2017; Zhang et al., 2021) .

Network: These give an overview of the composition of the system, which includes hydraulic fixtures as intrinsic properties of the pipes, maximum and minimum elevations of the hydraulic control elements, number of components by pipe characteristic, quantify and list of patterns, among others.

Topological: These describes the physical components of the network. Graph theory is the basis to analyse the topological metric. One of the characteristics that composes this subsection are link density, node degree, eccentricity, closeness centrality, among others.

Hydraulic: These required the simulation of the model network since the flow, pressure and demand are the basis for the analysis of this metric. This include high, low and average demands and pressure per node, resident time, entropy, and others.

Water quality: These focus on the water age along the system and hydraulic control elements such as tanks, nodes, dead-end pipeline, and others.

Economic: These are evaluated with the results of the model network under certain scenario to determine the maintenance, operation and repairment cost in the assets. These metric compromises pump energy, greenhouse gas emission and network cost.

The aforementioned metrics have been evaluated by separated (Set de characteristics) to find optimal solutions to address a given scenario. Below some studies that implement these metrics to cope with the challenges of water systems:

1. (Meng et al., 2018; Perelman & Ostfeld, 2005; Zarghami et al., 2019)(Meng et al., 2018; Perelman & Ostfeld, 2005; Zarghami et al., 2019)(Meng et al., 2018; Perelman & Ostfeld, 2005; Zarghami et al., 2019)(Meng et al., 2018; Perelman & Ostfeld, 2005; Zarghami et al., 2019)Topological and hydraulic variables to understand and analyse parameters such connectivity, robustness, resilience, entropy(Meng et al., 2018; Perelman & Ostfeld, 2005; Zarghami et al., 2019).

2. Analysis of centrality metrics as a factor for identifying key physical domain characteristics; these allow quantifying the importance of vertices or edges (Giustolisi et al., 2019).
3. The behaviour of the water system can be described in three main parameters; first, is for a physical law; describe the flow and their relationships with the pipes and hydraulic control elements; second, with a consumer demand; third, with the system layout (Ostfeld et al., 2002). This criterion compromises topological and hydraulic characteristics, which has generated breakthroughs for a better understanding to the network's behaviour.
4. Determination of the influence of residence time on water quality by means of hydraulic metrics (Shamsaei et al., 2013)
5. Evaluation of the reliability of the system through topological and hydraulic metrics (Ostfeld et al., 2002)
6. The degree of node distribution may determine redundancy in system connectivity (Ulusoy et al., 2018)
7. How the role of the topology can be represented along with graph theory to characterize the performance of the network in terms of water quality and hydraulic metrics(Torres et al., 2016) .
8. Resilience measurement by mean of the node degree as topological indicator (Newman, 2010).
9. Evaluation of robustness and entropy is quantified by integrating topological and hydraulic metrics. Both are integrated assess how the impact on disruptive events can be mitigated (Zarghami et al., 2019).
10. (Alvisi & Franchini, 2010; Diao et al., 2022; van Thienen & Vertommen, 2015)(Alvisi & Franchini, 2010; Diao et al., 2022; van Thienen & Vertommen, 2015)(Alvisi & Franchini, 2010; Diao et al., 2022; van Thienen & Vertommen, 2015)(Alvisi & Franchini, 2010; Diao et al., 2022; van Thienen & Vertommen, 2015)Many others (Alvisi & Franchini, 2010; Diao et al., 2022; van Thienen & Vertommen, 2015)

In addition to the above solutions for each specific scenario, literature shows a fundamental role of the characteristics in the DMAs design(Armand et al., 2018; Cassidy et al., 2021; Galdiero et al., 2016; Gonelas et al., 2017; Hajebi et al., 2016; Wright et al., 2014). In which the integration of two or more metrics has even contributed to the optimization of the design(Chondronasios et al., 2017). The DMA's forming is highly variant since include different combination of set of characteristics so that there are not a better or worse design established.

District metered areas

(Thienen & Montiel, 2014).(Thienen & Montiel, 2014).(Thienen & Montiel, 2014).(Thienen & Montiel, 2014).The District Metered Area is a small part of the network with a provision to monitor the water supplied and consumed individually(DTK HYDRONET, 2019). By creating DMAs provide a great number of advantages, including isolation to manage the water contamination, reduction of the non-revenue, water pressure management simplification, and signals of existing leakage (Bui et al., 2020). Which is valuable information for identifying anomalies in a water distribution system (Thienen & Montiel, 2014).

Applications and methods to design DMAs are numerous. So that a clear way to show its scope and use, some studies are presented as follows:

Based on graph theory: The goal of this theory is to find the smallest possible set of the vertex that cover all the edges in throw the closely related problems (Yuan & Kuo, 1998) (Bondy & Murty, 1978),
Based on genetic algorithm: This method evaluates a multivariate problem in a pro to provide an excellent solution to the optimization problem. this method has to evaluate the objectives functions is tested through MATLAB's and the result of the genetic algorithm are compared win an algorithm (c++) development in an early stage to provide the reduction pressure on the system. (Elsevier Enhanced Reader, 2016.)
Based on water leakage management: The form to design the DMA in this method is part of the DMA simulation; then, the definition of limits from the measurement points is performed. If the created DMAs are very large, the installation of gauges and virtual closures is performed. As a form of verification, the pressure variation in the delimited areas is evaluated. (Özdemir, 2018)
Based on multi-agent Metaphor: This is based on a collection and analysis of independent agents interacting through discrete events. When the agent is being defined and the relationship established is necessary to generate individual actions at a specific time. A final step consists in to observe the model and recording what is happening. It Will possess the properties of the networks (Cordeiro, 2011).
Based on box covering algorithm: This algorithm consists in to find the number of boxes; these boxes are defined with a group of unions needed to consider the entire network in various lengths. This includes the use of hydraulic variables such as the hydraulic gradient line. (Kevin Vargas, Camilo Salcedo, Juan Saldarriaga, 2019.)
Some of these methods have been applied of the different cases of study around the work; however, none of them has generated a direct relationship with specific network characteristics.
In the Netherlands:
For this case of study, the design of the DMAs will be created from the Gondwana platform. Gondwana is a software tool that has been developed to make changes in the design of DMAs. The definition of the objectives influences the efficiency of the changed design. Each time the tool changes, the interactive process generates an optimal performance design of the DMAs. (Van Laarhoven K, Ina Vertommen, van Thienen P, Henk de Kater, Schaap Peter, Gardien Dennis, Van den Boom Michael and Ralf de Groot5 , 2020). As a result, the platform gives a score of the best DMAs created. This score will depend on the methods of DMAs applied for Gondwana and fixed in each network.
The classification of the network will be a development from the characterization and the DMAs design. The idea is to create the relationship between the last two steps to create a better understanding of the networks.

Chapter 3 A case study

The area of interest is well known as Den Haag (The Hague), located in the Netherlands. A small city with a population of over half a million (Approximately 550.000 inhabitants) in which is considered the third-largest urban area in the Netherlands (citypopulation, 2022; statista, 2022). It is the capital of the province of South Holland. Den Haag is divided into eight main districts composed by multiples neighbourhoods.

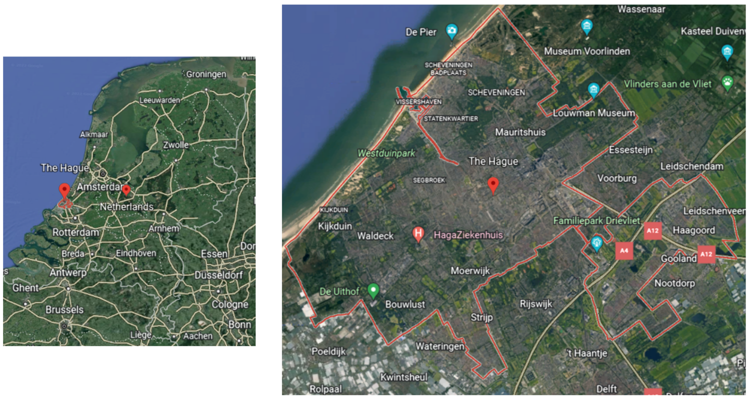


Figure 1 Geographic location of the municipality of Den Haag

This city is supplied by the water company Dunea (Dunea Waterbedrijf) in terms of drinking water. Its water distribution network has more than 750 kms of length with a number of nodes 19.000 approximately. The number of connections is about 580.000 with a demand of about 105 cubic meters per hour. Unlike most of the cities around of world, Den Haag has great part under sea level and therefore is a flat area. In this sense, its water distribution system is operated as pumping system. The minimum elevation of the network is about -7 meters under sea level and a maximum level of 15 meters approximately. Below a graphical representation of Den Haag:

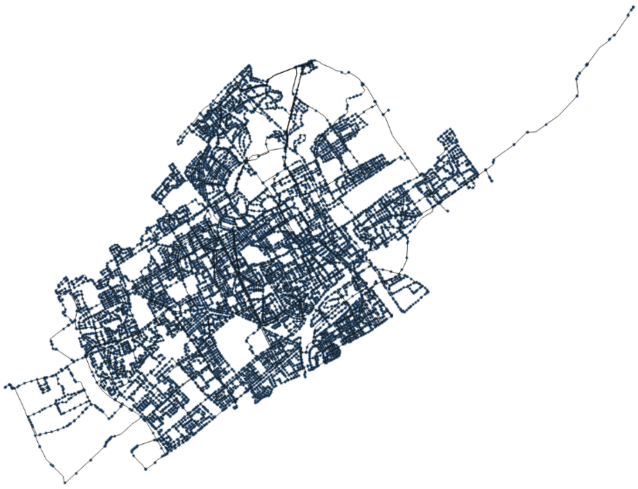


Figure 2 Representation of pipelines and nodes of the water distribution network of Den Haag

Chapter 4 Research Methodology

4.1 Preparation of the case studies

(*Gondwana - KWR*, n.d.)(*Gondwana - KWR*, n.d.)(*Gondwana - KWR*, n.d.)(*Gondwana - KWR*, n.d.)Den Haag network was taken as a base network. Building on this network was generated multiple new networks model (NM). This step was possible by using Gondwana Tool and Epanet. Gondwana is a optimization platform created by KWR (*Gondwana - KWR*, n.d.) to bring the gap between the academic research and practical implementability. It platform allows to redesigning the networks for different purposes. Using Gondwana was to modify and generate many networks with different variables simultaneously. It should be noted that the changes made to the networks consider the good hydraulic performance of the networks as a constraint. This good performance is evaluated with the ability to redesign the existing network without generating negative pressures and supply the demand in the system.

The variables considered in the changes are described as follows:

1. The diameter of all pipes
2. The base demand of the nodes
3. The elevation
4. The length of all pipes
5. Skeletonize the network model

Variables such as elevation and demand were modified by using Epanet environment. This modification was due to the time required, which included an optimization process of networks. With all the variables was possible to make different scenarios where they were combinate to obtain the new NM.

The values to get as an input to change the network were based on:

1. Diameter = select the medium and biggest values of the diameters on the existing network and give them input to the new networks (diameters in mm 0, 150, 400, 500, 800, 1200, 1600).
2. Demand = increase in the base demand of the total network from 5 to 30 %
3. Elevation = increase the elevation in some nodes from 10m to 25m
4. Length = select the medium and biggest values of the length on the existing network and give as an input of the new networks(length in m 30, 50, 100, 150, 180, 200, 300, 500, 1000).

Gondwana NM generation

The procedure to change the variable to generate the new NM from Gondwana is described as follows:
Input

a. **Hydraulic model=** simulation model of Den Haag.
b. **Datasets=** the two datasets included with the generation were:

Diameters= The size of the diameters allowed to change in the input network.

Diameter(mm)	0.00001	150	400	500	800	1200	1600

Table 1:List of input diameters

Length = The length of the pipes allowed to change in the input network.

Length (m)	30	50	100	150	180	200	300	500	1000

Table 2:List of input lengths

c. **Decision variables=** There are two decision variables applied simultaneously. Those had been related with the input values (the length and diameter in the pipes) allowed to add a dataset and change all the pipes of the model network system.

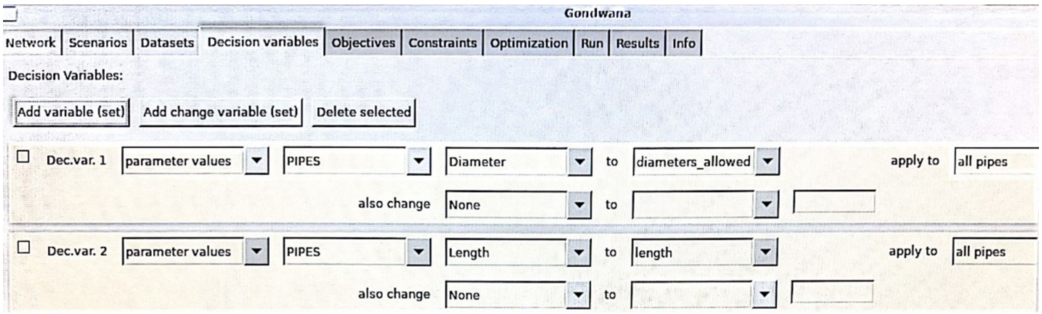


Figure 3: Gondwana, decision variables to generate new NM

Objective function=found the Minimum size of the pipes. This size is defined as the length per diameter of each pipe. this size is related to the pipe wall surface, not to pipe volume. This is a choice, and you should explain this.

$$Arg \min f(p) = p(L\ x\ D)$$

Equation 1: Objective to generate new NMs

Where: P= pipes, L= length, D= Diameter

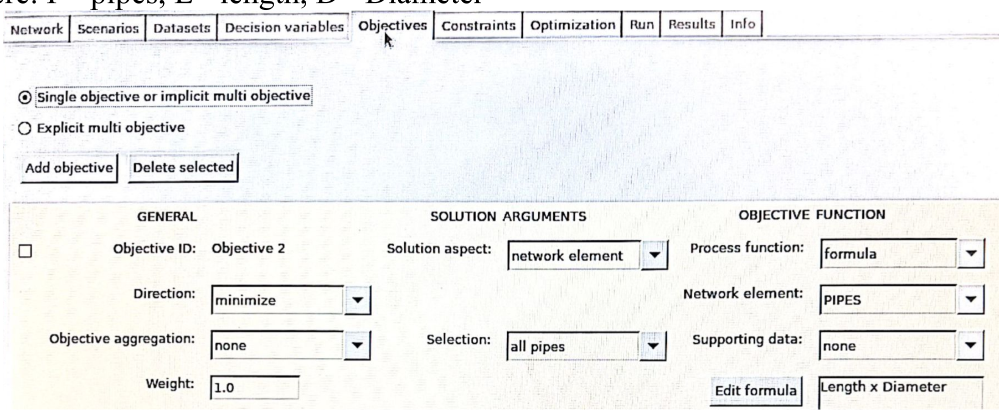


Figure 4:Gondwana,objective function to generate new NMs

d. **Constraints=** The constraints had been defined as a minimum value of pressure in all the nodes. This value was 6 mca.

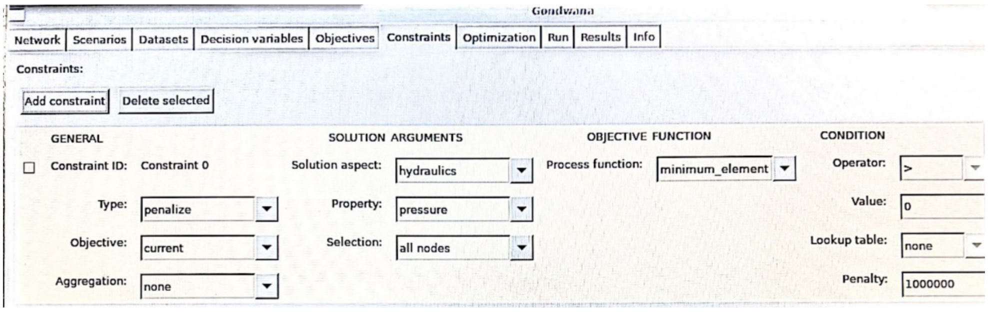


Figure 5:Gondwana, constrains to generate new NMs Change the grapth

e. **Optimization process=** In this case, the genetic algorithm created a new NM with a good hydraulic behavior. This hydraulic behavior is giving by the capacity to supply the total demand of the system. The number of generations assigned corresponds to the number of new models created based on the objective function and the constraints.

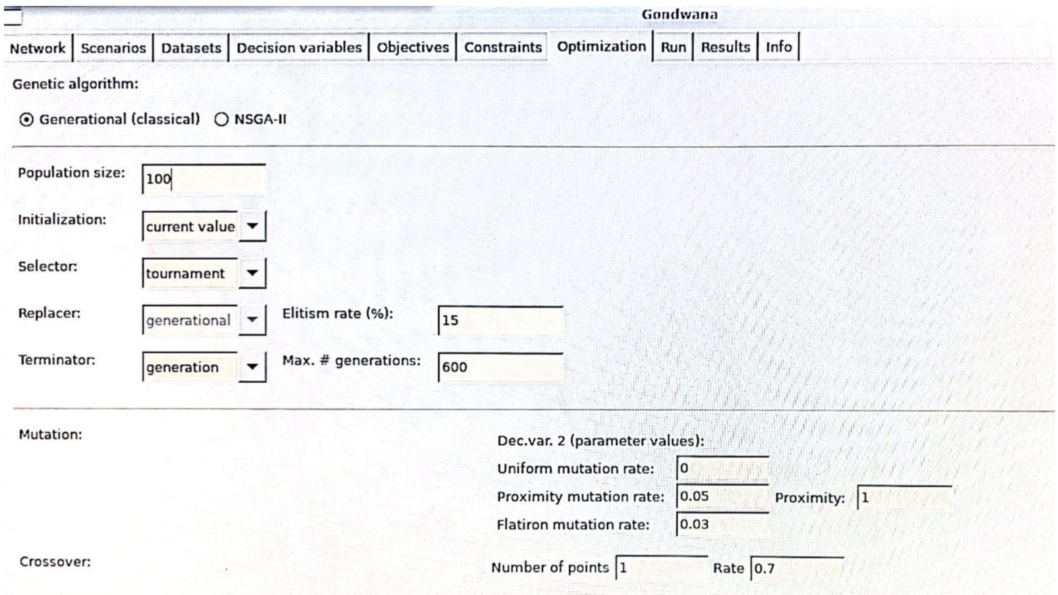


Figure 6:Gondwana, definition of the optimization process Change the grapth

Output: More than 600 NM were generated in this step.

Python and Epanet NM generation

Twenty NM of the dataset generated by Gondwana was modified by using a python script to change the base demand and the elevation of the node. The value of the base demand changed from 5% to 10% in 5 NM, from 10% to 20 % in 10 NM and from 20% to 30% at the last 15 NM.

Skeletonization of the NM

The skeletonisation (Sk) of the NM was applied to all the 600 networks generated by Gondwana and modified in Epanet. The main objective of this step was to create morphology changes on the NMs without a hydraulic impact.

The Sk had made with a python script and the use of the library Water Network Tool for Resilience (WNTR). WNTR is a Python package. It is designed to analyse impacts in the WDS, simulate disaster scenarios and analyse. It can help explore systems' capabilities to find solutions to disasters while generating planning for building resilient systems over time(Water Network Tool for Resilience (WNTR) User Manual: Version 0.2.3 | Science Inventory | US EPA, 2019)

The Sk of the networks model is based on three steps. First Branch trimming: this step receives as an input value the NM and a threshold of the size of the diameters that you want to remove. Based on that threshold, all undersized pipes are removed. The demand related to the junctions connected for the pipes removed is moving to the closer nodes. Second, the Series pipe merges: the condition to merge the pipes is that both pipes have to be in series and if they are below the pipe diameter threshold, the demand is moved to the nearest node on the model. The last step is parallel pipe merge: if both pipes in parallel are below the pipe diameter threshold, they are merged.

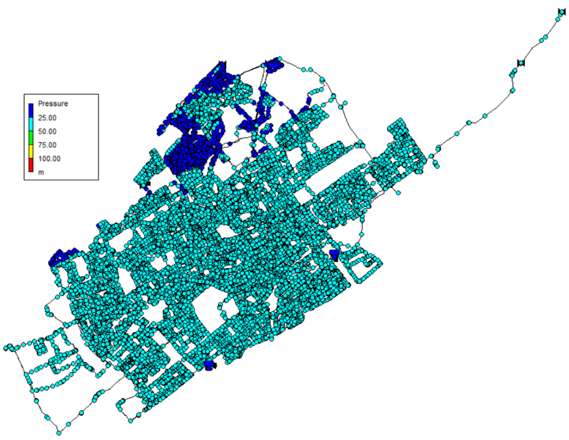


Figure 7:Den Haag model network

Figure 8: Den Haag model network skeletonized

	Nodes	Pipes
Den Haag	18997	21029
Den Haag Sk	4550	6505

The figure 5 and 6 shows the Den Haag network and the Den Haag skeletonized. For this research the threshold value to Sk the NM used was 304 mm or 12 in.

4.2 Characterization development.

Selection of the characteristics

The selection of the characteristics (CH) to analyse in the water distribution system was based on the possible components considered in a hydraulic network analysis of the epanet program. That means that the components to contemplate in this research are related to the network model. In most cases, this model has been built to simulate the water distribution network and evaluate the behaviour (modelling of the water system, it can be measured with variables as water resident time and water age among others)and performance (The hydraulic evaluation of the networks determines the performance of a network as a function of the flow velocity in the pipes and the pressure at the nodes (Hajibabaei et al., 2019)) in multiple scenarios (Priyanka Jawale et al., 2018).

The principal components to consider in this research describe variables as hydraulic modelling (Udhane et al., 2018)water security and water quality (Clark, 1992,Johnson et al., 2008).Furthermore, attributes such as topological analysis and evaluation of the network's economic cost are analysed using graph theory (Torres et al., 2016).These attributes require a hydraulic simulation.

The CH considered in this research are as follows:

Network characteristics:

These characteristics describe the general composition of the network. The 14 CH consider are:

Minimal elevation, maximum elevation, number of nodes, number of nodes with demand below or equal to 0, total demand, number of pipes, number of tanks, number of reservoirs, number of pumps, number of valves, number of patterns, list of patterns, length in relation with roughness, total length, average length,

Network Topological metrics:

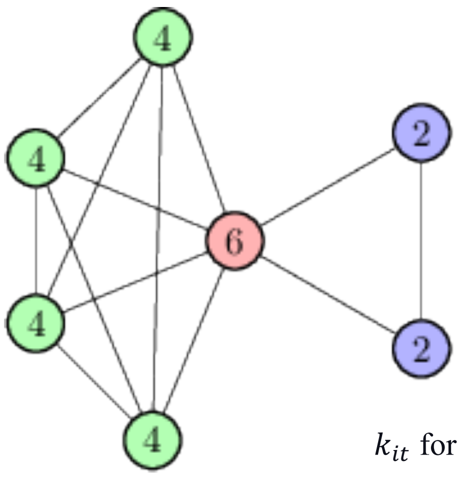
The topographical characteristics are based on graph theory. They describe the physical composition of the network; the CH considered are:

- Node degree: The node degree is determined by the number of connections (links, edges) that each node has with the rest of the network. In many cases, the distribution of node degrees is heterogeneous across networks. Nodes with a higher number of connections are considered putative hubs that complement the network's connectivity(Fundamentals of Brain Network Analysis ,2016).

$$k_i = \sum_{j \neq i} Aij$$

Equation 2:Node degree(Node Degree and Strength, 2016)

Where: k_i= the degree of nodes, i = number of edges connecting the nodes with all other nodes, j= 1 …. N -1 nodes, A is symmetric for an undirected graph.



$k_{i1} = 6$

$k_{i2} = 4$

$k_{i3} = 2$

k_{it} for the hole network = $k_{it} = \sum \frac{k_{i1}+k_{i2\ldots\ldots}}{nk}$

Where: k_{it}=average node degree, nk= total number of k_i

Figure 9:Example of node degree ((Maciej Serda, 2013)

Each node of the NM gives a node degree value; therefore, the average of the node degree was taken as a unique value for the hole NM.

- Link density (LD): describes the ratio of actual connections to connections in a real network (Newman, 2010). It is considered a connected network when the LD value is higher than a network with a lower density. In other words, the higher the LD, the network resists more failures in its edges.

$$LD = \frac{2i}{n(n-1)}$$

Equation 3:Link density(Newman, 2010)

Where: i = number of edges, n= number of nodes.

- Bridges: It is considered a bridge link that causes disconnection from the network when removing it. These links play an essential role in guaranteeing the connection of the networks. Therefore, it can be said that the vulnerability of a network with fewer bridges would be affected if any of them were removed(Tian et al., 2017).
- Articulation point: A node in a network is considered a articulation point when its removal disconnects the network or increases the number of connected components in the network(Ausiello et al., 2012).

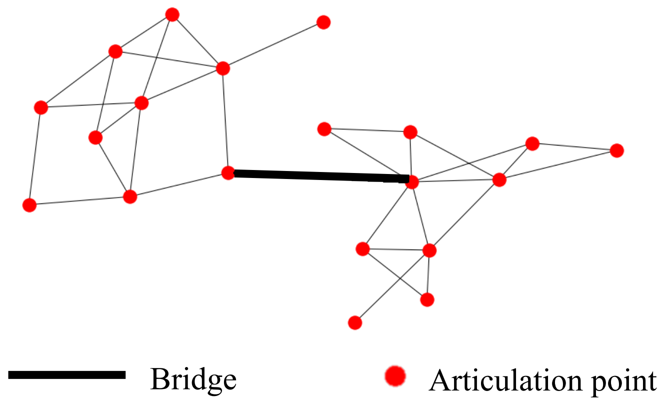


Figure 10: Example of bridge and articulation point(Python - Joining Two Networkx Graphs on a Single Edge - Stack Overflow, 2018.)

- Eccentricity: The eccentricity in a connected network is the maximum distance from node v to node u. It is calculated by summing the nodes on the shortest path from node v to all other nodes in the network. Eccentricity measures centrality and allows for evaluating the ease of access and flow control. (Krnc et al., 2018).
  - $E(G) = \sum_{v \in V(G)}^{n} \varepsilon(\upsilon)$

Equation 4:Eccentricity(Imran et al., 2018)

Where: G= Graph, ε(υ) = represents the eccentricity of vertex v, v = reservoir

- Closeness centrality: It measures the centrality of a node or link in a network. It is evaluated by assessing which nodes and links are the most compelling by determining the shortest paths between all nodes. The average centrality is the average centrality value between all evaluated nodes (Quuen Mary, 2008)

$$CC(i) = \frac{N-1}{\sum_j d(i,j)}$$

Equation 5:Closeness centrality (Network Analysis of Protein Interaction Data, n.d.)

Where: d = is the length of the shortest path between nodes i and j in the network, N = number of nodes.

- Average shortest path length: It is the sum of path length between all pairs of nodes evaluated in a network d(u,v)(Math Insight,Marinescu, 2017).

$$l = \frac{1}{N(N-1)} \sum_{i \neq j} dij$$

Equation 6: Average shortest path length(Math Insight, 2017)

Where: N= number of nodes, dij= length between node i and j.

Hydraulic metrics

Hydraulic CHs are based on demand, pressure, and flow. The calculation of these metrics requires the simulation of the network hydraulics to visualise the system performance under normal or abnormal conditions.

Therefore, the first CHs included are maximum pressure, minimum pressure, pressure and specific pressure threshold, total flow, average velocity, total demand and expected demand. Additional Ch are included as follow:

- Average demand per node:

$$Avrd = \frac{Td}{N}$$

Equation 7:Average demand per node

Where: Avrd=average demand, N= number of nodes (with demand different to 0), Td= total demand

- Average residence time

$$Avr = \frac{Vn}{\overline{D}}$$

Equation 8:Average residence time

Where: Avr=average resident time, Vn= volume of the network, $\overline{D}$= average demand per hour.

- Water service availability or fraction of delivered volume (FDV): it is the ratio between the total demand delivered to a node in all simulation runs, divided by the simulation runs, and the sum of the forecasted demand per consuming node in all simulation runs(Ostfeld et al., 2002).

$$FDVj = \frac{\sum_{i=1}^{N} \text{Vij}}{VT} \quad \forall\ of\ NN$$

Equation 9: Water service availability(Ostfeld et al., 2002)

Where: FDVj= the ratio of delivered volume per node, N=number of simulations, V= volume supply, j= consumer, i run, VT= total volume to supply a customer, NN= number of consumer nodes((with demand different to 0).

- Todini index : This characteristic of resilience is described as the ability of a system to overcome failures while satisfying the demands and pressures at the nodes. It is calculated as the surplus power (energy per unit time) at each node(Todini, 2000).

$$P_{tot} = \gamma \sum_{k=1}^{nr} Q_K H_K$$

Equation 10:Todini(Todini, 2000)

Where: γ= specific weigth of the water, Hk: head, Qk: discharge, nk: number of reservoirs, k= reservoir.

- Entropy: This CH can describe de redundancy and reaability of the system. The entropy can measure the network's capacity to respond to the failure of one of the links by finding multiple alternatives paths(Awumah et al., 1990).

$$S_j = -\sum_{i \in \overline{U}j} \left[\left(\frac{q_{ij}}{Q_j} ln \frac{q_{ij}}{Q_j} \right) \right] + \sum_{i \in \overline{U}j} \left[\frac{q_{ij}}{Q_J} \right] ln a_{ij}$$

Equation 11:Entropy(Awumah et al., 1990)

Where Sj= redundancy at node j, qij= flow from node i to j, Qj= Total flow in node j, $\overline{U}$j= Set of end nodes upstream of the links impinging on node j, a_{ij}= = effective number of independent paths from the node i to demand node j,

Water quality metrics

Water quality characteristics are based on assessing the concentration or age of the water; these require simulation of the water quality**.**

- Water age: Water age determines the time it takes for water to travel from the supply source to its consumers. This CH is considered an indicator of water quality due to the reactions between the pipe walls and the water(Kourbasis et al., 2020).

$$t_i = \frac{q_{ij(t_j+t_{ij})+} q_{ik(t_k+t_{ik})+STs}}{I_i}$$

Equation 12:Water age(Kourbasis et al., 2020)

- Population impacted: This CH describes the population impacted by quality over 24 hours of age.

Economic metrics

- Network cost: This CH is related to the annual cost of maintenance and operation of the network concerning pipes, tanks, valves and pumps.

- Greenhouse gas emission: This CH calculates the annual GHG emissions of each pipe in the network.

- Pump energy: This CH determines the cost of energy required for the operation of the pump.

4.3 Estimate the characteristics in the model networks

A python script was implemented to assess the components of the characterization. This script has a function to import the network, evaluate the formula of each component and determine a unique value. The principal libraries to simulate the hydraulic network analysis and evaluate the network's topology were Water Network Tool for Resilience-WNTR and NetworkX. WNTR is based on EPANET, a tool to simulate the movement and fate of the drinking water distribution system. (WNTR 0.1.4 Documentation.). NetworkX is a tool that contributes to the creation, manipulation, and study of the structure and dynamics of complex networks. (NetworkX Documentation).

Additionally, libraries such as Pandas, math, and NumPy were used to perform the operations necessary to generate the unique value of the features. Libraries such as Os and Json were used to allocate input files and store output data in both. json and Excel formats.

To check the good development of the characteristics in the script was necessary to validate the results of the network from the EPANET environment and the results of the functions used in the script for each component. That means that the CH were simulated in Epanet and cooperated with the script results.

As a result of the estimation and validation of the characteristics, a table of results of 38 characteristics was created employing a python script.

Charactetistic / Network name	DenHaag
Roughness (Milimeters)	0.01, 0.05, 0.08, 0.1, 0.2, 0.5, 0.76, 1.0, 2.0
Length (m)	34.96, 118406.11, 3.62, 662.03, 206821.37, 50209.38, 1523.75, 10.0, 412614.17
Average Length (m)	37.58074048
Total Length (m)	790285.3915
Total Flow (CMH (cubic meters/hr))	122.1300888
Total demand (CMH (cubic meters/hr))	104.6403427
Min elevation (m)	-7.000089
Max elevation(m)	14.840499
Number of nodes (units)	18999
Number of nodes demand >=0	5716
Average demand (units)	0.005507383
Average demand per node (units)	0.018306568
list of patterns (units)	FLAT, HRgem2015, K376692298RYS, L311058575DEH, K311114138DEH, N435980WSS, K120330939LVO
Valves segmentation (units)	80
Average Residence Time (units)	512736752.1
Node degree (degree)	2.216
Link density (units)	5.83189E-05
Articulation points (units)	114
Bridges (units)	111

Charactetistic / Network name	DenHaag
Eccentricity (m)	141.7763158
Maximun eccentricity (m)	192
Closeness eccentricity (m)	0.000129144
Average Shortest path length (m)	76.68815953
Max pressure (m (meters))	38.62200546
Pressure and specific pressure threshold 30m (m)	24.07368421
Water service availability (CMH (cubic meters/hr))	142975
Expected demand (CMH (cubic meters/hr))	35.54831429
Todini index, threshold (hours)	-0.02554143
Entropy (units)	2
Water age (hour)	27105.7832
Water age at 48 hours (hours)	7.529528618
Population that is impacted by water age greater than 24 hours (hours)	18997
Network cost (units)	10167583.41
Greenhouse gas emission (C02)	11380528.82
Pump energy (C02)	898510706.6

Table 3: Characterization of Den Haag

4.4 Analysis of the characteristics

Descriptive statistics analysis

The descriptive statistics analysis of the characteristics started with describing and understanding the relation of the variables in the same population. The equations applied were mean (central value of the population), median (the geometric position of the population), mode (the value with more repetition) and standard deviation (represent the dispersion or aggregation of data).

As a result of this descriptive statistical analysis was possible to determine for each CH the minimum and maximum values, the range of uncertainty, the frequency, the dispersion values and the central position.

Additionally, the GAUS bell was applied. This function represents the normal distribution of a set of data. It is seen graphically as a bell, and its values are distributed in low, medium and high values. For the component, have the pertinent graph. The Gaussian function is defined as:

$$f(x) = \frac{1}{\sigma * \sqrt{2 * \pi}} * e^{-\frac{1}{2}*\left(\frac{x1-mean}{\sigma}\right)^2}$$

Equation 13: Gaussian function

Where:
x1 = each value in the variable
σ = standard deviation

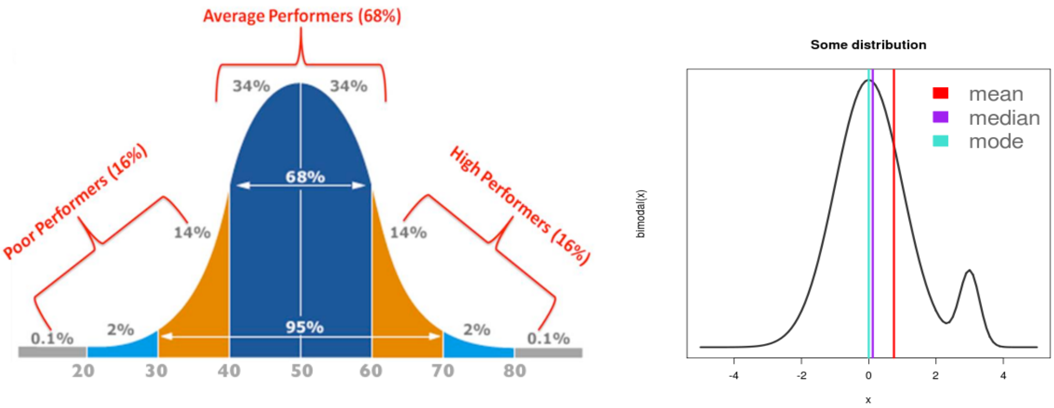


Figure 11: Example of Gaussian bell curve

The information plotted in the Gauss bell was also used in a box whisker diagram. The box diagrams can add many CH and visualize them in a single plot, unlike the Gaussian bell, where each component had a corresponding graph.

Figure 12: Example of box whisker diagram

Figure 2 shows how many variables can be a plot in the same graph on the left side. This is an easy visualization of the data. The right side is how the data can be evaluated depending on where the value fits in the graph.

Correlation analysis

This first analysis of the values was necessary to comprehend the components in the networks. Starting on this study to archive the first objective was also necessary to analyse the components in terms of how are the components correlated? And which are the similarity or differences in the components? To answer those questions was necessary to involve a multivariate analysis.This was the approach in two different ways:

Correlation Matrix: The correlation matrix shows how two or more variables are evaluated using Pearson's correlation. It measures the degree of the linear relationship between each pair of elements or variables. As a result of the evaluation a matrix is created with values between -1 and 1, usually the elements are positively correlated (Minitab).

For this case study the correlation matrix was created with the 38 components of the characterization. in it the correlation between the variables was evidenced and classified as follows:

Uncorrelated: between -0.2 to 0 and 0 to 0.2
low correlation: between -0.2 to -0.4 and 0.2 to 0.4
moderate correlation: between -0.4 to 0.6 and 0.4 to 0.6
high correlation: between -0.6 to -0.8 and 0.6 to 0.8
very high correlation between -0.8 to -0.99 and 0.8 to 0.99
perfect correlation -1 and 1

Pair plot: A pairs plot is a matrix used to graphically understand the pairwise relationship between different variables in a dataset. This matrix represents the scatter plot graph, a mathematical diagram that uses Cartesian coordinates to represent the values of a data set.(Touchette et al., 1985)

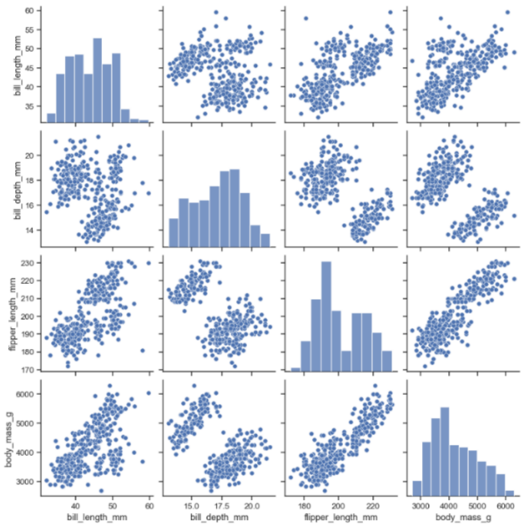


Figure 13: Example of a Pair pot graph

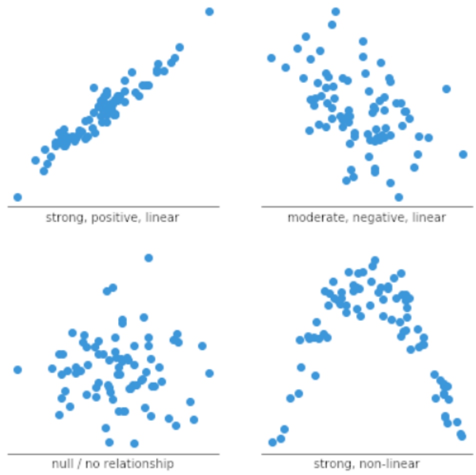


Figure 14:\classification of correlation in the data base on pair plot graph

Principal component Analysis

Base on the 38 components implemented in the characterization step and the complexity to the multi dimensions consist in all the components and the correlation with each other was necessary to apply the principal component analysis (PCA). PCA use a mathematical projection of the base data set to retains most of the variance among the data points(Dien, 2012).

To implement this method was necessary to made the following steps (Built In, 2021).

1. Normalize the data: each component in the characterization were normalized, this was with the propose to change the actual value of the data set to a common scale, without distorting differences in the ranges of values.

2. Calculate the covariance matrix: The calculation of the covariance aims to understand how the characteristics vary, in other words, to see if there is any relationship between them. As a result, a correlation matrix is obtained from this step.

3. Calculate the eigenvectors and eigenvalues: Based on the covariance matrix, the orthogonal of each vector is obtained, which represents the principal axes. The highest value corresponds to the highest variability, so an axis with a lower value will capture the lowest variability of the data.

   Arranging the eigenvalues in descending order will sort the principal component with the highest variability in the data set. Therefore, the first column will be the principal component.

4. Selection of the components: the number of components needed to describe the information was evaluated using the variance ratio. This determines the weight of each component. The number of components to choose will equal the weight of the components that, when added together, give a greater than 80 % starting from component 1.

Elbow method

The elbow method helps select the optimal number of clusters in the data(Syakur et al., 2018); therefore, after having determined the number of necessary components representing more than 80% of the information through the PCA, the next step is to apply this method. An arm represents the elbow method, and when it generates the elbow bend, it gives the value of the number of clusters needed for each dataset.

Silhouette method

The silhouette method is used to evaluate the quality of the clusters created by using algorithms such as K-means. Two measures determine the score of each cluster. The first is the average distance between the observed and all other data in the same cluster. The second is the average distance between the observed data and the nearest cluster. For our case, the number of clusters to be made is determined by the elbow method explained above. Based on the results, it should determine which CHs are grouped in each cluster.

Implementing methods such as the principal components elbow method and silhouette method was necessary to reduce the number of dimensions, identify the number of clusters and represent their relationship to the characteristics using the silhouette method. In other words, it is possible to reduce the 38 characteristics into fewer dimensions; this will depend on the variance of the data found in the principal component.

To develop these methods was necessary to use a python script that takes the information from the table of CH per network and evaluates the principal components. Based on the results, build the elbow method and silhouette method. The principal libraries used to assess this code were sklearn, seaborn, yellowbrick, and matplotlib.

4.5 Definition of an Optimal DMAs design

For this case study, the design of the DMAs was development base on the Gondwana software platform. The optimal DMAs definition has been preforming as follows:

Inputs

a. **Hydraulic model**= simulation model of the network created in EPANET.

Figure 15:Initial Interface of Gondwana

b. **Decision variables**= the decision variables had been defined as the range of the possible number of DMAs to be designed in the network.

In the United Kingdom, the minimum and maximum size for a DMA is typically equal to 500 and 5000 customer connections(Morrison, Tooms Stepthen, and Rogers Dewi 2007), (Ferrari, Savic, and Becciu 2013). In comparison, In the Netherland, the started values to define the number of DMAs to split the network has related to a wide variety of configuration ranges such as pipeline lengths from 3 to 459 km to DMAs and number of connections from 50 to 110,000 per DMAs. (Van Laarhoven et al. n.d.).

The minimum number of DMAs for this case study was fixed based on the number of customer connections per district. The number of connections was calculated on the basis of the base demand and the average consumption per connection per day.

The total base demand of the study area is 165122974.2 litres per day, and the average consumption per connection is 280 litres per day. The equation to find the total number of connections in the network is described as follow:

$$Ctot = \frac{total\ base\ demand}{average\ consume\ per\ conection}$$

Equation 14:Total number of connections

$$Ctot = \frac{165122974.2}{280}$$
$$Ctot = 589724.907$$

Where: Ctot is the total number of connections

Based on the maximum number of connections in each DMAs mentioned for Van Laarhoven, that is equal to 110000, where the relation between the total number of connections and the minimum connections per DMA was created.

$$DMA(\min) = \frac{Ctot}{Cmax}$$

Equation 15:minimum numbers of DMA

$$DMA(\min) = \frac{589724.907}{110000}$$

$$DMA(\min) = 5.36$$

Where: DMA (min) = Minimum number of DMAs in a network, Ctot = total connections Cmax= maximum number of connections per DMA

The above equation represents the possible number of DMAs for the Den Haag network.

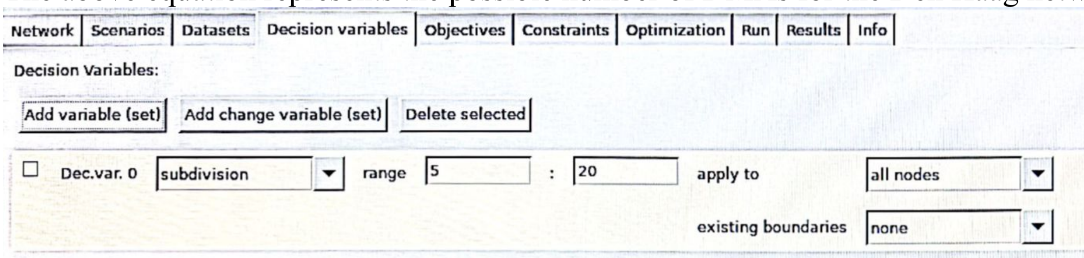


Figure 16: Gondwana, definition of decision variables

c. **Objectives function**= The objectives applied to this research are:

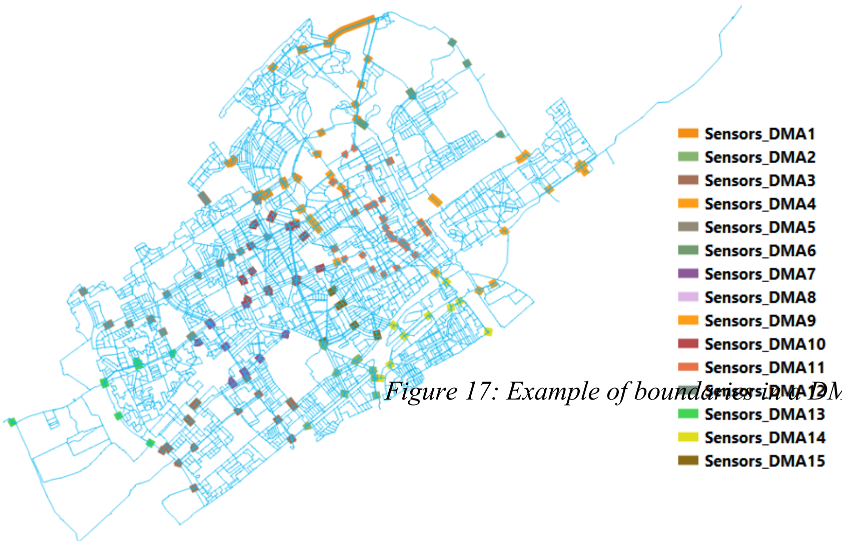
Objective 1 = Minimize the number of boundaries in the definition of DMAs. Defined with the following function.

$$Arg \min f(p) = \sum_{p1}^{n} (b(Pn)$$

Equation 16: Objective function 1 to design DMAs

Where p = pipes in the network, b = boundaries for DMA, Pn = pipes to delimitate a DMA, Arg min f (p): When f (p) takes the minimum value, the value of Pn

This objective 1 can be explained as a total number of sensors that have to be installed with de proposal to delimited an area and create a DMA in the all network.



	Number of sensors
Sensors_DMA1	8
Sensors_DMA2	10
Sensors_DMA3	12
Sensors_DMA4	15
Sensors_DMA5	9
Sensors_DMA6	3
Sensors_DMA7	5
Sensors_DMA8	9
Sensors_DMA9	3
Sensors_DMA10	3
Sensors_DMA11	1
Sensors_DMA12	4
Sensors_DMA13	12
Sensors_DMA14	6
Sensors_DMA15	7
Total Sensors	107

Figure 17: Example of boundaries in a DMA.

Figure 6 shows an example of objective 1. Each division of DMA is divided by the rest of the network for one or more boundaries (in this case a, sensors). This network was divided into 15 DMAs representing a total of 107 boundaries.

Objective 2 = Maximize the lowest boundaries flow in the pipes that are delimitate a DMA.

$$Arg \max f(p) \ = \sum_{p1}^{n} (FlowDMA1(Pn) + (FlowDMA2(Pn) + (FlowDMA..n(Pn)$$

Equation 17:Objective function 2 to design DMAs

Where: p = pipes in the network, Pn = pipes to delimitate a DMA, Flow = lowest flow in the boundaries of each DMA

Objective 2 can be described as a subdivision of each DMA's total demand in an equal form. This whit the purpose of finding the lowest consume in an area but are the same time five an appropriate size to the hole DMAs.

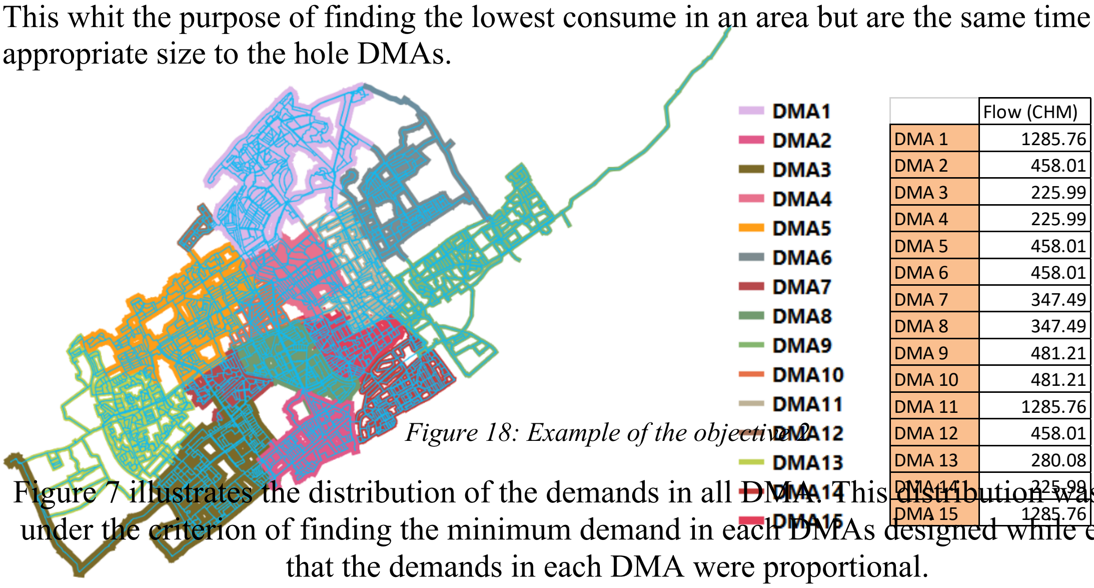


Figure 18: Example of the objective 2

Figure 7 illustrates the distribution of the demands in all DMA. This distribution was created under the criterion of finding the minimum demand in each DMAs designed while ensuring that the demands in each DMA were proportional.

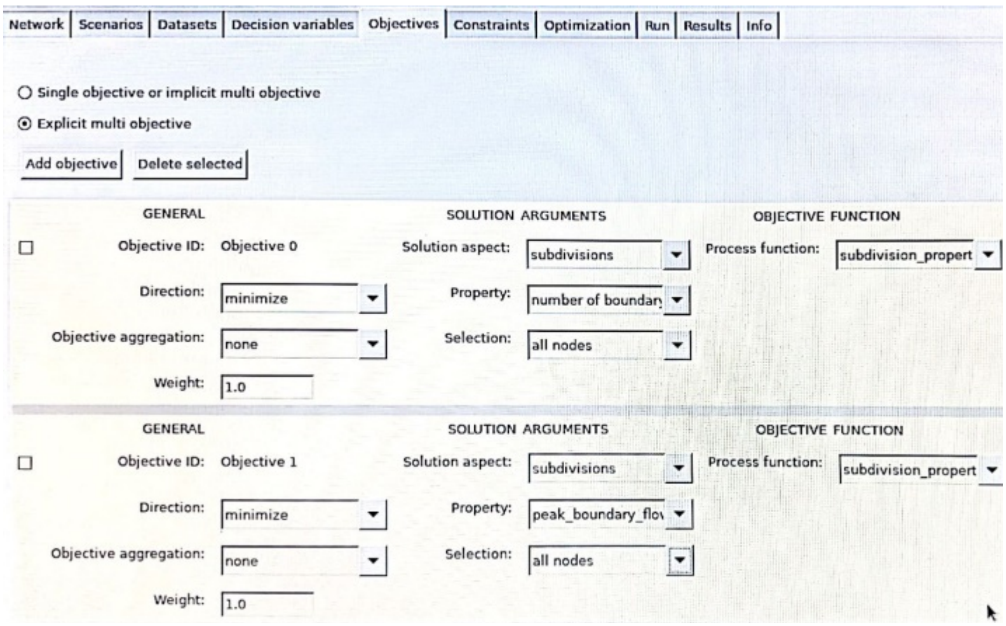


Figure 19: Gondwana, definition of objectives function

d. **Constraints** = the constraint have been defined as a maximum number of boundaries on the DMAS in a network. This maximum value is 500

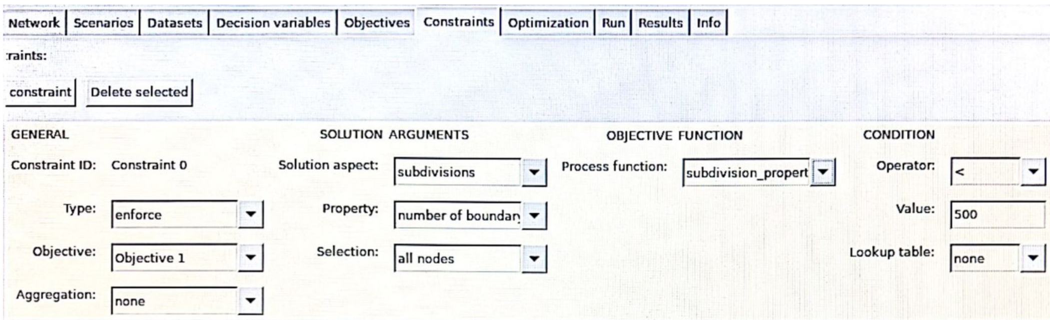


Figure 20:Gondwana, definition of constraints

Optimization Process

To perform the optimization progress Gondwana had been implemented a genetic optimization algorithm. The conceptual approach of the optimization process used by Gondwana is shown in figure:

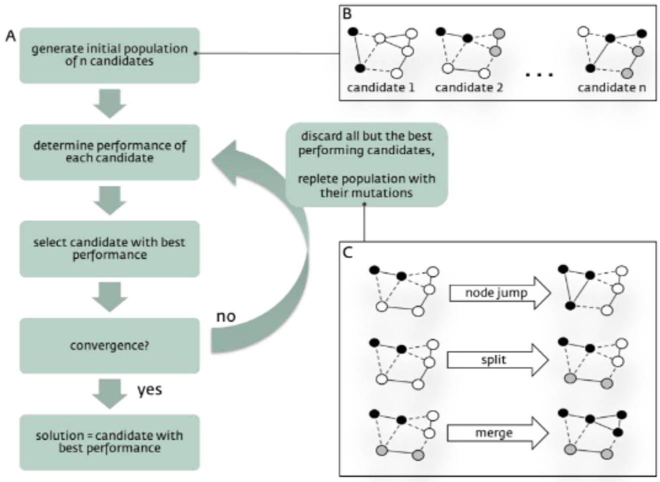


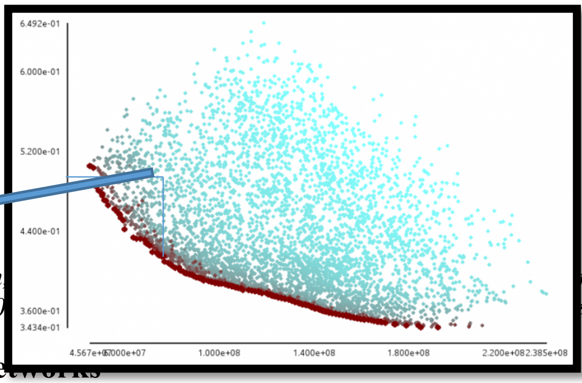
Figure 21: flowchart of the steps followed by the optimization platform Gondwana(van Laarhoven et al., n.d.).

(van Thienen & Vertommen, 2015)(van Thienen & Vertommen, 2015)(van Thienen & Vertommen, 2015)(van Thienen & Vertommen, 2015)As you can see in figure 8, the first step to DMAs design starts with generating n candidates. This means that after defining the inputs. After that the first step to the optimization process is to creates multiple possible DMA divisions (a population of individuals). These combinations are generated through graph theory, which identifies the possible pipelines that allow generating the division of the network in different numbers of DMAs. The second step is to evaluate the performance of each of the candidates. This will determine and classify the best candidates concerning the target function, decision variables and defined constraints. In the third step, the best-performing candidates are kept, whereas the rest have been replaced. new solutions are generated in two ways, either by making small variations on the solutions chosen as good, this is called mutation, or by making combinations between the elements of the solutions that work well; at this process call reproduction. The second and third steps are repeated, always taking the candidates with the best performance and made the mutation or reproduction; this optimisation process ends when the determined convergence value is fulfilled or the proposed number of generations is realised(van Thienen & Vertommen, 2015).

Output

As a result of the DMA design, a Pareto front is generated. It represented the generations run and evaluated to reach a certain number of generations or the convergence of the objectives. The last generation represented in the Pareto shows the best solutions generated, so there is no single optimal solution. For our case, the optimal solution was taken as the closest solution to the origin evaluated in the last generation.

Optimal solution



(van Thienen & Vertommen ... n & Vertommen, 2015)(van Thienen & Vertommen, 20 ... nen & Vertommen, 2015)

DMAs design for the networks

For the implementation of the DMA design, 60 networks were sampled. These 60 networks were selected from the population created in section 4.1, where reference is made to generating new NM from a base network.

The results of these DMAs designs were extracted by developing a python script that takes the JSON file (created by Gondwana) with the results of all generations and extracts only the information of the last generation; consequently, the "optimal solution" defined before. the main libraries used to get and organize the information were Json, Os, Numpy and Pandas

The principal variables extracted were the number of DMAs, number of nodes per DMAs, number of sensors to be installed in the system as a boundary to design DMAs and the score of the Pareto front.

As a result, a description of each DMAs design for a network, the following table was crate to consolidate the information.

	Number of DMAs	Nodes related with DMAs	Sensors to be install	X (xd)	Y (xd)
Network 1	10	{'8': 556, '7': 266, '10': 435, '5': 458, '2': 216, '1': 579, '4': 969, '9': 396, '3': 379, '6': 299}	304	337	5313.474
Network 2	20	{'2': 1761, '18': 139, '4': 751, '3': 1591, '7': 16, '13': 3, '5': 5, '8': 89, '14': 94, '1': 290, '9': 139, '19': 65, '16': 14, '15': 29, '20': 19, '10': 12, '11': 13, '6': 1, '17': 1, '12': 3}	266	294	3722.459
Network 3	11	{'5': 998, '3': 301, '10': 635, '11': 632, '7': 588, '8': 582, '9': 611, '2': 663, '4': 559, '6': 495, '1': 634}	304	318	3145.657

Table 4: Example of Description of DMAs design

Chapter 5 Results

5.1 Characterization of many networks

To know the general characteristics variation, behavior and range were required to involve different networks. Those networks are located in different cities because the components of the characterization will change and give the idea of the variables that influence the performance of the components.

These networks represent different cities, small, medium and large and therefore, their composition is different in nodes, pipes and valves, and reservoirs. Additionally, their consumption patterns and total demand vary considerably.
From the above, it can be said that the variables considered as topological and the hydraulic, water safety, water quality and economic variables were simulated and could determine the maximum and average minimum ranges of the networks in general.

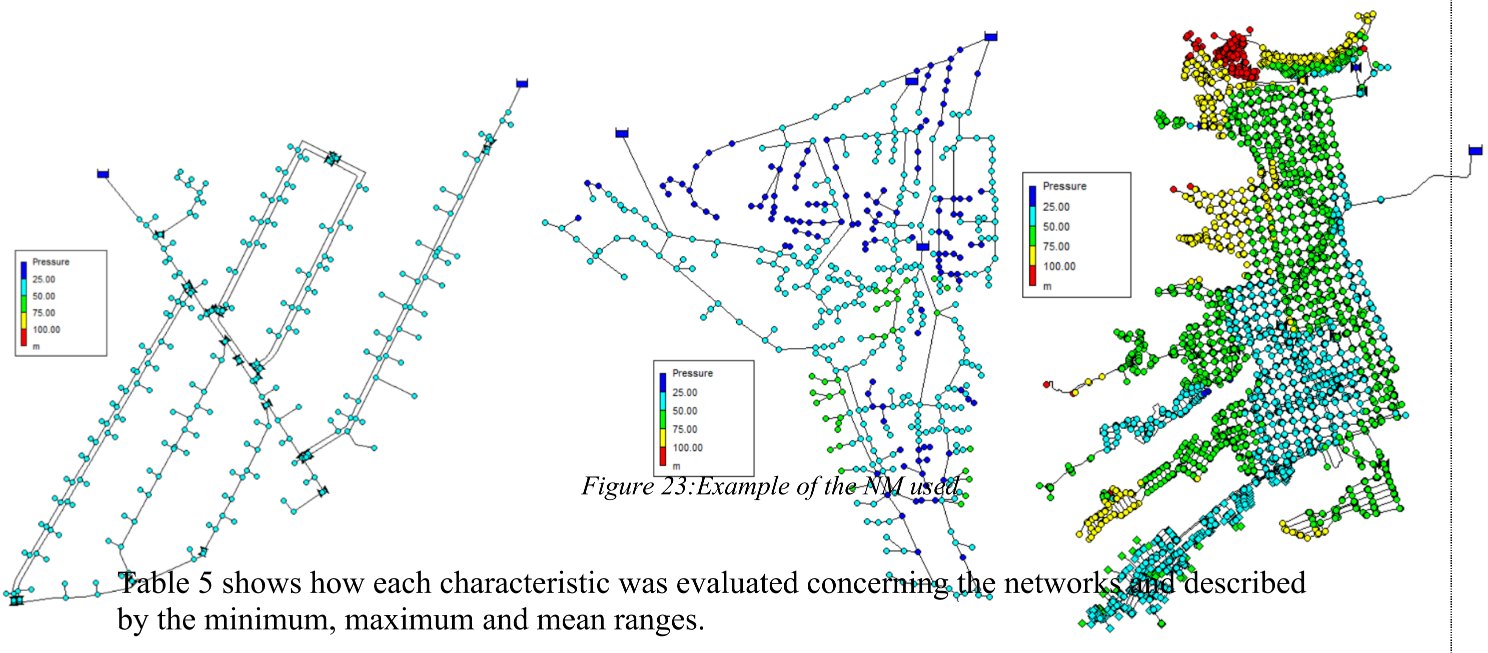


Figure 23:Example of the NM used

Table 5 shows how each characteristic was evaluated concerning the networks and described by the minimum, maximum and mean ranges.

	Characteristic	units	minimum	median	maximum
1	Average Length	m	8.08	223.68	5309.33
2	Total Length	m	8062.3	86034.35	12309660
3	Total Flow	CMH	0	10.29	313.48
4	Total demand	CMH	0.05	33.45	191.42
5	Min elevation	m	-7	0	1349.5
6	Max elevation	m	0	23	1762.32
7	Number of nodes	units	10	634	18999
8	Number of nodes demand less than or equal 0	units	0	330	5716
9	Average demand	units	0	0.01	5.71
10	Average demand per node	units	0	0.02	5.68
11	Number of pipes	units	17	674	21029
12	Number of tanks	units	0	0	4
13	Number of reservoir	units	0	1	4
14	Number of pumps	units	0	0	7
15	Number of valves	units	0	2.5	140
16	Number of patterns	units	0	4.5	1328
17	Valves segmentation	units	80	80	80
18	Average Residence Time	units	18628.35	57304173.9	1.43357E+11
19	Node degree	degree	1.93	2.22	4.2
20	Link density	units	0	0	0.13
21	Articulation points	units	3	114	1454
22	Bridges	units	4	111	1838
23	Eccentricity	m	4.58	67.53	156.4
24	Maximun eccentricity	m	6	92.5	207
25	Closeness eccentricity	m	0	0	0.16
26	Average Shortest path length	m	2.59	30.73	81.37
27	Max pressure	m	4.7	48.71	211719.78
28	Pressure and specific pressure threshold 30m	m	0	2.31	49.16
29	Water service availability	CMH	1	43552	6951257
30	Expected demand	CMH	-0.13	10.28	57.13
31	Todini index, threshold	hours	-5.07E+28	0.3	40281.48
32	Entropy	units	0	2	393
33	Water age	hour	0	26917.61	53445.85
34	Water age at 48 hours	hours	0	7.48	16.67
35	Population that is impacted by water age greater than 24 hours	hours	10	634	18997
36	Network cost	units	67161.21	1609085.4	142848610.2
37	Greenhouse gas emission	C02	47946.17	2340483.07	174504337.6
38	Pump energy	C02	0	0	6025146384

Table 5:Range values of different networks

Figure 22 represents the ranges of each characteristic using the box whisker plot, which graphically shows the ranges of the values and whether out-of-range or unusual values are represented.

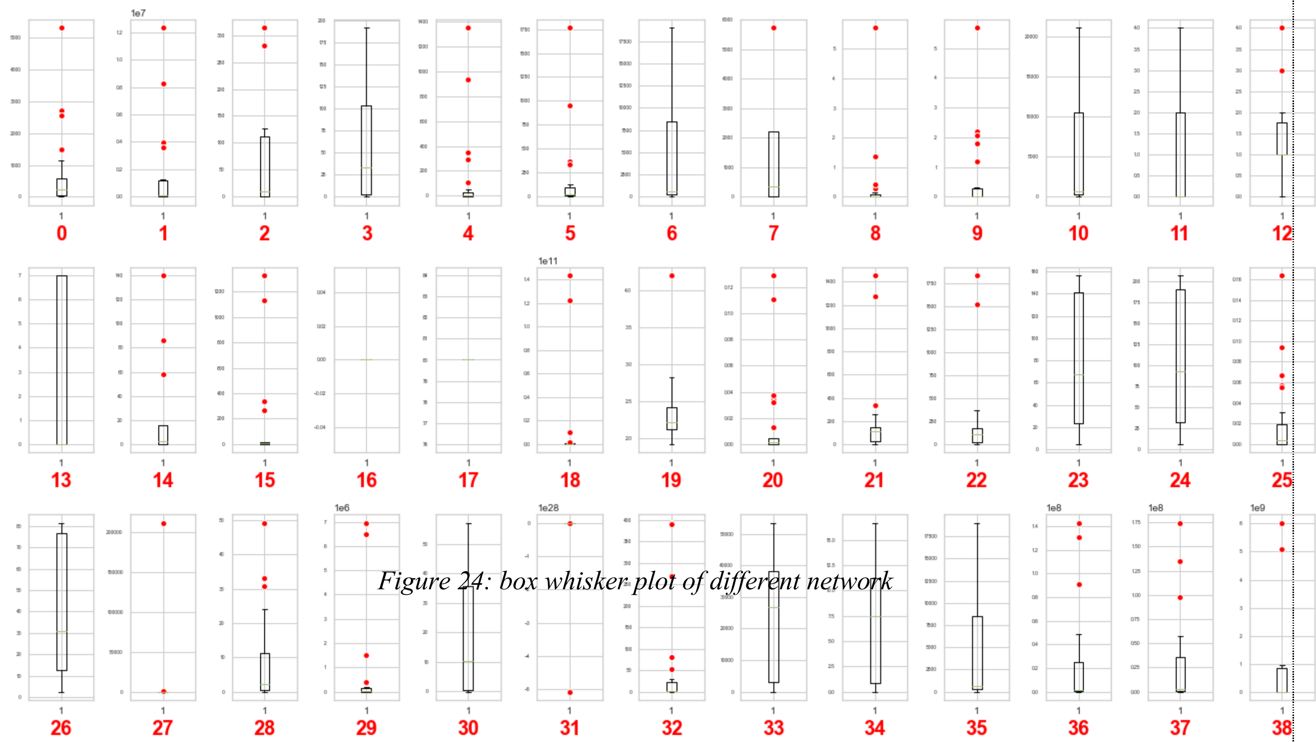


Figure 24: box whisker plot of different network

An additional analysis was carried out for the characteristics of the different networks. This analysis corresponds to the correlation matrix, in which the different characteristics of the networks were compared and classified into three main tables. The first one with characteristics with no correlation, the second one with characteristics with low correlation, the third one with medium correlation and finally the variables with high correlation.

Correlation matrix

	Average Length	Total Length	Total Flow	Total demand	Min elevation	Max elevation	Number of nodes	Nodes demand <=0	Average demand	Average demand per node	Number of pipes	Average Residence Time	Node degree	Link density	Articulation points	Bridges	Eccentricity	Maximun eccentricity	Closeness eccentricity	Average Shortest path length	Max pressure	Pressure and specific pressure threshold 30m	Water service availability	Expected demand	Todini index, threshold	Network cost	Green-house gas emission	Pump energy
Average Length	1	-0.05	0.36	0.06	-0.06	-0.07	-0.24	-0.26	0.81	-0.1	-0.24	-0.13	0.83	0.79	-0.27	-0.26	-0.48	-0.48	0.67	-0.46	-0.02	-0.18	-0.2	0.22	-0.32	-0.01	0.02	0.02
Total Length	-0.05	1	0.58	0.28	-0.2	-0.19	0.68	0.66	-0.14	-0.22	0.69	-0.14	-0.03	-0.22	-0.15	-0.16	0.48	0.48	-0.27	0.51	0.59	-0.18	-0.14	0.52	0.1	1	0.98	-0.02
Total Flow	0.36	0.58	1	0.53	-0.06	-0.04	0.61	0.6	0.48	-0.27	0.61	-0.18	0.45	0.15	-0.21	-0.22	0.36	0.37	-0.02	0.38	0.61	-0.09	-0.19	0.74	0.14	0.6	0.62	-0.09
Total demand	0.06	0.28	0.53	1	0.25	0.3	0.51	0.51	0.25	-0.26	0.51	-0.28	0.17	-0.08	-0.07	-0.09	0.45	0.47	-0.21	0.44	-0.12	0.03	-0.08	0.64	0.19	0.28	0.25	0.1
Min elevation	-0.06	-0.2	-0.06	0.25	1	0.99	-0.26	-0.25	-0.06	0.42	-0.26	0.07	-0.01	0	-0.04	-0.05	-0.15	-0.13	0	-0.2	-0.08	0.45	-0.14	-0.36	-0.11	-0.19	-0.17	-0.16
Max elevation	-0.07	-0.19	-0.04	0.3	0.99	1	-0.25	-0.24	-0.06	0.38	-0.25	0.05	-0.04	-0.01	-0.05	-0.06	-0.15	-0.12	0	-0.2	-0.08	0.44	-0.15	-0.37	-0.09	-0.18	-0.17	-0.13
Number of nodes	-0.24	0.68	0.61	0.51	-0.26	-0.25	1	0.99	-0.2	-0.28	1	-0.18	-0.14	-0.32	-0.08	-0.1	0.78	0.78	-0.38	0.81	0.35	-0.15	-0.08	0.68	0.14	0.68	0.66	0.01
Nodes demand >= 0	-0.26	0.66	0.6	0.51	-0.25	-0.24	0.99	1	-0.21	-0.31	0.99	-0.2	-0.17	-0.33	0.01	-0.01	0.83	0.83	-0.4	0.85	0.35	-0.14	0.01	0.67	0.15	0.66	0.64	-0.01
Average demand	0.81	-0.14	0.48	0.25	-0.06	-0.06	-0.2	-0.21	1	-0.03	-0.2	-0.09	0.81	0.62	-0.15	-0.14	-0.33	-0.33	0.44	-0.32	-0.06	-0.16	-0.1	0.4	0.05	-0.13	-0.12	0.09
Average demand per node	-0.1	-0.22	-0.27	-0.26	0.42	0.38	-0.28	-0.31	-0.03	1	-0.28	-0.13	-0.14	0	-0.07	-0.07	-0.3	-0.29	0.05	-0.3	-0.09	-0.05	-0.15	-0.33	0.04	-0.23	-0.23	0.18
Number of pipes	-0.24	0.69	0.61	0.51	-0.26	-0.25	1	0.99	-0.2	-0.28	1	-0.18	-0.13	-0.32	-0.09	-0.11	0.77	0.77	-0.38	0.8	0.35	-0.15	-0.1	0.68	0.14	0.69	0.67	0.01
Average Residence Time	-0.13	-0.14	-0.18	-0.28	0.07	0.05	-0.18	-0.2	-0.09	-0.13	-0.18	1	0.11	-0.11	-0.11	-0.1	-0.24	-0.24	-0.1	-0.23	-0.05	0.48	-0.11	-0.28	0.06	-0.15	-0.15	-0.13
Node degree	0.83	-0.03	0.45	0.17	-0.01	-0.04	-0.14	-0.17	0.81	-0.14	-0.13	0.11	1	0.68	-0.31	-0.3	-0.38	-0.38	0.52	-0.36	-0.05	-0.01	-0.24	0.39	-0.24	-0.01	0	-0.09
Link density	0.79	-0.22	0.15	-0.08	0	-0.01	-0.32	-0.33	0.62	0	-0.32	-0.11	0.68	1	-0.23	-0.22	-0.54	-0.54	0.96	-0.52	-0.09	-0.27	-0.16	0.03	-0.71	-0.22	-0.22	0.01
Articulation points	-0.27	-0.15	-0.21	-0.07	-0.04	-0.05	-0.08	0.01	-0.15	-0.07	-0.09	-0.11	-0.31	-0.23	1	1	0.48	0.47	-0.27	0.48	-0.05	-0.16	0.97	-0.04	0.11	-0.16	-0.16	-0.19
Bridges	-0.26	-0.16	-0.22	-0.09	-0.05	-0.06	-0.1	-0.01	-0.14	-0.07	-0.11	-0.1	-0.3	-0.22	1	1	0.46	0.45	-0.25	0.45	-0.05	-0.17	0.97	-0.05	0.1	-0.17	-0.18	-0.18
Eccentricity	-0.48	0.48	0.36	0.45	-0.15	-0.15	0.78	0.83	-0.33	-0.3	0.77	-0.24	-0.38	-0.54	0.48	0.46	1	1	-0.62	1	0.26	-0.07	0.45	0.55	0.26	0.47	0.45	-0.16
Maximun eccentricity	-0.48	0.48	0.37	0.47	-0.13	-0.12	0.78	0.83	-0.33	-0.29	0.77	-0.24	-0.38	-0.54	0.47	0.45	1	1	-0.62	0.99	0.26	-0.06	0.44	0.54	0.26	0.47	0.45	-0.16
Closeness eccentricity	0.67	-0.27	-0.02	-0.21	0	0	-0.38	-0.4	0.44	0.05	-0.38	-0.1	0.52	0.96	-0.27	-0.25	-0.62	-0.62	1	-0.6	-0.11	-0.3	-0.19	-0.15	-0.76	-0.27	-0.27	0.08
Average Shortest path length	[illegible]	[illegible]	0.38	0.44	-0.2	-0.2	0.81	0.85	-0.32	-0.3	[illegible]	[illegible]	-0.36	-0.52	0.48	0.45	1	0.99	-0.6	1	0.27	-0.11	0.44	0.56	0.25	0.5	0.48	-0.14
Max pressure	[illegible]	[illegible]	0.61	-0.12	-0.08	-0.08	0.35	0.35	-0.06	-0.09	[illegible]	[illegible]	-0.05	-0.09	-0.05	-0.05	0.26	0.26	-0.11	0.27	1	-0.01	-0.03	0.21	0.04	0.61	0.67	-0.09
Pressure and specific pressure threshold 30m	-0.18	-0.18	-0.09	0.03	0.45	0.44	-0.15	-0.14	-0.16	-0.05	-0.15	0.48	-0.01	-0.27	-0.16	-0.17	-0.07	-0.06	-0.3	-0.11	-0.01	1	-0.17	-0.24	0.15	-0.16	-0.13	-0.2
Water service availability	-0.2	-0.14	-0.19	-0.08	-0.14	-0.15	-0.08	0.01	-0.1	-0.15	-0.1	-0.11	-0.24	-0.16	0.97	0.97	0.45	0.44	-0.19	0.44	-0.03	-0.17	1	0.04	0.07	-0.15	-0.15	-0.15
Expected demand	0.22	0.52	0.74	0.64	-0.36	-0.37	0.68	0.67	0.4	-0.33	0.68	-0.28	0.39	0.03	-0.04	-0.05	0.55	0.54	-0.15	0.56	0.21	-0.24	0.04	1	0.19	0.51	0.49	0.04
Todini	-0.32	0.1	0.14	0.19	-0.11	-0.09	0.14	0.15	0.05	0.04	0.14	0.06	-0.24	-0.71	0.11	0.1	0.26	0.26	-0.76	0.25	0.04	0.15	0.07	0.19	1	0.11	0.11	0.09
Network cost	-0.01	1	0.6	0.28	-0.19	-0.18	0.68	0.66	-0.13	-0.23	0.69	-0.15	-0.01	-0.22	-0.16	-0.17	0.47	0.47	-0.27	0.5	0.61	-0.16	-0.15	0.51	0.11	1	0.99	-0.03
Greenhouse gas emission	0.02	0.98	0.62	0.25	-0.17	-0.17	0.66	0.64	-0.12	-0.23	0.67	-0.15	0	-0.22	-0.16	-0.18	0.45	0.45	-0.27	0.48	0.67	-0.13	-0.15	0.49	0.11	0.99	1	-0.04
Pump energy	0.02	-0.02	-0.09	0.1	-0.16	-0.13	0.01	-0.01	0.09	0.18	0.01	-0.18	0.09	0.01	-0.19	-0.18	-0.16	-0.16	0.08	-0.14	-0.09	-0.2	-0.15	0.04	0.09	-0.03	-0.04	1

No negative correlation = -0.2 – 0

No positive correlation = 0 – 0.2

Table 6: Matrix of the CH with no correlation for many networks

Correlation matrix

	Average Length	Total Length	Total Flow	Total demand	Min elevation	Max elevation	Number of nodes	Nodes demand <=0	Average demand	Average demand per node	Number of pipes	Average Residence Time	Node degree	Link density	Articulation points	Bridges	Eccentricity	Maximun eccentricity	Closeness eccentricity	Average Shortest path length	Max pressure	Pressure and specific pressure threshold 30m	Water service availability	Expected demand	Todini index, threshold	Network cost	Green-house gas emission	Pump energy
Average Length	1	-0.05	0.36	0.06	-0.06	-0.07	-0.24	-0.26	0.81	-0.1	-0.24	-0.13	0.83	0.79	-0.27	-0.26	-0.48	-0.48	0.67	-0.46	-0.02	-0.18	-0.2	0.22	-0.32	-0.01	0.02	0.02
Total Length	-0.05	1	0.58	0.28	-0.2	-0.19	0.68	0.66	-0.14	-0.22	0.69	-0.14	-0.03	-0.22	-0.15	-0.16	0.48	0.48	-0.27	0.51	0.59	-0.18	-0.14	0.52	0.1	1	0.98	-0.02
Total Flow	0.36	0.58	1	0.53	-0.06	-0.04	0.61	0.6	0.48	-0.27	0.61	-0.18	0.45	0.15	-0.21	-0.22	0.36	0.37	-0.02	0.38	0.61	-0.09	-0.19	0.74	0.14	0.6	0.62	-0.09
Total demand	0.06	0.28	0.53	1	0.25	0.3	0.51	0.51	0.25	-0.26	0.51	-0.28	0.17	-0.08	-0.07	-0.09	0.45	0.47	-0.21	0.44	-0.12	0.03	-0.08	0.64	0.19	0.28	0.25	0.1
Min elevation	-0.06	-0.2	-0.06	0.25	1	0.99	-0.26	-0.25	-0.06	0.42	-0.26	0.07	-0.01	0	-0.04	-0.05	-0.15	-0.13	0	-0.2	-0.08	0.45	-0.14	-0.36	-0.11	-0.19	-0.17	-0.16
Max elevation	-0.07	-0.19	-0.04	0.3	0.99	1	-0.25	-0.24	-0.06	0.38	-0.25	0.05	-0.04	-0.01	-0.05	-0.06	-0.15	-0.12	0	-0.2	-0.08	0.44	-0.15	-0.37	-0.09	-0.18	-0.17	-0.13
Number of nodes	-0.24	0.68	0.61	0.51	-0.26	-0.25	1	0.99	-0.2	-0.28	1	-0.18	-0.14	-0.32	-0.08	-0.1	0.78	0.78	-0.38	0.81	0.35	-0.15	-0.08	0.68	0.14	0.68	0.66	0.01
Nodes demand >= 0	-0.26	0.66	0.6	0.51	-0.25	-0.24	0.99	1	-0.21	-0.31	0.99	-0.2	-0.17	-0.33	0.01	-0.01	0.83	0.83	-0.4	0.85	0.35	-0.14	0.01	0.67	0.15	0.66	0.64	-0.01
Average demand	0.81	-0.14	0.48	0.25	-0.06	-0.06	-0.2	-0.21	1	-0.03	-0.2	-0.09	0.81	0.62	-0.15	-0.14	-0.33	-0.33	0.44	-0.32	-0.06	-0.16	-0.1	0.4	0.05	-0.13	-0.12	0.09
Average demand per node	-0.1	-0.22	-0.27	-0.26	0.42	0.38	-0.28	-0.31	-0.03	1	-0.28	-0.13	-0.14	0	-0.07	-0.07	-0.3	-0.29	0.05	-0.3	-0.09	-0.05	-0.15	-0.33	0.04	-0.23	-0.23	0.18
Number of pipes	-0.24	0.69	0.61	0.51	-0.26	-0.25	1	0.99	-0.2	-0.28	1	-0.18	-0.13	-0.32	-0.09	-0.11	0.77	0.77	-0.38	0.8	0.35	-0.15	-0.1	0.68	0.14	0.69	0.67	0.01
Average Residence Time	-0.13	-0.14	-0.18	-0.28	0.07	0.05	-0.18	-0.2	-0.09	-0.13	-0.18	1	0.11	-0.11	-0.11	-0.1	-0.24	-0.24	-0.1	-0.23	-0.05	0.48	-0.11	-0.28	0.06	-0.15	-0.15	-0.13
Node degree	0.83	-0.03	0.45	0.17	-0.01	-0.04	-0.14	-0.17	0.81	-0.14	-0.13	0.11	1	0.68	-0.31	-0.3	-0.38	-0.38	0.52	-0.36	-0.05	-0.01	-0.24	0.39	-0.24	-0.01	0	-0.09
Link density	0.79	-0.22	0.15	-0.08	0	-0.01	-0.32	-0.33	0.62	0	-0.32	-0.11	0.68	1	-0.23	-0.22	-0.54	-0.54	0.96	-0.52	-0.09	-0.27	-0.16	0.03	-0.71	-0.22	-0.22	0.01
Articulation points	-0.27	-0.15	-0.21	-0.07	-0.04	-0.05	-0.08	0.01	-0.15	-0.07	-0.09	-0.11	-0.31	-0.23	1	1	0.48	0.47	-0.27	0.48	-0.05	-0.16	0.97	-0.04	0.11	-0.16	-0.16	-0.19
Bridges	-0.26	-0.16	-0.22	-0.09	-0.05	-0.06	-0.1	-0.01	-0.14	-0.07	-0.11	-0.1	-0.3	-0.22	1	1	0.46	0.45	-0.25	0.45	-0.05	-0.17	0.97	-0.05	0.1	-0.17	-0.18	-0.18
Eccentricity	-0.48	0.48	0.36	0.45	-0.15	-0.15	0.78	0.83	-0.33	-0.3	0.77	-0.24	-0.38	-0.54	0.48	0.46	1	1	-0.62	1	0.26	-0.07	0.45	0.55	0.26	0.47	0.45	-0.16
Maximun eccentricity	-0.48	0.48	0.37	0.47	-0.13	-0.12	0.78	0.83	-0.33	-0.29	0.77	-0.24	-0.38	-0.54	0.47	0.45	1	1	-0.62	0.99	0.26	-0.06	0.44	0.54	0.26	0.47	0.45	-0.16
Closeness eccentricity	0.67	-0.27	-0.02	-0.21	0	0	-0.38	-0.4	0.44	0.05	-0.38	-0.1	0.52	0.96	-0.27	-0.25	-0.62	-0.62	1	-0.6	-0.11	-0.3	-0.19	-0.15	-0.76	-0.27	-0.27	0.08

Weak negative correlation = -0.2 – -0.4

Weak positive correlation = 0.2 – 0.4

Table 7: Matrix of the CH with weak correlation for many networks

Correlation matrix

	Average Length	Total Length	Total Flow	Total demand	Min elevation	Max elevation	Number of nodes	Nodes demand <=0	Average demand	Average demand per node	Number of pipes	Average Residence Time	Node degree	Link density	Articulation points	Bridges	Eccentricity	Maximun eccentricity	Closeness eccentricity	Average Shortest path length	Max pressure	Pressure and specific pressure threshold 30m	Water service availability	Expected demand	Todini index, threshold	Network cost	Green-house gas emission	Pump energy
Average Length	1	-0.05	0.36	0.06	-0.06	-0.07	-0.24	-0.26	0.81	-0.1	-0.24	-0.13	0.83	0.79	-0.27	-0.26	-0.48	-0.48	0.67	-0.46	-0.02	-0.18	-0.2	0.22	-0.32	-0.01	0.02	0.02
Total Length	-0.05	1	0.58	0.28	-0.2	-0.19	0.68	0.66	-0.14	-0.22	0.69	-0.14	-0.03	-0.22	-0.15	-0.16	0.48	0.48	-0.27	0.51	0.59	-0.18	-0.14	0.52	0.1	1	0.98	-0.02
Total Flow	0.36	0.58	1	0.53	-0.06	-0.04	0.61	0.6	0.48	-0.27	0.61	-0.18	0.45	0.15	-0.21	-0.22	0.36	0.37	-0.02	0.38	0.61	-0.09	-0.19	0.74	0.14	0.6	0.62	-0.09
Total demand	0.06	0.28	0.53	1	0.25	0.3	0.51	0.51	0.25	-0.26	0.51	-0.28	0.17	-0.08	-0.07	-0.09	0.45	0.47	-0.21	0.44	-0.12	0.03	-0.08	0.64	0.19	0.28	0.25	0.1
Min elevation	-0.06	-0.2	-0.06	0.25	1	0.99	-0.26	-0.25	-0.06	0.42	-0.26	0.07	-0.01	0	-0.04	-0.05	-0.15	-0.13	0	-0.2	-0.08	0.45	-0.14	-0.36	-0.11	-0.19	-0.17	-0.16
Max elevation	-0.07	-0.19	-0.04	0.3	0.99	1	-0.25	-0.24	-0.06	0.38	-0.25	0.05	-0.04	-0.01	-0.05	-0.06	-0.15	-0.12	0	-0.2	-0.08	0.44	-0.15	-0.37	-0.09	-0.18	-0.17	-0.13
Number of nodes	-0.24	0.68	0.61	0.51	-0.26	-0.25	1	0.99	-0.2	-0.28	1	-0.18	-0.14	-0.32	-0.08	-0.1	0.78	0.78	-0.38	0.81	0.35	-0.15	-0.08	0.68	0.14	0.68	0.66	0.01
Nodes demand >= 0	-0.26	0.66	0.6	0.51	-0.25	-0.24	0.99	1	-0.21	-0.31	0.99	-0.2	-0.17	-0.33	0.01	-0.01	0.83	0.83	-0.4	0.85	0.35	-0.14	0.01	0.67	0.15	0.66	0.64	-0.01
Average demand	0.81	-0.14	0.48	0.25	-0.06	-0.06	-0.2	-0.21	1	-0.03	-0.2	-0.09	0.81	0.62	-0.15	-0.14	-0.33	-0.33	0.44	-0.32	-0.06	-0.16	-0.1	0.4	0.05	-0.13	-0.12	0.09
Average demand per node	-0.1	-0.22	-0.27	-0.26	0.42	0.38	-0.28	-0.31	-0.03	1	-0.28	-0.13	-0.14	0	-0.07	-0.07	-0.3	-0.29	0.05	-0.3	-0.09	-0.05	-0.15	-0.33	0.04	-0.23	-0.23	0.18
Number of pipes	-0.24	0.69	0.61	0.51	-0.26	-0.25	1	0.99	-0.2	-0.28	1	-0.18	-0.13	-0.32	-0.09	-0.11	0.77	0.77	-0.38	0.8	0.35	-0.15	-0.1	0.68	0.14	0.69	0.67	0.01
Average Residence Time	-0.13	-0.14	-0.18	-0.28	0.07	0.05	-0.18	-0.2	-0.09	-0.13	-0.18	1	0.11	-0.11	-0.11	-0.1	-0.24	-0.24	-0.1	-0.23	-0.05	0.48	-0.11	-0.28	0.06	-0.15	-0.15	-0.13
Node degree	0.83	-0.03	0.45	0.17	-0.01	-0.04	-0.14	-0.17	0.81	-0.14	-0.13	0.11	1	0.68	-0.31	-0.3	-0.38	-0.38	0.52	-0.36	-0.05	-0.01	-0.24	0.39	-0.24	-0.01	0	-0.09
Link density	0.79	-0.22	0.15	-0.08	0	-0.01	-0.32	-0.33	0.62	0	-0.32	-0.11	0.68	1	-0.23	-0.22	-0.54	-0.54	0.96	-0.52	-0.09	-0.27	-0.16	0.03	-0.71	-0.22	-0.22	0.01
Articulation points	-0.27	-0.15	-0.21	-0.07	-0.04	-0.05	-0.08	0.01	-0.15	-0.07	-0.09	-0.11	-0.31	-0.23	1	1	0.48	0.47	-0.27	0.48	-0.05	-0.16	0.97	-0.04	0.11	-0.16	-0.16	-0.19
Bridges	-0.26	-0.16	-0.22	-0.09	-0.05	-0.06	-0.1	-0.01	-0.14	-0.07	-0.11	-0.1	-0.3	-0.22	1	1	0.46	0.45	-0.25	0.45	-0.05	-0.17	0.97	-0.05	0.1	-0.17	-0.18	-0.18
Eccentricity	[illegible]	0.48	0.36	0.45	-0.15	-0.15	0.78	0.83	-0.33	-0.3	[illegible]	[illegible]	[illegible]	-0.54	0.48	0.46	1	1	-0.62	1	0.26	-0.07	0.45	0.55	0.26	0.47	0.45	-0.16
Maximun eccentric[illegible]	[illegible]	0.48	0.37	0.47	-0.13	-0.12	0.78	0.83	-0.33	-0.29	[illegible]	[illegible]	[illegible]	-0.54	0.47	0.45	1	1	-0.62	0.99	0.26	-0.06	0.44	0.54	0.26	0.47	0.45	-0.16
Closeness eccentri[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.4	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.11	-0.3	-0.19	-0.15	-0.76	-0.27	-0.27	0.08
Average Shortest path length	-0.46	0.51	0.38	0.44	-0.2	-0.2	0.81	0.85	-0.32	-0.3	0.8	-0.23	-0.36	-0.52	0.48	0.45	1	0.99	-0.6	1	0.27	-0.11	0.44	0.56	0.25	0.5	0.48	-0.14
Max pressure	-0.02	0.59	0.61	-0.12	-0.08	-0.08	0.35	0.35	-0.06	-0.09	0.35	-0.05	-0.05	-0.09	-0.05	-0.05	0.26	0.26	-0.11	0.27	1	-0.01	-0.03	0.21	0.04	0.61	0.67	-0.09
Pressure and specific pressure threshold 30m	-0.18	-0.18	-0.09	0.03	0.45	0.44	-0.15	-0.14	-0.16	-0.05	-0.15	0.48	-0.01	-0.27	-0.16	-0.17	-0.07	-0.06	-0.3	-0.11	-0.01	1	-0.17	-0.24	0.15	-0.16	-0.13	-0.2
Water service availability	-0.2	-0.14	-0.19	-0.08	-0.14	-0.15	-0.08	0.01	-0.1	-0.15	-0.1	-0.11	-0.24	-0.16	0.97	0.97	0.45	0.44	-0.19	0.44	-0.03	-0.17	1	0.04	0.07	-0.15	-0.15	-0.15
Expected demand	0.22	0.52	0.74	0.64	-0.36	-0.37	0.68	0.67	0.4	-0.33	0.68	-0.28	0.39	0.03	-0.04	-0.05	0.55	0.54	-0.15	0.56	0.21	-0.24	0.04	1	0.19	0.51	0.49	0.04
Todini	-0.32	0.1	0.14	0.19	-0.11	-0.09	0.14	0.15	0.05	0.04	0.14	0.06	-0.24	-0.71	0.11	0.1	0.26	0.26	-0.76	0.25	0.04	0.15	0.07	0.19	1	0.11	0.11	0.09
Network cost	-0.01	1	0.6	0.28	-0.19	-0.18	0.68	0.66	-0.13	-0.23	0.69	-0.15	-0.01	-0.22	-0.16	-0.17	0.47	0.47	-0.27	0.5	0.61	-0.16	-0.15	0.51	0.11	1	0.99	-0.03
Greenhouse gas emission	0.02	0.98	0.62	0.25	-0.17	-0.17	0.66	0.64	-0.12	-0.23	0.67	-0.15	0	-0.22	-0.16	-0.18	0.45	0.45	-0.27	0.48	0.67	-0.13	-0.15	0.49	0.11	0.99	1	-0.04
Pump energy	0.02	-0.02	-0.09	0.1	-0.16	-0.13	0.01	-0.01	0.09	0.18	0.01	-0.13	-0.09	0.01	-0.19	-0.18	-0.16	-0.16	0.08	-0.14	-0.09	-0.2	-0.15	0.04	0.09	-0.03	-0.04	1

Moderate negative correlation = -0.4 – -0.6

Moderate positive correlation = 0.4 – 0.6

Table 8: Matrix of the CH with moderate correlation for many networks

Correlation matrix

	Average Length	Total Length	Total Flow	Total demand	Min elevation	Max elevation	Number of nodes	Nodes demand <=0	Average demand	Average demand per node	Number of pipes	Average Residence Time	Node degree	Link density	Articulation points	Bridges	Eccentricity	Maximun eccentricity	Closeness eccentricity	Average Shortest path length	Max pressure	Pressure and specific pressure threshold 30m	Water service availability	Expected demand	Todini index, threshold	Network cost	Green-house gas emission	Pump energy
Average Length	1	-0.05	0.36	0.06	-0.06	-0.07	-0.24	-0.26	0.81	-0.1	-0.24	-0.13	0.83	0.79	-0.27	-0.26	-0.48	-0.48	0.67	-0.46	-0.02	-0.18	-0.2	0.22	-0.32	-0.01	0.02	0.02
Total Length	-0.05	1	0.58	0.28	-0.2	-0.19	0.68	0.66	-0.14	-0.22	0.69	-0.14	-0.03	-0.22	-0.15	-0.16	0.48	0.48	-0.27	0.51	0.59	-0.18	-0.14	0.52	0.1	1	0.98	-0.02
Total Flow	0.36	0.58	1	0.53	-0.06	-0.04	0.61	0.6	0.48	-0.27	0.61	-0.18	0.45	0.15	-0.21	-0.22	0.36	0.37	-0.02	0.38	0.61	-0.09	-0.19	0.74	0.14	0.6	0.62	-0.09
Total demand	0.06	0.28	0.53	1	0.25	0.3	0.51	0.51	0.25	-0.26	0.51	-0.28	0.17	-0.08	-0.07	-0.09	0.45	0.47	-0.21	0.44	-0.12	0.03	-0.08	0.64	0.19	0.28	0.25	0.1
Min elevation	-0.06	-0.2	-0.06	0.25	1	0.99	-0.26	-0.25	-0.06	0.42	-0.26	0.07	-0.01	0	-0.04	-0.05	-0.15	-0.13	0	-0.2	-0.08	0.45	-0.14	-0.36	-0.11	-0.19	-0.17	-0.16
Max elevation	-0.07	-0.19	-0.04	0.3	0.99	1	-0.25	-0.24	-0.06	0.38	-0.25	0.05	-0.04	-0.01	-0.05	-0.06	-0.15	-0.12	0	-0.2	-0.08	0.44	-0.15	-0.37	-0.09	-0.18	-0.17	-0.13
Number of nodes	-0.24	0.68	0.61	0.51	-0.26	-0.25	1	0.99	-0.2	-0.28	1	-0.18	-0.14	-0.32	-0.08	-0.1	0.78	0.78	-0.38	0.81	0.35	-0.15	-0.08	0.68	0.14	0.68	0.66	0.01
Nodes demand >= 0	-0.26	0.66	0.6	0.51	-0.25	-0.24	0.99	1	-0.21	-0.31	0.99	-0.2	-0.17	-0.33	0.01	-0.01	0.83	0.83	-0.4	0.85	0.35	-0.14	0.01	0.67	0.15	0.66	0.64	-0.01
Average demand	0.81	-0.14	0.48	0.25	-0.06	-0.06	-0.2	-0.21	1	-0.03	-0.2	-0.09	0.81	0.62	-0.15	-0.14	-0.33	-0.33	0.44	-0.32	-0.06	-0.16	-0.1	0.4	0.05	-0.13	-0.12	0.09
Average demand per node	-0.1	-0.22	-0.27	-0.26	0.42	0.38	-0.28	-0.31	-0.03	1	-0.28	-0.13	-0.14	0	-0.07	-0.07	-0.3	-0.29	0.05	-0.3	-0.09	-0.05	-0.15	-0.33	0.04	-0.23	-0.23	0.18
Number of pipes	-0.24	0.69	0.61	0.51	-0.26	-0.25	1	0.99	-0.2	-0.28	1	-0.18	-0.13	-0.32	-0.09	-0.11	0.77	0.77	-0.38	0.8	0.35	-0.15	-0.1	0.68	0.14	0.69	0.67	0.01
Average Residence Time	-0.13	-0.14	-0.18	-0.28	0.07	0.05	-0.18	-0.2	-0.09	-0.13	-0.18	1	0.11	-0.11	-0.11	-0.1	-0.24	-0.24	-0.1	-0.23	-0.05	0.48	-0.11	-0.28	0.06	-0.15	-0.15	-0.13
Node degree	0.83	-0.03	0.45	0.17	-0.01	-0.04	-0.14	-0.17	0.81	-0.14	-0.13	0.11	1	0.68	-0.31	-0.3	-0.38	-0.38	0.52	-0.36	-0.05	-0.01	-0.24	0.39	-0.24	-0.01	0	-0.09
Link density	0.79	-0.22	0.15	-0.08	0	-0.01	-0.32	-0.33	0.62	0	-0.32	-0.11	0.68	1	-0.23	-0.22	-0.54	-0.54	0.96	-0.52	-0.09	-0.27	-0.16	0.03	-0.71	-0.22	-0.22	0.01
Articulation points	-0.27	-0.15	-0.21	-0.07	-0.04	-0.05	-0.08	0.01	-0.15	-0.07	-0.09	-0.11	-0.31	-0.23	1	1	0.48	0.47	-0.27	0.48	-0.05	-0.16	0.97	-0.04	0.11	-0.16	-0.16	-0.19
Bridges	-0.26	-0.16	-0.22	-0.09	-0.05	-0.06	-0.1	-0.01	-0.14	-0.07	-0.11	-0.1	-0.3	-0.22	1	1	0.46	0.45	-0.25	0.45	-0.05	-0.17	0.97	-0.05	0.1	-0.17	-0.18	-0.18
Eccentricity	-0.48	0.48	0.36	0.45	-0.15	-0.15	0.78	0.83	-0.33	-0.3	0.77	-0.24	-0.38	-0.54	0.48	0.46	1	1	-0.62	1	0.26	-0.07	0.45	0.55	0.26	0.47	0.45	-0.16
Maximun eccentricity	[illegible]	[illegible]	0.37	0.47	-0.13	-0.12	0.78	0.83	-0.33	-0.29	[illegible]	[illegible]	[illegible]	-0.54	0.47	0.45	1	1	-0.62	0.99	0.26	-0.06	0.44	0.54	0.26	0.47	0.45	-0.16
Closeness eccentricity	[illegible]	[illegible]	-0.02	-0.21	0	0	-0.38	-0.4	0.44	0.05	[illegible]	[illegible]	[illegible]	0.96	-0.27	-0.25	-0.62	-0.62	1	-0.6	-0.11	-0.3	-0.19	-0.15	-0.76	-0.27	-0.27	0.08
Average Shortest path length	-0.46	0.51	0.38	0.44	-0.2	-0.2	0.81	0.85	-0.32	-0.3	0.8	-0.23	-0.36	-0.52	0.48	0.45	1	0.99	-0.6	1	0.27	-0.11	0.44	0.56	0.25	0.5	0.48	-0.14
Max pressure	-0.02	0.59	0.61	-0.12	-0.08	-0.08	0.35	0.35	-0.06	-0.09	0.35	-0.05	-0.05	-0.09	-0.05	-0.05	0.26	0.26	-0.11	0.27	1	-0.01	-0.03	0.21	0.04	0.61	0.67	-0.09
Pressure and specific pressure threshold 30m	-0.18	-0.18	-0.09	0.03	0.45	0.44	-0.15	-0.14	-0.16	-0.05	-0.15	0.48	-0.01	-0.27	-0.16	-0.17	-0.07	-0.06	-0.3	-0.11	-0.01	1	-0.17	-0.24	0.15	-0.16	-0.13	-0.2
Water service availability	-0.2	-0.14	-0.19	-0.08	-0.14	-0.15	-0.08	0.01	-0.1	-0.15	-0.1	-0.11	-0.24	-0.16	0.97	0.97	0.45	0.44	-0.19	0.44	-0.03	-0.17	1	0.04	0.07	-0.15	-0.15	-0.15
Expected demand	0.22	0.52	0.74	0.64	-0.36	-0.37	0.68	0.67	0.4	-0.33	0.68	-0.28	0.39	0.03	-0.04	-0.05	0.55	0.54	-0.15	0.56	0.21	-0.24	0.04	1	0.19	0.51	0.49	0.04
Todini	-0.32	0.1	0.14	0.19	-0.11	-0.09	0.14	0.15	0.05	0.04	0.14	0.06	-0.24	-0.71	0.11	0.1	0.26	0.26	-0.76	0.25	0.04	0.15	0.07	0.19	1	0.11	0.11	0.09
Network cost	-0.01	1	0.6	0.28	-0.19	-0.18	0.68	0.66	-0.13	-0.23	0.69	-0.15	-0.01	-0.22	-0.16	-0.17	0.47	0.47	-0.27	0.5	0.61	-0.16	-0.15	0.51	0.11	1	0.99	-0.03
Greenhouse gas emission	0.02	0.98	0.62	0.25	-0.17	-0.17	0.66	0.64	-0.12	-0.23	0.67	-0.15	0	-0.22	-0.16	-0.18	0.45	0.45	-0.27	0.48	0.67	-0.13	-0.15	0.49	0.11	0.99	1	-0.04
Pump energy	0.02	-0.02	-0.09	0.1	-0.16	-0.13	0.01	-0.01	0.09	0.18	0.01	-0.13	-0.09	0.01	-0.19	-0.18	-0.16	-0.16	0.08	-0.14	-0.09	-0.2	-0.15	0.04	0.09	-0.03	-0.04	1

Strong negative correlation = -0.6 – -0.8

Strong positive correlation = 0.6 – 0.8

Table 9: Matrix of the GH with strong correlation for many network

Correlation matrix

	Average Length	Total Length	Total Flow	Total demand	Min elevation	Max elevation	Number of nodes	Nodes demand <=0	Average demand	Average demand per node	Number of pipes	Average Residence Time	Node degree	Link density	Articulation points	Bridges	Eccentricity	Maximun eccentricity	Closeness eccentricity	Average Shortest path length	Max pressure	Pressure and specific pressure threshold 30m	Water service availability	Expected demand	Todini index, threshold	Network cost	Green-house gas emission	Pump energy
Average Length	1	-0.05	0.36	0.06	-0.06	-0.07	-0.24	-0.26	0.81	-0.1	-0.24	-0.13	0.83	0.79	-0.27	-0.26	-0.48	-0.48	0.67	-0.46	-0.02	-0.18	-0.2	0.22	-0.32	-0.01	0.02	0.02
Total Length	-0.05	1	0.58	0.28	-0.2	-0.19	0.68	0.66	-0.14	-0.22	0.69	-0.14	-0.03	-0.22	-0.15	-0.16	0.48	0.48	-0.27	0.51	0.59	-0.18	-0.14	0.52	0.1	1	0.98	-0.02
Total Flow	0.36	0.58	1	0.53	-0.06	-0.04	0.61	0.6	0.48	-0.27	0.61	-0.18	0.45	0.15	-0.21	-0.22	0.36	0.37	-0.02	0.38	0.61	-0.09	-0.19	0.74	0.14	0.6	0.62	-0.09
Total demand	0.06	0.28	0.53	1	0.25	0.3	0.51	0.51	0.25	-0.26	0.51	-0.28	0.17	-0.08	-0.07	-0.09	0.45	0.47	-0.21	0.44	-0.12	0.03	-0.08	0.64	0.19	0.28	0.25	0.1
Min elevation	-0.06	-0.2	-0.06	0.25	1	0.99	-0.26	-0.25	-0.06	0.42	-0.26	0.07	-0.01	0	-0.04	-0.05	-0.15	-0.13	0	-0.2	-0.08	0.45	-0.14	-0.36	-0.11	-0.19	-0.17	-0.16
Max elevation	-0.07	-0.19	-0.04	0.3	0.99	1	-0.25	-0.24	-0.06	0.38	-0.25	0.05	-0.04	-0.01	-0.05	-0.06	-0.15	-0.12	0	-0.2	-0.08	0.44	-0.15	-0.37	-0.09	-0.18	-0.17	-0.13
Number of nodes	-0.24	0.68	0.61	0.51	-0.26	-0.25	1	0.99	-0.2	-0.28	1	-0.18	-0.14	-0.32	-0.08	-0.1	0.78	0.78	-0.38	0.81	0.35	-0.15	-0.08	0.68	0.14	0.68	0.66	0.01
Nodes demand >= 0	-0.26	0.66	0.6	0.51	-0.25	-0.24	0.99	1	-0.21	-0.31	0.99	-0.2	-0.17	-0.33	0.01	-0.01	0.83	0.83	-0.4	0.85	0.35	-0.14	0.01	0.67	0.15	0.66	0.64	-0.01
Average demand	0.81	-0.14	0.48	0.25	-0.06	-0.06	-0.2	-0.21	1	-0.03	-0.2	-0.09	0.81	0.62	-0.15	-0.14	-0.33	-0.33	0.44	-0.32	-0.06	-0.16	-0.1	0.4	0.05	-0.13	-0.12	0.09
Average demand per node	-0.1	-0.22	-0.27	-0.26	0.42	0.38	-0.28	-0.31	-0.03	1	-0.28	-0.13	-0.14	0	-0.07	-0.07	-0.3	-0.29	0.05	-0.3	-0.09	-0.05	-0.15	-0.33	0.04	-0.23	-0.23	0.18
Number of pipes	-0.24	0.69	0.61	0.51	-0.26	-0.25	1	0.99	-0.2	-0.28	1	-0.18	-0.13	-0.32	-0.09	-0.11	0.77	0.77	-0.38	0.8	0.35	-0.15	-0.1	0.68	0.14	0.69	0.67	0.01
Average Residence Time	-0.13	-0.14	-0.18	-0.28	0.07	0.05	-0.18	-0.2	-0.09	-0.13	-0.18	1	0.11	-0.11	-0.11	-0.1	-0.24	-0.24	-0.1	-0.23	-0.05	0.48	-0.11	-0.28	0.06	-0.15	-0.15	-0.13
Node degree	0.83	-0.03	0.45	0.17	-0.01	-0.04	-0.14	-0.17	0.81	-0.14	-0.13	0.11	1	0.68	-0.31	-0.3	-0.38	-0.38	0.52	-0.36	-0.05	-0.01	-0.24	0.39	-0.24	-0.01	0	-0.09
Link density	0.79	-0.22	0.15	-0.08	0	-0.01	-0.32	-0.33	0.62	0	-0.32	-0.11	0.68	1	-0.23	-0.22	-0.54	-0.54	0.96	-0.52	-0.09	-0.27	-0.16	0.03	-0.71	-0.22	-0.22	0.01
Articulation points	-0.27	-0.15	-0.21	-0.07	-0.04	-0.05	-0.08	0.01	-0.15	-0.07	-0.09	-0.11	-0.31	-0.23	1	1	0.48	0.47	-0.27	0.48	-0.05	-0.16	0.97	-0.04	0.11	-0.16	-0.16	-0.19
Bridges	-0.26	-0.16	-0.22	-0.09	-0.05	-0.06	-0.1	-0.01	-0.14	-0.07	-0.11	-0.1	-0.3	-0.22	1	1	0.46	0.45	-0.25	0.45	-0.05	-0.17	0.97	-0.05	0.1	-0.17	-0.18	-0.18
Eccentricity	-0.48	0.48	0.36	0.45	-0.15	-0.15	0.78	0.83	-0.33	-0.3	0.77	-0.24	-0.38	-0.54	0.48	0.46	1	1	-0.62	1	0.26	-0.07	0.45	0.55	0.26	0.47	0.45	-0.16
Maximun eccentricity	-0.48	0.48	0.37	0.47	-0.13	-0.12	0.78	0.83	-0.33	-0.29	0.77	-0.24	-0.38	-0.54	0.47	0.45	1	1	-0.62	0.99	0.26	-0.06	0.44	0.54	0.26	0.47	0.45	-0.16
Closeness eccentricity	[illegible]	[illegible]	[illegible]	-0.21	0	0	-0.38	-0.4	0.44	0.05	-0.38	[illegible]	[illegible]	0.96	-0.27	-0.25	-0.62	-0.62	1	-0.6	-0.11	-0.3	-0.19	-0.15	-0.76	-0.27	-0.27	0.08
Average Shortest path	[illegible]	[illegible]	[illegible]	0.44	-0.2	-0.2	0.81	0.85	-0.32	-0.3	0.8	[illegible]	[illegible]	-0.52	0.48	0.45	1	0.99	-0.6	1	0.27	-0.11	0.44	0.56	0.25	0.5	0.48	-0.14
Max pressure	[illegible]	[illegible]	[illegible]	-0.12	-0.08	-0.08	0.35	0.35	-0.06	-0.09	0.35	[illegible]	[illegible]	-0.09	-0.05	-0.05	0.26	0.26	-0.11	0.27	1	-0.01	-0.03	0.21	0.04	0.61	0.67	-0.09
Pressure and specific pressure threshold 30m	-0.18	-0.18	-0.09	0.03	0.45	0.44	-0.15	-0.14	-0.16	-0.05	-0.15	0.48	-0.01	-0.27	-0.16	-0.17	-0.07	-0.06	-0.3	-0.11	-0.01	1	-0.17	-0.24	0.15	-0.16	-0.13	-0.2
Water service availability	-0.2	-0.14	-0.19	-0.08	-0.14	-0.15	-0.08	0.01	-0.1	-0.15	-0.1	-0.11	-0.24	-0.16	0.97	0.97	0.45	0.44	-0.19	0.44	-0.03	-0.17	1	0.04	0.07	-0.15	-0.15	-0.15
Expected demand	0.22	0.52	0.74	0.64	-0.36	-0.37	0.68	0.67	0.4	-0.33	0.68	-0.28	0.39	0.03	-0.04	-0.05	0.55	0.54	-0.15	0.56	0.21	-0.24	0.04	1	0.19	0.51	0.49	0.04
Todini	-0.32	0.1	0.14	0.19	-0.11	-0.09	0.14	0.15	0.05	0.04	0.14	0.06	-0.24	-0.71	0.11	0.1	0.26	0.26	-0.76	0.25	0.04	0.15	0.07	0.19	1	0.11	0.11	0.09
Network cost	-0.01	1	0.6	0.28	-0.19	-0.18	0.68	0.66	-0.13	-0.23	0.69	-0.15	-0.01	-0.22	-0.16	-0.17	0.47	0.47	-0.27	0.5	0.61	-0.16	-0.15	0.51	0.11	1	0.99	-0.03
Greenhouse gas emission	0.02	0.98	0.62	0.25	-0.17	-0.17	0.66	0.64	-0.12	-0.23	0.67	-0.15	0	-0.22	-0.16	-0.18	0.45	0.45	-0.27	0.48	0.67	-0.13	-0.15	0.49	0.11	0.99	1	-0.04
Pump energy	0.02	-0.02	-0.09	0.1	-0.16	-0.13	0.01	-0.01	0.09	0.18	0.01	-0.13	-0.09	0.01	-0.19	-0.18	-0.16	-0.16	0.08	-0.14	-0.09	-0.2	-0.15	0.04	0.09	-0.03	-0.04	1

Very strong negative correlation= -0.8 - -1

Very strong positive correlation= 0.8 - 1

Table 10: Matrix of the CHI with very strong correlation for many network

Correlation matrix

	Average Length	Total Length	Total Flow	Total demand	Min elevation	Max elevation	Number of nodes	Nodes demand <=0	Average demand	Average demand per node	Number of pipes	Average Residence Time	Node degree	Link density	Articulation points	Bridges	Eccentricity	Maximun eccentricity	Closeness eccentricity	Average Shortest path length	Max pressure	Pressure and specific pressure threshold 30m	Water service availability	Expected demand	Todini index, threshold	Network cost	Green-house gas emission	Pump energy
Average Length	1	-0.05	0.36	0.06	-0.06	-0.07	-0.24	-0.26	0.81	-0.1	-0.24	-0.13	0.83	0.79	-0.27	-0.26	-0.48	-0.48	0.67	-0.46	-0.02	-0.18	-0.2	0.22	-0.32	-0.01	0.02	0.02
Total Length	-0.05	1	0.58	0.28	-0.2	-0.19	0.68	0.66	-0.14	-0.22	0.69	-0.14	-0.03	-0.22	-0.15	-0.16	0.48	0.48	-0.27	0.51	0.59	-0.18	-0.14	0.52	0.1	1	0.98	-0.02
Total Flow	0.36	0.58	1	0.53	-0.06	-0.04	0.61	0.6	0.48	-0.27	0.61	-0.18	0.45	0.15	-0.21	-0.22	0.36	0.37	-0.02	0.38	0.61	-0.09	-0.19	0.74	0.14	0.6	0.62	-0.09
Total demand	0.06	0.28	0.53	1	0.25	0.3	0.51	0.51	0.25	-0.26	0.51	-0.28	0.17	-0.08	-0.07	-0.09	0.45	0.47	-0.21	0.44	-0.12	0.03	-0.08	0.64	0.19	0.28	0.25	0.1
Min elevation	-0.06	-0.2	-0.06	0.25	1	0.99	-0.26	-0.25	-0.06	0.42	-0.26	0.07	-0.01	0	-0.04	-0.05	-0.15	-0.13	0	-0.2	-0.08	0.45	-0.14	-0.36	-0.11	-0.19	-0.17	-0.16
Max elevation	-0.07	-0.19	-0.04	0.3	0.99	1	-0.25	-0.24	-0.06	0.38	-0.25	0.05	-0.04	-0.01	-0.05	-0.06	-0.15	-0.12	0	-0.2	-0.08	0.44	-0.15	-0.37	-0.09	-0.18	-0.17	-0.13
Number of nodes	-0.24	0.68	0.61	0.51	-0.26	-0.25	1	0.99	-0.2	-0.28	1	-0.18	-0.14	-0.32	-0.08	-0.1	0.78	0.78	-0.38	0.81	0.35	-0.15	-0.08	0.68	0.14	0.68	0.66	0.01
Nodes demand >= 0	-0.26	0.66	0.6	0.51	-0.25	-0.24	0.99	1	-0.21	-0.31	0.99	-0.2	-0.17	-0.33	0.01	-0.01	0.83	0.83	-0.4	0.85	0.35	-0.14	0.01	0.67	0.15	0.66	0.64	-0.01
Average demand	0.81	-0.14	0.48	0.25	-0.06	-0.06	-0.2	-0.21	1	-0.03	-0.2	-0.09	0.81	0.62	-0.15	-0.14	-0.33	-0.33	0.44	-0.32	-0.06	-0.16	-0.1	0.4	0.05	-0.13	-0.12	0.09
Average demand per node	-0.1	-0.22	-0.27	-0.26	0.42	0.38	-0.28	-0.31	-0.03	1	-0.28	-0.13	-0.14	0	-0.07	-0.07	-0.3	-0.29	0.05	-0.3	-0.09	-0.05	-0.15	-0.33	0.04	-0.23	-0.23	0.18
Number of pipes	-0.24	0.69	0.61	0.51	-0.26	-0.25	1	0.99	-0.2	-0.28	1	-0.18	-0.13	-0.32	-0.09	-0.11	0.77	0.77	-0.38	0.8	0.35	-0.15	-0.1	0.68	0.14	0.69	0.67	0.01
Average Residence Time	-0.13	-0.14	-0.18	-0.28	0.07	0.05	-0.18	-0.2	-0.09	-0.13	-0.18	1	0.11	-0.11	-0.11	-0.1	-0.24	-0.24	-0.1	-0.23	-0.05	0.48	-0.11	-0.28	0.06	-0.15	-0.15	-0.13
Node degree	0.83	-0.03	0.45	0.17	-0.01	-0.04	-0.14	-0.17	0.81	-0.14	-0.13	0.11	1	0.68	-0.31	-0.3	-0.38	-0.38	0.52	-0.36	-0.05	-0.01	-0.24	0.39	-0.24	-0.01	0	-0.09
Link density	0.79	-0.22	0.15	-0.08	0	-0.01	-0.32	-0.33	0.62	0	-0.32	-0.11	0.68	1	-0.23	-0.22	-0.54	-0.54	0.96	-0.52	-0.09	-0.27	-0.16	0.03	-0.71	-0.22	-0.22	0.01
Articulation points	-0.27	-0.15	-0.21	-0.07	-0.04	-0.05	-0.08	0.01	-0.15	-0.07	-0.09	-0.11	-0.31	-0.23	1	1	0.48	0.47	-0.27	0.48	-0.05	-0.16	0.97	-0.04	0.11	-0.16	-0.16	-0.19
Bridges	-0.26	-0.16	-0.22	-0.09	-0.05	-0.06	-0.1	-0.01	-0.14	-0.07	-0.11	-0.1	-0.3	-0.22	1	1	0.46	0.45	-0.25	0.45	-0.05	-0.17	0.97	-0.05	0.1	-0.17	-0.18	-0.18
Eccentricity	-0.48	0.48	0.36	0.45	-0.15	-0.15	0.78	0.83	-0.33	-0.3	0.77	-0.24	-0.38	-0.54	0.48	0.46	1	1	-0.62	1	0.26	-0.07	0.45	0.55	0.26	0.47	0.45	-0.16
Maximun eccentricity	-0.48	0.48	0.37	0.47	-0.13	-0.12	0.78	0.83	-0.33	-0.29	0.77	-0.24	-0.38	-0.54	0.47	0.45	1	1	-0.62	0.99	0.26	-0.06	0.44	0.54	0.26	0.47	0.45	-0.16
Closeness eccentricity	[illegible]	-0.27	-0.02	-0.21	0	0	-0.38	-0.4	0.44	0.05	-0.38	-0.1	0.52	0.96	-0.27	-0.25	-0.62	-0.62	1	-0.6	-0.11	-0.3	-0.19	-0.15	-0.76	-0.27	-0.27	0.08
Average Shortest path length	[illegible]	0.51	0.38	0.44	-0.2	-0.2	0.81	0.85	-0.32	-0.3	0.8	-0.23	-0.36	-0.52	0.48	0.45	1	0.99	-0.6	1	0.27	-0.11	0.44	0.56	0.25	0.5	0.48	-0.14
Max pressure	[illegible]	0.59	0.61	-0.12	-0.08	-0.08	0.35	0.35	-0.06	-0.09	0.35	-0.05	-0.05	-0.09	-0.05	-0.05	0.26	0.26	-0.11	0.27	1	-0.01	-0.03	0.21	0.04	0.61	0.67	-0.09
Pressure and specific pressure threshold 30m	-0.18	-0.18	-0.09	0.03	0.45	0.44	-0.15	-0.14	-0.16	-0.05	-0.15	0.48	-0.01	-0.27	-0.16	-0.17	-0.07	-0.06	-0.3	-0.11	-0.01	1	-0.17	-0.24	0.15	-0.16	-0.13	-0.2
Water service availability	-0.2	-0.14	-0.19	-0.08	-0.14	-0.15	-0.08	0.01	-0.1	-0.15	-0.1	-0.11	-0.24	-0.16	0.97	0.97	0.45	0.44	-0.19	0.44	-0.03	-0.17	1	0.04	0.07	-0.15	-0.15	-0.15
Expected demand	0.22	0.52	0.74	0.64	-0.36	-0.37	0.68	0.67	0.4	-0.33	0.68	-0.28	0.39	0.03	-0.04	-0.05	0.55	0.54	-0.15	0.56	0.21	-0.24	0.04	1	0.19	0.51	0.49	0.04
Todini	-0.32	0.1	0.14	0.19	-0.11	-0.09	0.14	0.15	0.05	0.04	0.14	0.06	-0.24	-0.71	0.11	0.1	0.26	0.26	-0.76	0.25	0.04	0.15	0.07	0.19	1	0.11	0.11	0.09
Network cost	-0.01	1	0.6	0.28	-0.19	-0.18	0.68	0.66	-0.13	-0.23	0.69	-0.15	-0.01	-0.22	-0.16	-0.17	0.47	0.47	-0.27	0.5	0.61	-0.16	-0.15	0.51	0.11	1	0.99	-0.03
Greenhouse gas emission	0.02	0.98	0.62	0.25	-0.17	-0.17	0.66	0.64	-0.12	-0.23	0.67	-0.15	0	-0.22	-0.16	-0.18	0.45	0.45	-0.27	0.48	0.67	-0.13	-0.15	0.49	0.11	0.99	1	-0.04
Pump energy	0.02	-0.02	-0.09	0.1	-0.16	-0.13	0.01	-0.01	0.09	0.18	0.01	-0.13	-0.09	0.01	-0.19	-0.18	-0.16	-0.16	0.08	-0.14	-0.09	-0.2	-0.15	0.04	0.09	-0.03	-0.04	1

Perfect positive correlation

Table 11: Matrix of the GH with perfect correlation for many network

5.2 Characterization of Den Haag

Descriptive analysis

As a result of the characterisation of the 600 networks, the box whisker plot is shown in figure 23; this graph is represented in the form of boxes drawing for each CH the minimum medium and maximum values for all the model networks. In addition, table 10 shows the ranges of each characteristic. It should be noted that the numbering of the box whisker plot corresponds to the numbered characteristics in table 10.

Figure 25: box whisker plot of different model networks

	Characteristic	units	minimum	median	maximum
1	Average Length	m	78.94	82.59	118.74
2	Total Length	m	772410.7	784243.56	787305.88
3	Total Flow	CMH	106.74	134.86	173.13
4	Total demand	CMH	4.19	4.19	5.08
5	Min elevation	m	-7	-7	-6.15
6	Max elevation	m	14.78	14.78	14.84
7	Number of nodes	units	4552	7505	7949
8	Number of nodes demand less than or equal 0	units	837	1388	1487
9	Average demand	units	0	0	0
10	Average demand per node	units	0	0	0.01
11	Number of pipes	units	6505	9501	9958
12	Number of tanks	units	2	2	2
13	Number of reservoir	units	1	1	1
14	Number of pumps	units	7	7	7
15	Number of valves	units	16	16	16
16	Number of patterns	units	13	13	13
17	Valves segmentation	units	80	80	80
18	Average Residence Time	units	1.31E+09	2301722273	2650827616
19	Node degree	degree	2.51	2.54	2.87
20	Link density	units	0	0	0
21	Articulation points	units	36	51	87
22	Bridges	units	33	48	86
23	Eccentricity	m	76.03	83.3	104.34
24	Maximun eccentricity	m	99	111	151
25	Closeness eccentricity	m	0	0	0
26	Average Shortest path length	m	38.4	43.98	46.71
27	Max pressure	m	23.18	461.38	90651.75
28	Pressure and specific pressure threshold 30m	m	0	2.58	8.74
29	Water service availability	CMH	[illegible]	1391	1490
30	Expected demand	CMH	1.23	1.23	1.23
31	Todini index, threshold	hours	-58.16	74.58	2187.32
32	Entropy	units	2	2	2
33	Water age	hour	0	0	0
34	Water age at 48 hours	hours	0	0	0
35	Population that is impacted by water age greater than 24 hours	hours	4550	7503	7947
36	Network cost	units	10398191	11917199.4	12331817.01
37	Greenhouse gas emission	C02	12178951	15183230.3	16002992.08
38	Pump energy	C02	968502.8	968531.62	968537.78

Table 13: Range values of different networks

The implementation of the gaussian method of each one of the CH gives a result of three main graphs of the normal distribution in the data set.

Low distribution High distribution Linear values

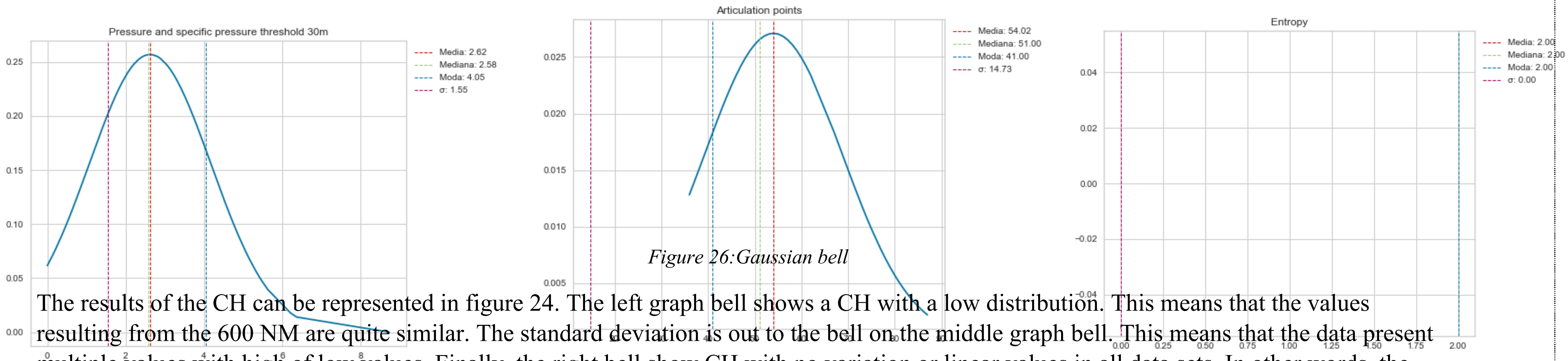


Figure 26:Gaussian bell

The results of the CH can be represented in figure 24. The left graph bell shows a CH with a low distribution. This means that the values resulting from the 600 NM are quite similar. The standard deviation is out to the bell on the middle graph bell. This means that the data present multiple values with high of low values. Finally, the right bell show CH with no variation or linear values in all data sets. In other words, the changes made to the network model do not affect these CHs.

The CH with the similarly left bell is Pressure and specific pressure threshold, max pressure, todini.

The CH with the similar middle bell is: Articulation points, Average demand per node, Average length, average resident time, average shortest path length, bridges, closeness eccentricity, eccentricity, greenhouse emission, link density, network cost, node degree, number of nodes demand less than or equal to 0, number of nodes, number of pipes, population that is impacted by water age greater than 24 hours, total demand, total flow, water service availability.

The CH with the similar right bell is: Entropy, expected demand, max elevation, min elevation, number of patterns, number of pumps, number of reservoirs, number of tanks, pump energy, total length, water age.

Correlation Analysis

As a result of the correlation analysis, the following tables represent the correlation of the CH classify and show in six tables. The first table shows the variables with no correlation. This corresponds to the data between -0.2 to 0 and 0 to 0.2.

Correlation matrix

	Average Length	Total Length	Total Flow	Total demand	Min elevation	Max elevation	Number of nodes	Nodes demand <=0	Average demand	Average demand per node	Number of pipes	Average Residence Time	Node degree	Link density	Articulation points	Bridges	Eccentricity	Maximun eccentricity	Closeness eccentricity	Average Shortest path length	Max pressure	Pressure and specific pressure threshold 30m	Water service availability	Expected demand	Todini index, threshold	Network cost	Green-house gas emission	Pump energy
Average Length	1	-0.95	0.36	0.83	0.35	-0.43	-0.99	-0.93	0.97	0.88	-0.99	-0.95	1	1	0.9	0.9	0.84	0.85	0.99	0.34	-0.19	0.22	-0.93	0.05	-0.06	-0.97	-0.96	0.23
Total Length	-0.95	1	-0.37	-0.86	-0.33	0.41	0.94	0.91	-0.95	-0.89	0.95	0.91	-0.95	-0.96	-0.84	-0.85	-0.75	-0.77	-0.95	-0.27	0.19	-0.13	0.91	-0.06	0.11	0.92	0.91	-0.24
Total Flow	0.36	-0.37	1	0.52	0.48	-0.07	-0.32	-0.35	0.44	0.43	-0.32	-0.28	0.37	0.38	0.29	0.34	0.18	0.25	0.38	-0.2	0.11	-0.16	-0.35	0.01	0.02	-0.27	-0.26	-0.03
Total demand	0.83	-0.86	0.52	1	0.32	-0.22	-0.78	-0.89	0.93	0.95	-0.78	-0.73	0.84	0.85	0.61	0.64	0.49	0.52	0.87	-0.09	-0.1	-0.18	-0.89	0.05	-0.12	-0.72	-0.7	0.22
Min elevation	0.35	-0.33	0.48	0.32	1	-0.1	-0.33	-0.22	0.35	0.27	-0.33	-0.27	0.35	0.35	0.48	0.49	0.37	0.41	0.36	0.01	0.12	0.05	-0.22	0.04	0.16	-0.28	-0.28	0.28
Max elevation	-0.43	0.41	-0.07	-0.22	-0.1	1	0.46	0.39	-0.36	-0.28	0.46	0.47	-0.42	-0.41	-0.44	-0.45	-0.48	-0.51	-0.4	-0.39	0.2	-0.3	0.39	0.03	0.17	0.5	0.5	-0.17
Number of nodes	-0.99	0.94	-0.32	-0.78	-0.33	0.46	1	0.9	-0.95	-0.84	1	0.97	-0.99	-0.98	-0.92	-0.92	-0.88	-0.89	-0.98	-0.43	0.21	-0.29	0.9	-0.05	0.05	0.99	0.98	-0.22
Nodes demand >= 0	-0.93	0.91	-0.35	-0.89	-0.22	0.39	0.9	1	-0.96	-0.97	0.9	0.81	-0.94	-0.94	-0.7	-0.72	-0.61	-0.63	-0.95	-0.05	0.23	0.02	1	-0.04	0.13	0.84	0.83	-0.2

No negative correlation = -0.2 – 0 No positive correlation = 0 - 0.2

Table 13: Matrix of the CH with no correlation for different model networks

The second table shows the CH with a weak correlation. This correlation corresponds to the CH with values from -0.4 to -0.2 and 0.2 to 0.4.

Correlation matrix

	Average Length	Total Length	Total Flow	Total demand	Min elevation	Max elevation	Number of nodes	Nodes demand <=0	Average demand	Average demand per node	Number of pipes	Average Residence Time	Node degree	Link density	Articulation points	Bridges	Eccentricity	Maximun eccentricity	Closeness eccentricity	Average Shortest path length	Max pressure	Pressure and specific pressure threshold 30m	Water service availability	Expected demand	Todini index, threshold	Network cost	Green-house gas emission	Pump energy
Average Length	1	-0.95	0.36	0.83	0.35	-0.43	-0.99	-0.93	0.97	0.88	-0.99	-0.95	1	1	0.9	0.9	0.84	0.85	0.99	0.34	-0.19	0.22	-0.93	0.05	-0.06	-0.97	-0.96	0.23
Total Length	-0.95	1	-0.37	-0.86	-0.33	0.41	0.94	0.91	-0.95	-0.89	0.95	0.91	-0.95	-0.96	-0.84	-0.85	-0.75	-0.77	-0.95	-0.27	0.19	-0.13	0.91	-0.06	0.11	0.92	0.91	-0.24
Total Flow	0.36	-0.37	1	0.52	0.48	-0.07	-0.32	-0.35	0.44	0.43	-0.32	-0.28	0.37	0.38	0.29	0.34	0.18	0.25	0.38	-0.2	0.11	-0.16	-0.35	0.01	0.02	-0.27	-0.26	-0.03
Total demand	0.83	-0.86	0.52	1	0.32	-0.22	-0.78	-0.89	0.93	0.95	-0.78	-0.73	0.84	0.85	0.61	0.64	0.49	0.52	0.87	-0.09	-0.1	-0.18	-0.89	0.05	-0.12	-0.72	-0.7	0.22
Min elevation	0.35	-0.33	0.48	0.32	1	-0.1	-0.33	-0.22	0.35	0.27	-0.33	-0.27	0.35	0.35	0.48	0.49	0.37	0.41	0.36	0.01	0.12	0.05	-0.22	0.04	0.16	-0.28	-0.28	0.28
Max elevation	-0.43	0.41	-0.07	-0.22	-0.1	1	0.46	0.39	-0.36	-0.28	0.46	0.47	-0.42	-0.41	-0.44	-0.45	-0.48	-0.51	-0.4	-0.39	0.2	-0.3	0.39	0.03	0.17	0.5	0.5	-0.17
Number of nodes	-0.99	0.94	-0.32	-0.78	-0.33	0.46	1	0.9	-0.95	-0.84	1	0.97	-0.99	-0.98	-0.92	-0.92	-0.88	-0.89	-0.98	-0.43	0.21	-0.29	0.9	-0.05	0.05	0.99	0.98	-0.22
Nodes demand >= 0	-0.93	0.91	-0.35	-0.89	-0.22	0.39	0.9	1	-0.96	-0.97	0.9	0.81	-0.94	-0.94	-0.7	-0.72	-0.61	-0.63	-0.95	-0.05	0.23	0.02	1	-0.04	0.13	0.84	0.83	-0.2
Average demand	0.97	-0.95	0.44	0.93	0.35	-0.36	-0.95	-0.96	1	0.96	-0.95	-0.89	0.98	0.98	0.81	0.82	0.72	0.74	0.99	0.16	-0.16	0.05	-0.96	0.06	-0.09	-0.9	-0.89	0.24
Average demand per node	0.88	-0.89	0.43	0.95	0.27	-0.28	-0.84	-0.97	0.96	1	-0.84	-0.74	0.9	0.91	0.64	0.66	0.51	0.54	0.93	-0.09	-0.15	-0.14	-0.97	0.04	-0.13	-0.76	-0.75	0.21
Number of pipes	-0.99	0.95	-0.32	-0.78	-0.33	0.46	1	0.9	-0.95	-0.84	1	0.97	-0.99	-0.98	-0.92	-0.92	-0.88	-0.89	-0.98	-0.43	0.21	-0.29	0.9	-0.05	0.05	0.99	0.99	-0.22
Average Residence Time	-0.95	0.91	-0.28	-0.73	-0.27	0.47	0.97	0.81	-0.89	-0.74	0.97	1	-0.94	-0.93	-0.91	-0.91	-0.91	-0.91	-0.91	-0.56	0.22	-0.36	0.81	-0.06	0.05	0.99	0.98	-0.21
Node degree	1	-0.95	0.37	0.84	0.35	-0.42	-0.99	-0.94	0.98	0.9	-0.99	-0.94	1	1	0.88	0.89	0.82	0.84	1	0.31	-0.19	0.19	-0.94	0.05	-0.06	-0.96	-0.95	0.23
Link density	1	-0.96	0.38	0.85	0.35	-0.41	-0.98	-0.94	0.98	0.91	-0.98	-0.93	1	1	0.87	0.88	0.8	0.82	1	0.28	-0.18	0.17	-0.94	0.05	-0.06	-0.95	-0.95	0.23
Articulation points	0.9	-0.84	0.29	0.61	0.48	-0.44	-0.92	-0.7	0.81	0.64	-0.92	-0.91	0.88	0.87	1	1	0.94	0.95	0.86	0.55	-0.13	0.43	-0.7	0.05	0.09	-0.93	-0.93	0.35
Bridges	0.9	-0.85	0.34	0.64	0.49	-0.45	-0.92	-0.72	0.82	0.66	-0.92	-0.91	0.89	0.88	1	1	0.93	0.94	0.87	0.52	-0.13	0.4	-0.72	0.06	0.09	-0.92	-0.92	0.35
Eccentricity	0.84	-0.75	0.18	0.49	0.37	-0.48	-0.88	-0.61	0.72	0.51	-0.88	-0.91	0.82	0.8	0.94	0.93	1	0.99	0.78	0.72	-0.18	0.54	-0.61	0.04	0.06	-0.91	-0.91	0.27
Maximun eccentricity	0.85	-0.77	0.25	0.52	0.41	-0.51	-0.89	-0.63	0.74	0.54	-0.89	-0.91	0.84	0.82	0.95	0.94	0.99	1	0.8	0.68	-0.15	0.54	-0.63	0.04	0.07	-0.92	-0.92	0.29
Closeness eccer[illegible]	[illegible]	[illegible]	0.38	0.87	0.36	-0.4	-0.98	-0.95	0.99	0.93	[illegible]	[illegible]	[illegible]	1	0.86	0.87	0.78	0.8	1	0.24	-0.18	0.13	-0.95	0.05	-0.06	-0.94	-0.93	0.24
Average Shortest pa[illegible]	[illegible]	[illegible]	-0.2	-0.09	0.01	-0.39	-0.43	-0.05	0.16	-0.09	[illegible]	[illegible]	[illegible]	0.28	0.55	-0.52	0.72	0.68	0.24	1	-0.11	0.77	-0.05	0.03	0.06	-0.54	-0.55	0.06
Max pressur[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.16	-0.15	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.15	[illegible]	-0.11	1	0.01	0.23	0.05	0.7	0.22	0.22	0.13
Pressure and specific pressure threshold 30m	0.22	-0.13	-0.16	-0.18	0.05	-0.3	-0.29	0.02	0.05	-0.14	-0.29	-0.36	0.19	0.17	0.43	0.4	0.54	0.54	0.13	0.77	0.01	1	0.02	0.05	0.2	-0.38	-0.4	0.15
Water service availability	-0.93	0.91	-0.35	-0.89	-0.22	0.39	0.9	1	-0.96	-0.97	0.9	0.81	-0.94	-0.94	-0.7	-0.72	-0.61	-0.63	-0.95	-0.05	0.23	0.02	1	-0.04	0.13	0.84	0.83	-0.2
Expected demand	0.05	-0.06	0.01	0.05	0.04	0.03	-0.05	-0.04	0.06	0.04	-0.05	-0.06	0.05	0.05	0.05	0.06	0.04	0.04	0.05	0.03	0.05	0.05	-0.04	1	0.02	-0.05	-0.05	0.04
Todini	-0.06	0.11	0.02	-0.12	0.16	0.17	0.05	0.13	-0.09	-0.13	0.05	0.05	-0.06	-0.06	0.09	0.09	0.06	0.07	-0.06	0.06	0.7	0.2	0.13	0.02	1	0.04	0.04	0.13
Network cost	-0.97	0.92	-0.27	-0.72	-0.28	0.5	0.99	0.84	-0.9	-0.76	0.99	0.99	-0.96	-0.95	-0.93	-0.92	-0.91	-0.92	-0.94	-0.54	0.22	-0.38	0.84	-0.05	0.04	1	1	-0.22
Greenhouse gas emission	-0.96	0.91	-0.26	-0.7	-0.28	0.5	0.98	0.83	-0.89	-0.75	0.99	0.98	-0.95	-0.95	-0.93	-0.92	-0.91	-0.92	-0.93	-0.55	0.22	-0.4	0.83	-0.05	0.04	1	1	-0.22
Pump energy	0.23	-0.24	-0.03	0.22	0.28	-0.17	-0.22	-0.2	0.24	0.21	-0.22	-0.21	0.23	0.23	0.35	0.35	0.27	0.29	0.24	0.06	0.13	0.15	-0.2	0.04	0.13	-0.22	-0.22	1

Weak negative correlation = -0.2 - -0.4

Weak positive correlation = 0.2 - 0.4

Table 14: Matrix of the CH with weak correlation for different model networks

The third table presents the CH with a moderate correlation. It corresponds to values from -0.6 to -0.4 and 0.4 to 0.6.

Correlation matrix

	Average Length	Total Length	Total Flow	Total demand	Min elevation	Max elevation	Number of nodes	Nodes demand <=0	Average demand	Average demand per node	Number of pipes	Average Residence Time	Node degree	Link density	Articulation points	Bridges	Eccentricity	Maximun eccentricity	Closeness eccentricity	Average Shortest path length	Max pressure	Pressure and specific pressure threshold 30m	Water service availability	Expected demand	Todini index, threshold	Network cost	Green-house gas emission	Pump energy
Average Length	1	-0.95	0.36	0.83	0.35	-0.43	-0.99	-0.93	0.97	0.88	-0.99	-0.95	1	1	0.9	0.9	0.84	0.85	0.99	0.34	-0.19	0.22	-0.93	0.05	-0.06	-0.97	-0.96	0.23
Total Length	-0.95	1	-0.37	-0.86	-0.33	0.41	0.94	0.91	-0.95	-0.89	0.95	0.91	-0.95	-0.96	-0.84	-0.85	-0.75	-0.77	-0.95	-0.27	0.19	-0.13	0.91	-0.06	0.11	0.92	0.91	-0.24
Total Flow	0.36	-0.37	1	0.52	0.48	-0.07	-0.32	-0.35	0.44	0.43	-0.32	-0.28	0.37	0.38	0.29	0.34	0.18	0.25	0.38	-0.2	0.11	-0.16	-0.35	0.01	0.02	-0.27	-0.26	-0.03
Total demand	0.83	-0.86	0.52	1	0.32	-0.22	-0.78	-0.89	0.93	0.95	-0.78	-0.73	0.84	0.85	0.61	0.64	0.49	0.52	0.87	-0.09	-0.1	-0.18	-0.89	0.05	-0.12	-0.72	-0.7	0.22
Min elevation	0.35	-0.33	0.48	0.32	1	-0.1	-0.33	-0.22	0.35	0.27	-0.33	-0.27	0.35	0.35	0.48	0.49	0.37	0.41	0.36	0.01	0.12	0.05	-0.22	0.04	0.16	-0.28	-0.28	0.28
Max elevation	-0.43	0.41	-0.07	-0.22	-0.1	1	0.46	0.39	-0.36	-0.28	0.46	0.47	-0.42	-0.41	-0.44	-0.45	-0.48	-0.51	-0.4	-0.39	0.2	-0.3	0.39	0.03	0.17	0.5	0.5	-0.17
Number of nodes	-0.99	0.94	-0.32	-0.78	-0.33	0.46	1	0.9	-0.95	-0.84	1	0.97	-0.99	-0.98	-0.92	-0.92	-0.88	-0.89	-0.98	-0.43	0.21	-0.29	0.9	-0.05	0.05	0.99	0.98	-0.22
Nodes demand >= 0	-0.93	0.91	-0.35	-0.89	-0.22	0.39	0.9	1	-0.96	-0.97	0.9	0.81	-0.94	-0.94	-0.7	-0.72	-0.61	-0.63	-0.95	-0.05	0.23	0.02	1	-0.04	0.13	0.84	0.83	-0.2
Average demand	0.97	-0.95	0.44	0.93	0.35	-0.36	-0.95	-0.96	1	0.96	-0.95	-0.89	0.98	0.98	0.81	0.82	0.72	0.74	0.99	0.16	-0.16	0.05	-0.96	0.06	-0.09	-0.9	-0.89	0.24
Average demand per node	0.88	-0.89	0.43	0.95	0.27	-0.28	-0.84	-0.97	0.96	1	-0.84	-0.74	0.9	0.91	0.64	0.66	0.51	0.54	0.93	-0.09	-0.15	-0.14	-0.97	0.04	-0.13	-0.76	-0.75	0.21
Number of pipes	-0.99	0.95	-0.32	-0.78	-0.33	0.46	1	0.9	-0.95	-0.84	1	0.97	-0.99	-0.98	-0.92	-0.92	-0.88	-0.89	-0.98	-0.43	0.21	-0.29	0.9	-0.05	0.05	0.99	0.99	-0.22
Average Residence Time	-0.95	0.91	-0.28	-0.73	-0.27	0.47	0.97	0.81	-0.89	-0.74	0.97	1	-0.94	-0.93	-0.91	-0.91	-0.91	-0.91	-0.91	-0.56	0.22	-0.36	0.81	-0.06	0.05	0.99	0.98	-0.21
Node degree	1	-0.95	0.37	0.84	0.35	-0.42	-0.99	-0.94	0.98	0.9	-0.99	-0.94	1	1	0.88	0.89	0.82	0.84	1	0.31	-0.19	0.19	-0.94	0.05	-0.06	-0.96	-0.95	0.23
Link density	1	-0.96	0.38	0.85	0.35	-0.41	-0.98	-0.94	0.98	0.91	-0.98	-0.93	1	1	0.87	0.88	0.8	0.82	1	0.28	-0.18	0.17	-0.94	0.05	-0.06	-0.95	-0.95	0.23
Articulation points	0.9	-0.84	0.29	0.61	0.48	-0.44	-0.92	-0.7	0.81	0.64	-0.92	-0.91	0.88	0.87	1	1	0.94	0.95	0.86	0.55	-0.13	0.43	-0.7	0.05	0.09	-0.93	-0.93	0.35
Bridges	0.9	-0.85	0.34	0.64	0.49	-0.45	-0.92	-0.72	0.82	0.66	-0.92	-0.91	0.89	0.88	1	1	0.93	0.94	0.87	0.52	-0.13	0.4	-0.72	0.06	0.09	-0.92	-0.92	0.35
Eccentricity	0.84	-0.75	0.18	0.49	0.37	-0.48	-0.88	-0.61	0.72	0.51	-0.88	-0.91	0.82	0.8	0.94	0.93	1	0.99	0.78	0.72	-0.18	0.54	-0.61	0.04	0.06	-0.91	-0.91	0.27

Moderate negative correlation = -0.4 – -0.6 Moderate positive correlation = 0.4 – 0.6

Table 15: Matrix of the CH with moderate correlation for different model networks

The fourth table is showing the variables with strong correlation. The values of this correlation start from -0.6 to -0.8 and 0.6 to 0.8.

Correlation matrix

	Average Length	Total Length	Total Flow	Total demand	Min elevation	Max elevation	Number of nodes	Nodes demand <=0	Average demand	Average demand per node	Number of pipes	Average Residence Time	Node degree	Link density	Articulation points	Bridges	Eccentricity	Maximun eccentricity	Closeness eccentricity	Average Shortest path length	Max pressure	Pressure and specific pressure threshold 30m	Water service availability	Expected demand	Todini index, threshold	Network cost	Green-house gas emission	Pump energy
Average Length	1	-0.95	0.36	0.83	0.35	-0.43	-0.99	-0.93	0.97	0.88	-0.99	-0.95	1	1	0.9	0.9	0.84	0.85	0.99	0.34	-0.19	0.22	-0.93	0.05	-0.06	-0.97	-0.96	0.23
Total Length	-0.95	1	-0.37	-0.86	-0.33	0.41	0.94	0.91	-0.95	-0.89	0.95	0.91	-0.95	-0.96	-0.84	-0.85	-0.75	-0.77	-0.95	-0.27	0.19	-0.13	0.91	-0.06	0.11	0.92	0.91	-0.24
Total Flow	0.36	-0.37	1	0.52	0.48	-0.07	-0.32	-0.35	0.44	0.43	-0.32	-0.28	0.37	0.38	0.29	0.34	0.18	0.25	0.38	-0.2	0.11	-0.16	-0.35	0.01	0.02	-0.27	-0.26	-0.03
Total demand	0.83	-0.86	0.52	1	0.32	-0.22	-0.78	-0.89	0.93	0.95	-0.78	-0.73	0.84	0.85	0.61	0.64	0.49	0.52	0.87	-0.09	-0.1	-0.18	-0.89	0.05	-0.12	-0.72	-0.7	0.22
Min elevation	0.35	-0.33	0.48	0.32	1	-0.1	-0.33	-0.22	0.35	0.27	-0.33	-0.27	0.35	0.35	0.48	0.49	0.37	0.41	0.36	0.01	0.12	0.05	-0.22	0.04	0.16	-0.28	-0.28	0.28
Max elevation	-0.43	0.41	-0.07	-0.22	-0.1	1	0.46	0.39	-0.36	-0.28	0.46	0.47	-0.42	-0.41	-0.44	-0.45	-0.48	-0.51	-0.4	-0.39	0.2	-0.3	0.39	0.03	0.17	0.5	0.5	-0.17
Number of nodes	-0.99	0.94	-0.32	-0.78	-0.33	0.46	1	0.9	-0.95	-0.84	1	0.97	-0.99	-0.98	-0.92	-0.92	-0.88	-0.89	-0.98	-0.43	0.21	-0.29	0.9	-0.05	0.05	0.99	0.98	-0.22
Nodes demand >= 0	-0.93	0.91	-0.35	-0.89	-0.22	0.39	0.9	1	-0.96	-0.97	0.9	0.81	-0.94	-0.94	-0.7	-0.72	-0.61	-0.63	-0.95	-0.05	0.23	0.02	1	-0.04	0.13	0.84	0.83	-0.2
Average demand	0.97	-0.95	0.44	0.93	0.35	-0.36	-0.95	-0.96	1	0.96	-0.95	-0.89	0.98	0.98	0.81	0.82	0.72	0.74	0.99	0.16	-0.16	0.05	-0.96	0.06	-0.09	-0.9	-0.89	0.24
Average demand per node	0.88	-0.89	0.43	0.95	0.27	-0.28	-0.84	-0.97	0.96	1	-0.84	-0.74	0.9	0.91	0.64	0.66	0.51	0.54	0.93	-0.09	-0.15	-0.14	-0.97	0.04	-0.13	-0.76	-0.75	0.21
Number of pipes	-0.99	0.95	-0.32	-0.78	-0.33	0.46	1	0.9	-0.95	-0.84	1	0.97	-0.99	-0.98	-0.92	-0.92	-0.88	-0.89	-0.98	-0.43	0.21	-0.29	0.9	-0.05	0.05	0.99	0.99	-0.22
Average Residence Time	-0.95	0.91	-0.28	-0.73	-0.27	0.47	0.97	0.81	-0.89	-0.74	0.97	1	-0.94	-0.93	-0.91	-0.91	-0.91	-0.91	-0.91	-0.56	0.22	-0.36	0.81	-0.06	0.05	0.99	0.98	-0.21
Node degree	1	-0.95	0.37	0.84	0.35	-0.42	-0.99	-0.94	0.98	0.9	-0.99	-0.94	1	1	0.88	0.89	0.82	0.84	1	0.31	-0.19	0.19	-0.94	0.05	-0.06	-0.96	-0.95	0.23
Link density	1	-0.96	0.38	0.85	0.35	-0.41	-0.98	-0.94	0.98	0.91	-0.98	-0.93	1	1	0.87	0.88	0.8	0.82	1	0.28	-0.18	0.17	-0.94	0.05	-0.06	-0.95	-0.95	0.23
Articulation points	0.9	-0.84	0.29	0.61	0.48	-0.44	-0.92	-0.7	0.81	0.64	-0.92	-0.91	0.88	0.87	1	1	0.94	0.95	0.86	0.55	-0.13	0.43	-0.7	0.05	0.09	-0.93	-0.93	0.35
Bridges	0.9	-0.85	0.34	0.64	0.49	-0.45	-0.92	-0.72	0.82	0.66	-0.92	-0.91	0.89	0.88	1	1	0.93	0.94	0.87	0.52	-0.13	0.4	-0.72	0.06	0.09	-0.92	-0.92	0.35
Eccentricity	0.84	-0.75	0.18	0.49	0.37	-0.48	-0.88	-0.61	0.72	0.51	-0.88	-0.91	0.82	0.8	0.94	0.93	1	0.99	0.78	0.72	-0.18	0.54	-0.61	0.04	0.06	-0.91	-0.91	0.27
Maximun eccentricity	0.85	-0.77	0.25	0.52	0.41	-0.51	-0.89	-0.63	0.74	0.54	-0.89	-0.91	0.84	0.82	0.95	0.94	0.99	1	0.8	0.68	-0.15	0.54	-0.63	0.04	0.07	-0.92	-0.92	0.29
Closeness eccentricity	0.99	-0.95	0.38	0.87	0.36	-0.4	-0.98	-0.95	0.99	0.93	-0.98	-0.91	1	1	0.86	0.87	0.78	0.8	1	0.24	-0.18	0.13	-0.95	0.05	-0.06	-0.94	-0.93	0.24
Average Shortest pat[illegible]	[illegible]	[illegible]	-0.2	-0.09	0.01	-0.39	-0.43	-0.05	0.16	-0.09	[illegible]	[illegible]	[illegible]	0.28	0.55	0.52	0.72	0.68	0.24	1	-0.11	0.77	-0.05	0.03	0.06	-0.54	-0.55	0.06
Max pressure	[illegible]	[illegible]	0.11	-0.1	0.12	0.2	0.21	0.23	-0.16	-0.15	[illegible]	[illegible]	[illegible]	-0.18	-0.13	-0.13	-0.18	-0.15	-0.18	-0.11	1	0.01	0.23	0.05	0.7	0.22	0.22	0.13
Pressure and specific [illegible] threshold 30m	0.22	-0.13	-0.16	-0.18	0.05	-0.3	-0.29	0.02	0.05	-0.14	-0.29	-0.36	0.19	0.17	0.43	0.4	0.54	0.54	0.13	0.77	0.01	1	0.02	0.05	0.2	-0.38	-0.4	0.15
Water service availability	-0.93	0.91	-0.35	-0.89	-0.22	0.39	0.9	1	-0.96	-0.97	0.9	0.81	-0.94	-0.94	-0.7	-0.72	-0.61	-0.63	-0.95	-0.05	0.23	0.02	1	-0.04	0.13	0.84	0.83	-0.2
Expected demand	0.05	-0.06	0.01	0.05	0.04	0.03	-0.05	-0.04	0.06	0.04	-0.05	-0.06	0.05	0.05	0.05	0.06	0.04	0.04	0.05	0.03	0.05	0.05	-0.04	1	0.02	-0.05	-0.05	0.04
Todini	-0.06	0.11	0.02	-0.12	0.16	0.17	0.05	0.13	-0.09	-0.13	0.05	0.05	-0.06	-0.06	0.09	0.09	0.06	0.07	-0.06	0.06	0.7	0.2	0.13	0.02	1	0.04	0.04	0.13
Network cost	-0.97	0.92	-0.27	-0.72	-0.28	0.5	0.99	0.84	-0.9	-0.76	0.99	0.99	-0.96	-0.95	-0.93	-0.92	-0.91	-0.92	-0.94	-0.54	0.22	-0.38	0.84	-0.05	0.04	1	1	-0.22
Greenhouse gas emission	-0.96	0.91	-0.26	-0.7	-0.28	0.5	0.98	0.83	-0.89	-0.75	0.99	0.98	-0.95	-0.95	-0.93	-0.92	-0.91	-0.92	-0.93	-0.55	0.22	-0.4	0.83	-0.05	0.04	1	1	-0.22
Pump energy	0.23	-0.24	-0.03	0.22	0.28	-0.17	-0.22	-0.2	0.24	0.21	-0.22	-0.21	0.23	0.23	0.35	0.35	0.27	0.29	0.24	0.06	0.13	0.15	-0.2	0.04	0.13	-0.22	-0.22	1

Strong negative correlation = -0.6 - -0.8 Strong positive correlation= 0.6 – 0.8

Table 16: Matrix of the CH with strong correlation for different model networks

The fifth t table is showing the variables with very strong correlation. The values of this correlation start from -0.6 to -0.8 and 0.6 to 0.8.

Correlation matrix

	Average Length	Total Length	Total Flow	Total demand	Min elevation	Max elevation	Number of nodes	Nodes demand <=0	Average demand	Average demand per node	Number of pipes	Average Residence Time	Node degree	Link density	Articulation points	Bridges	Eccentricity	Maximun eccentricity	Closeness eccentricity	Average Shortest path length	Max pressure	Pressure and specific pressure threshold 30m	Water service availability	Expected demand	Todini index, threshold	Network cost	Green-house gas emission	Pump energy
Average Length	1	-0.95	0.36	0.83	0.35	-0.43	-0.99	-0.93	0.97	0.88	-0.99	-0.95	1	1	0.9	0.9	0.84	0.85	0.99	0.34	-0.19	0.22	-0.93	0.05	-0.06	-0.97	-0.96	0.23
Total Length	-0.95	1	-0.37	-0.86	-0.33	0.41	0.94	0.91	-0.95	-0.89	0.95	0.91	-0.95	-0.96	-0.84	-0.85	-0.75	-0.77	-0.95	-0.27	0.19	-0.13	0.91	-0.06	0.11	0.92	0.91	-0.24
Total Flow	0.36	-0.37	1	0.52	0.48	-0.07	-0.32	-0.35	0.44	0.43	-0.32	-0.28	0.37	0.38	0.29	0.34	0.18	0.25	0.38	-0.2	0.11	-0.16	-0.35	0.01	0.02	-0.27	-0.26	-0.03
Total demand	0.83	-0.86	0.52	1	0.32	-0.22	-0.78	-0.89	0.93	0.95	-0.78	-0.73	0.84	0.85	0.61	0.64	0.49	0.52	0.87	-0.09	-0.1	-0.18	-0.89	0.05	-0.12	-0.72	-0.7	0.22
Min elevation	0.35	-0.33	0.48	0.32	1	-0.1	-0.33	-0.22	0.35	0.27	-0.33	-0.27	0.35	0.35	0.48	0.49	0.37	0.41	0.36	0.01	0.12	0.05	-0.22	0.04	0.16	-0.28	-0.28	0.28
Max elevation	-0.43	0.41	-0.07	-0.22	-0.1	1	0.46	0.39	-0.36	-0.28	0.46	0.47	-0.42	-0.41	-0.44	-0.45	-0.48	-0.51	-0.4	-0.39	0.2	-0.3	0.39	0.03	0.17	0.5	0.5	-0.17
Number of nodes	-0.99	0.94	-0.32	-0.78	-0.33	0.46	1	0.9	-0.95	-0.84	1	0.97	-0.99	-0.98	-0.92	-0.92	-0.88	-0.89	-0.98	-0.43	0.21	-0.29	0.9	-0.05	0.05	0.99	0.98	-0.22
Nodes demand >= 0	-0.93	0.91	-0.35	-0.89	-0.22	0.39	0.9	1	-0.96	-0.97	0.9	0.81	-0.94	-0.94	-0.7	-0.72	-0.61	-0.63	-0.95	-0.05	0.23	0.02	1	-0.04	0.13	0.84	0.83	-0.2
Average demand	0.97	-0.95	0.44	0.93	0.35	-0.36	-0.95	-0.96	1	0.96	-0.95	-0.89	0.98	0.98	0.81	0.82	0.72	0.74	0.99	0.16	-0.16	0.05	-0.96	0.06	-0.09	-0.9	-0.89	0.24
Average demand per node	0.88	-0.89	0.43	0.95	0.27	-0.28	-0.84	-0.97	0.96	1	-0.84	-0.74	0.9	0.91	0.64	0.66	0.51	0.54	0.93	-0.09	-0.15	-0.14	-0.97	0.04	-0.13	-0.76	-0.75	0.21
Number of pipes	-0.99	0.95	-0.32	-0.78	-0.33	0.46	1	0.9	-0.95	-0.84	1	0.97	-0.99	-0.98	-0.92	-0.92	-0.88	-0.89	-0.98	-0.43	0.21	-0.29	0.9	-0.05	0.05	0.99	0.99	-0.22
Average Residence Time	-0.95	0.91	-0.28	-0.73	-0.27	0.47	0.97	0.81	-0.89	-0.74	0.97	1	-0.94	-0.93	-0.91	-0.91	-0.91	-0.91	-0.91	-0.56	0.22	-0.36	0.81	-0.06	0.05	0.99	0.98	-0.21
Node degree	1	-0.95	0.37	0.84	0.35	-0.42	-0.99	-0.94	0.98	0.9	-0.99	-0.94	1	1	0.88	0.89	0.82	0.84	1	0.31	-0.19	0.19	-0.94	0.05	-0.06	-0.96	-0.95	0.23
Link density	1	-0.96	0.38	0.85	0.35	-0.41	-0.98	-0.94	0.98	0.91	-0.98	-0.93	1	1	0.87	0.88	0.8	0.82	1	0.28	-0.18	0.17	-0.94	0.05	-0.06	-0.95	-0.95	0.23
Articulation points	0.9	-0.84	0.29	0.61	0.48	-0.44	-0.92	-0.7	0.81	0.64	-0.92	-0.91	0.88	0.87	1	1	0.94	0.95	0.86	0.55	-0.13	0.43	-0.7	0.05	0.09	-0.93	-0.93	0.35
Bridges	0.9	-0.85	0.34	0.64	0.49	-0.45	-0.92	-0.72	0.82	0.66	-0.92	-0.91	0.89	0.88	1	1	0.93	0.94	0.87	0.52	-0.13	0.4	-0.72	0.06	0.09	-0.92	-0.92	0.35
Eccentricity	0.84	-0.75	0.18	0.49	0.37	-0.48	-0.88	-0.61	0.72	0.51	-0.88	-0.91	0.82	0.8	0.94	0.93	1	0.99	0.78	0.72	-0.18	0.54	-0.61	0.04	0.06	-0.91	-0.91	0.27
Maximun eccentricity	0.85	-0.77	0.25	0.52	0.41	-0.51	-0.89	-0.63	0.74	0.54	-0.89	-0.91	0.84	0.82	0.95	0.94	0.99	1	0.8	0.68	-0.15	0.54	-0.63	0.04	0.07	-0.92	-0.92	0.29
Closeness eccentricity	0.99	-0.95	0.38	0.87	0.36	-0.4	-0.98	-0.95	0.99	0.93	-0.98	-0.91	1	1	0.86	0.87	0.78	0.8	1	0.24	-0.18	0.13	-0.95	0.05	-0.06	-0.94	-0.93	0.24
Average Shortest path	[illegible]	[illegible]	-0.2	-0.09	0.01	-0.39	-0.43	-0.05	0.16	-0.09	-0.43	[illegible]	[illegible]	0.28	0.55	0.52	0.72	0.68	0.24	1	-0.11	0.77	-0.05	0.03	0.06	-0.54	-0.55	0.06
Max pressure	[illegible]	[illegible]	[illegible]	-0.1	0.12	0.2	0.21	0.23	-0.16	-0.15	0.21	[illegible]	[illegible]	-0.18	-0.13	-0.13	-0.18	-0.15	-0.18	-0.11	1	0.01	0.23	0.05	0.7	0.22	0.22	0.13
Pressure and specific pressure threshold 30m	0.22	-0.13	-0.16	-0.18	0.05	-0.3	-0.29	0.02	0.05	-0.14	-0.29	-0.36	0.19	0.17	0.43	0.4	0.54	0.54	0.13	0.77	0.01	1	0.02	0.05	0.2	-0.38	-0.4	0.15
Water service availability	-0.93	0.91	-0.35	-0.89	-0.22	0.39	0.9	[illegible]	-0.96	-0.97	0.9	0.81	-0.94	-0.94	-0.7	-0.72	-0.61	-0.63	-0.95	-0.05	0.23	0.02	1	-0.04	0.13	0.84	0.83	-0.2
Expected demand	0.05	-0.06	0.01	0.05	0.04	0.03	-0.05	-0.04	0.06	0.04	-0.05	-0.06	0.05	0.05	0.05	0.06	0.04	0.04	0.05	0.03	0.05	0.05	-0.04	1	0.02	-0.05	-0.05	0.04
Todini	-0.06	0.11	0.02	-0.12	0.16	0.17	0.05	0.13	-0.09	-0.13	0.05	0.05	-0.06	-0.06	0.09	0.09	0.06	0.07	-0.06	0.06	0.7	0.2	0.13	0.02	1	0.04	0.04	0.13
Network cost	-0.97	0.92	-0.27	-0.72	-0.28	0.5	0.99	0.84	-0.9	-0.76	0.99	0.99	-0.96	-0.95	-0.93	-0.92	-0.91	-0.92	-0.94	-0.54	0.22	-0.38	0.84	-0.05	0.04	1	1	-0.22
Greenhouse gas emission	-0.96	0.91	-0.26	-0.7	-0.28	0.5	0.98	0.83	-0.89	-0.75	0.99	0.98	-0.95	-0.95	-0.93	-0.92	-0.91	-0.92	-0.93	-0.55	0.22	-0.4	0.83	-0.05	0.04	1	1	-0.22
Pump energy	0.23	-0.24	-0.03	0.22	0.28	-0.17	-0.22	-0.2	0.24	0.21	-0.22	-0.21	0.23	0.23	0.35	0.35	0.27	0.29	0.24	0.06	0.13	0.15	-0.2	0.04	0.13	-0.22	-0.22	1

Very strong negative correlation= -0.8 - -1

Very strong positive correlation= 0.8 – 1

Table 17: Matrix of the CH with very strong correlation for different model networks

The last table is showing the variables with perfect correlation. The values of this correlation -1 or 1

Correlation matrix

	Average Length	Total Length	Total Flow	Total demand	Min elevation	Max elevation	Number of nodes	Nodes demand <=0	Average demand	Average demand per node	Number of pipes	Average Residence Time	Node degree	Link density	Articulation points	Bridges	Eccentricity	Maximun eccentricity	Closeness eccentricity	Average Shortest path length	Max pressure	Pressure and specific pressure threshold 30m	Water service availability	Expected demand	Todini index, threshold	Network cost	Green-house gas emission	Pump energy
Average Length	1	-0.95	0.36	0.83	0.35	-0.43	-0.99	-0.93	0.97	0.88	-0.99	-0.95	1	1	0.9	0.9	0.84	0.85	0.99	0.34	-0.19	0.22	-0.93	0.05	-0.06	-0.97	-0.96	0.23
Total Length	-0.95	1	-0.37	-0.86	-0.33	0.41	0.94	0.91	-0.95	-0.89	0.95	0.91	-0.95	-0.96	-0.84	-0.85	-0.75	-0.77	-0.95	-0.27	0.19	-0.13	0.91	-0.06	0.11	0.92	0.91	-0.24
Total Flow	0.36	-0.37	1	0.52	0.48	-0.07	-0.32	-0.35	0.44	0.43	-0.32	-0.28	0.37	0.38	0.29	0.34	0.18	0.25	0.38	-0.2	0.11	-0.16	-0.35	0.01	0.02	-0.27	-0.26	-0.03
Total demand	0.83	-0.86	0.52	1	0.32	-0.22	-0.78	-0.89	0.93	0.95	-0.78	-0.73	0.84	0.85	0.61	0.64	0.49	0.52	0.87	-0.09	-0.1	-0.18	-0.89	0.05	-0.12	-0.72	-0.7	0.22
Min elevation	0.35	-0.33	0.48	0.32	1	-0.1	-0.33	-0.22	0.35	0.27	-0.33	-0.27	0.35	0.35	0.48	0.49	0.37	0.41	0.36	0.01	0.12	0.05	-0.22	0.04	0.16	-0.28	-0.28	0.28
Max elevation	-0.43	0.41	-0.07	-0.22	-0.1	1	0.46	0.39	-0.36	-0.28	0.46	0.47	-0.42	-0.41	-0.44	-0.45	-0.48	-0.51	-0.4	-0.39	0.2	-0.3	0.39	0.03	0.17	0.5	0.5	-0.17
Number of nodes	-0.99	0.94	-0.32	-0.78	-0.33	0.46	1	0.9	-0.95	-0.84	1	0.97	-0.99	-0.98	-0.92	-0.92	-0.88	-0.89	-0.98	-0.43	0.21	-0.29	0.9	-0.05	0.05	0.99	0.98	-0.22
Nodes demand >= 0	-0.93	0.91	-0.35	-0.89	-0.22	0.39	0.9	1	-0.96	-0.97	0.9	0.81	-0.94	-0.94	-0.7	-0.72	-0.61	-0.63	-0.95	-0.05	0.23	0.02	1	-0.04	0.13	0.84	0.83	-0.2
Average demand	0.97	-0.95	0.44	0.93	0.35	-0.36	-0.95	-0.96	1	0.96	-0.95	-0.89	0.98	0.98	0.81	0.82	0.72	0.74	0.99	0.16	-0.16	0.05	-0.96	0.06	-0.09	-0.9	-0.89	0.24
Average demand per node	0.88	-0.89	0.43	0.95	0.27	-0.28	-0.84	-0.97	0.96	1	-0.84	-0.74	0.9	0.91	0.64	0.66	0.51	0.54	0.93	-0.09	-0.15	-0.14	-0.97	0.04	-0.13	-0.76	-0.75	0.21
Number of pipes	-0.99	0.95	-0.32	-0.78	-0.33	0.46	1	0.9	-0.95	-0.84	1	0.97	-0.99	-0.98	-0.92	-0.92	-0.88	-0.89	-0.98	-0.43	0.21	-0.29	0.9	-0.05	0.05	0.99	0.99	-0.22

Perfect positive correlation

Table 18: Matrix of the CH with perfect correlation for different model networks

The pair plot graph was implemented to visualise the correlation of the characteristics. It shows how the data are correlated.

In the following, some of the perfectly correlated characteristics and the table with the characteristics with perfect correlation.

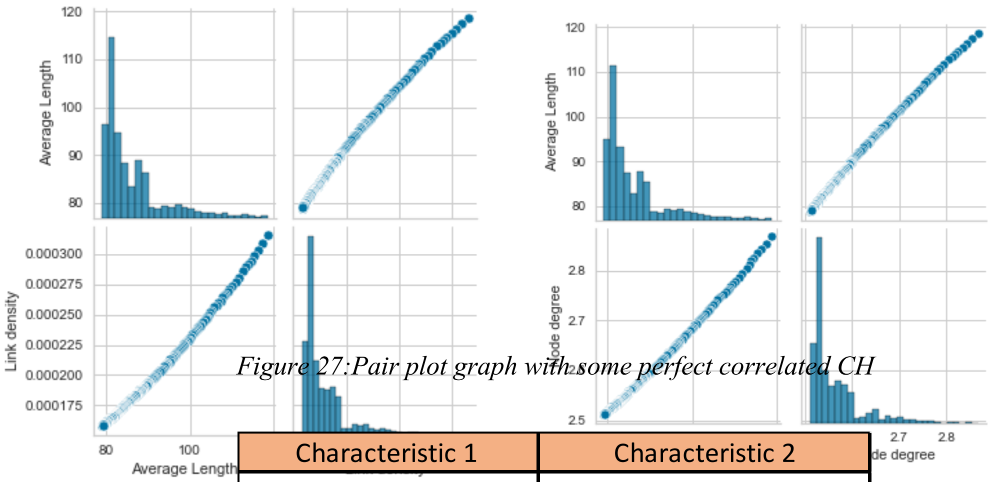


Figure 27:Pair plot graph with some perfect correlated CH

Characteristic 1	Characteristic 2
Average length	Node degree
Average length	Link density
Number of nodes	Number of pipes
Node with demand >=0	Water service avaliability
Node degree	Link density
Articulation points	Bridges
Closeness eccentricity	Node degree
Closeness eccentricity	Link density
Network cost	Green house emission

Table 19: Characteristics with perfect correlation

Principal Component analysis

For the principal component analysis, 38 features and 600 networks were taken as a dataset. These were evaluated, and as a result, the dimensionality of the features was reduced in three principal components. These three components are determined by the variance of the data found in the PCA. Table 21 below shows the variance values, whereby it is determined that more than 80% of the data can be represented in 3 dimensions. Additionally, the variance data is shown visually in percentage in figure 28.

	Variance	Sum variance
componet 1	0.65533007	
componet 2	0.11724901	0.7726
componet 3	0.07422969	0.8468
componet 4	0.03950605	0.8863
componet 5	0.036177	0.9225
componet 6	0.02593766	0.9484
componet 7	0.01567099	0.9641
componet 8	0.01106058	0.9752
componet 9	0.01035589	0.9855
componet 10	0.00438901	0.9899
componet 11	0.00339305	0.9933
componet 12	0.00276829	0.9961
componet 13	0.00144563	0.9975
componet 14	0.00101907	0.9985
componet 15	0.00083546	0.9994
componet 16	0.00030112	0.9997
componet 17	0.00015035	0.9998
componet 18	6.99E-05	0.9999
componet 19	5.01E-05	0.9999
componet 20	4.32E-05	1.0000
componet 21	1.26E-05	1.0000
componet 22	2.98E-06	1.0000

Table 20: explained variance

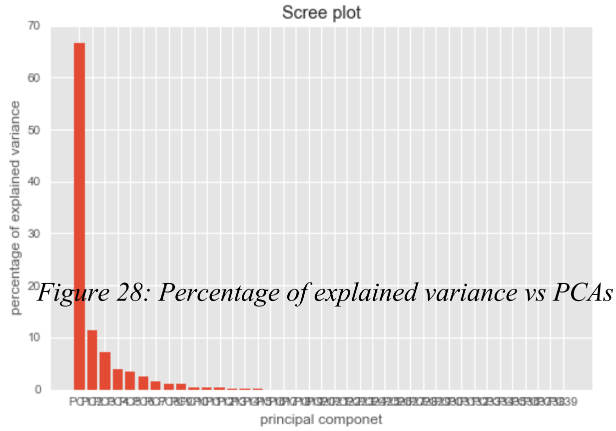


Figure 28: Percentage of explained variance vs PCAs

The representation of the data set in three PCA is as following:

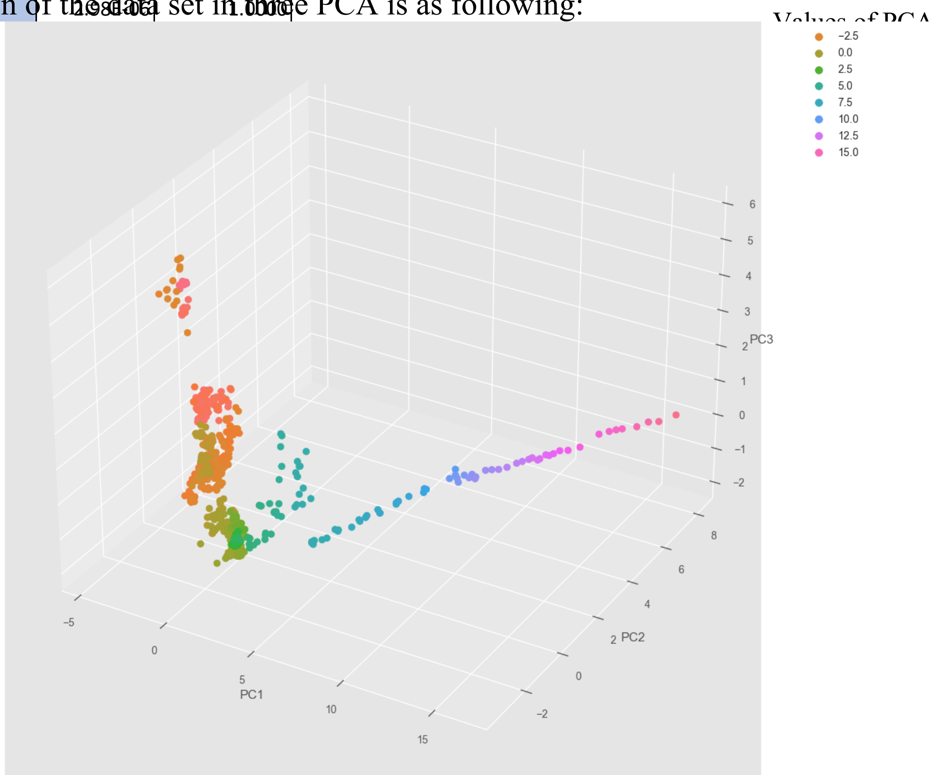
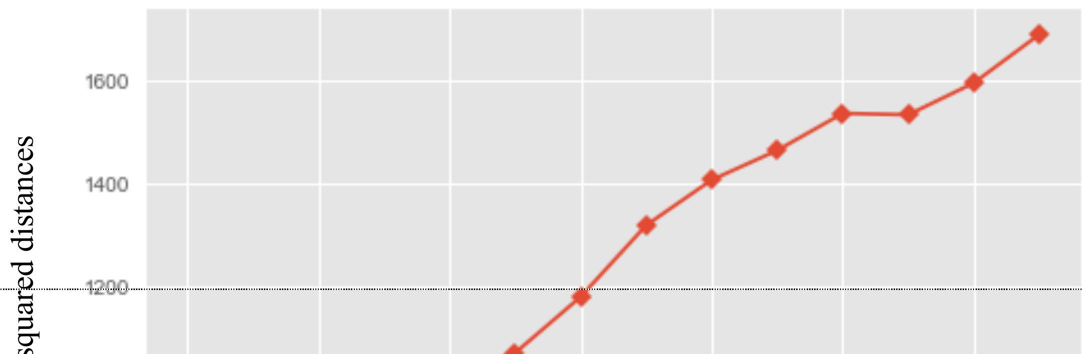


Figure 29: Representation of the data set in the best view

Elbow method

Using the elbow method, the number of clusters to be generated was determined. In this case, figure 30 shows the inflexion point of the arm, which is 4.



k

Figure 30: Elbow method

Silhouette method

The silhouette method was intended to identify how the data set and its dimensions, already reduced by the PCA method, can be grouped. For this case, 4 clusters were created. This number of clusters was determined by the elbow method. Figure 31 shows the silhouette representation of the data on the left-hand side and the clusters generated on the right-hand side.

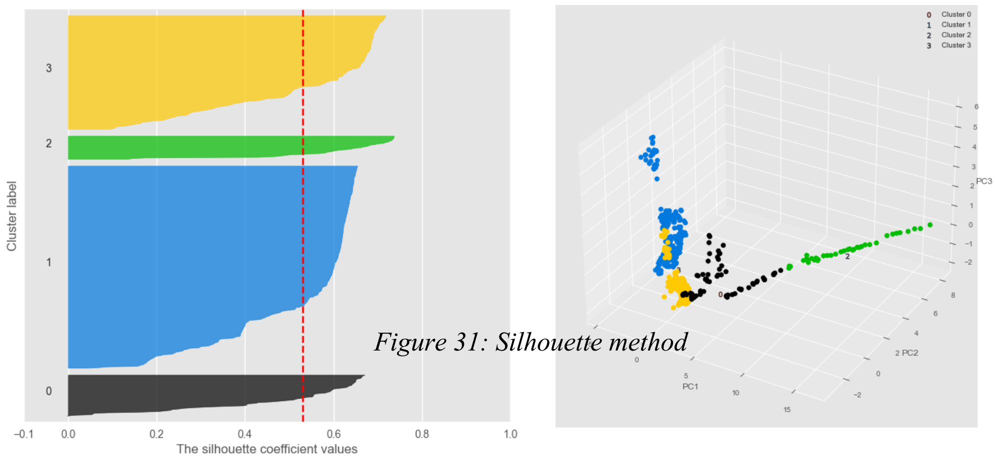


Figure 31: Silhouette method

As a result of the creation of the clusters, it was possible to relate the set of characteristics that make up each cluster. The composition of the clusters is given as follows. Cluster 2 is made up of 38 networks, cluster 0 is made up of 65 networks, cluster 3 is made up of 174 networks, and cluster 1 is made up of 318 networks. The min, max and average ranges of each cluster are shown below

Cluster 2	Average Length (m (meters))	Total Length (m (meters))	Total Flow (CMH (cubic meters/hr))	Total demand (CMH (cubic meters/hr))	Min elevation (m (meters))	Max elevation (m (meters))	Number of nodes (units)	Number of nodes demand less than or equal 0 (units)	Average demand (units)	Average demand per node (units)	Number of pipes (units)	Number of tanks (units)	Number of reservoir (units)	Number of pumps (units)	Number of valves (units)	Number of patterns (units)	Average Residence Time (units)
min	100.835	772411	153.9806	4.653786	-6.15	14.77619	4552	837	0.00081	0.003911	6505	2	1	7	16	13	1307564547
med	106.444	775280	163.0977	5.077018	-6.15	14.77619	5316	1090.5	0.00095	0.004656	7278	2	1	7	16	13	1339988856
max	118.741	778152	173.1269	5.077018	-6.15	14.77619	5738	1193	0.00112	0.006066	7704	2	1	7	16	13	1762836856
Cluster 0																	
Min	90.1878	776983	130.123	4.188128	-6.15	14.77619	5737	1200	0.00062	0.003002	7703	2	1	7	16	13	1445164522
media	95.7066	781555	140.9781	4.290084	-6.15	14.77619	6190	1306	0.0007	0.003292	8172	2	1	7	16	13	1769338676
max	100.868	783732	170.2203	4.769727	-6.15	14.77619	6707	1395	0.00083	0.003975	8690	2	1	7	16	13	1999228399
Cluster 3																	
min	83.3949	782027	110.273	4.188124	-7.000089	14.77619	6712	1333	0.00057	0.002947	8697	2	1	7	16	13	1965375606
med	87.0096	783540	131.8849	4.188128	-7.000089	14.77619	7012	1375	0.0006	0.003046	9005	2	1	7	16	13	2081483436
max	90.0833	784729	142.9641	4.188128	-6.15	14.77619	7406	1421	0.00062	0.003142	9401	2	1	7	16	13	[illegible]
Cluster 1																	
min	78.9436	783318	106.7448	4.187163	-7.000089	14.77619	7404	1362	0.00053	0.002816	9398	2	1	7	16	13	2275480559
med	81.2835	785670	133.02	4.188128	-7.000089	14.8405	7656	1411	0.00055	0.002968	9658	2	1	7	16	13	2443623104
max	83.421	787306	172.5852	4.188133	-6.15	14.8405	7949	1487	0.00057	0.003075	9958	2	1	7	16	13	2650827616

Cluster 2	Node degree (degree)	Link density (units)	Articulation points (units)	Bridges (units)	Eccentricity (m (meters))	Maximun eccentricity (m (meters))	Closeness eccentricity (m (meters))	Average Shortest path length (m (meters))	Max pressure (m (meters))	Pressure and specific pressure threshold 30m (m (meters))	Water service availability (CMH (cubic meters/hr))	Expected demand (CMH (cubic meters/hr))	Todini index, threshold (hours)	Entropy (units)	Water age (hour)	Water age at 48 hours (hours)	Population that is impacted by water age greater than 24 hours (hours)	Network cost (units)	Greenhouse gas emission (C02)	Pump energy (C02)
min	2.6928	0.0002	78	77	90.5424995	131	0.00040939	38.404098	36.34333	0.210526	840	1.234884	-11.566	2	0	0	4550	1E+07	1E+07	968533
med	2.7463	0.0003	85	84	97.7880929	140	0.00045495	43.312355	37.64314	0.578947	1093.5	1.234884	-1.9284	2	0	0	5314	1E+07	1E+07	968533
max	2.8676	0.0003	87	86	100.499035	145	0.00056575	44.877467	144.7036	3.315789	1196	1.234884	-0.3424	2	0	0	5736	1E+07	1E+07	968534
Cluster 0																				
Min	2.5978	0.0002	75	71	97.1433528	137	0.00032907	44.998831	143.4386	2.842105	1203	1.234884	-58.162	2	0	0	5735	1E+07	1E+07	968533
media	2.6464	0.0002	82	80	100.235897	143	0.00036746	45.813928	246.4359	4.631579	1309	1.234884	67.532	2	0	0	6188	1E+07	1E+07	968533
max	2.6929	0.0002	87	86	104.343166	151	0.00040681	46.598326	920.6823	5.421053	1398	1.234884	2187.3	2	0	0	6705	1E+07	1E+07	968534
Cluster 3																				
min	2.5446	0.0002	49	46	82.4802577	113	0.00029306	43.729391	213.8051	1.842105	1336	1.234884	18.314	2	0	0	6710	1E+07	1E+07	968510
med	2.5744	0.0002	55	53	86.941202	119	0.00030815	45.331764	756.4895	4.052632	1378	1.234884	51.983	2	0	0	7010	1E+07	1E+07	968533
max	2.5982	0.0002	75	71	99.0302308	141	0.00033384	46.707329	24863.69	8.736842	1424	1.234884	751.32	2	0	0	7404	1E+07	2E+07	968538
Cluster 1																				
min	2.5109	0.0002	36	33	76.0347648	99	0.00026411	41.130114	23.1781	0	1365	1.234884	21.41	2	0	0	7402	1E+07	1E+07	968503
med	2.5282	0.0002	43	41	80.0410959	105	0.00028224	42.730622	396.8112	1.473684	1414	1.234884	145.31	2	0	0	7654	1E+07	2E+07	968530
max	2.5445	0.0002	59	56	84.3305818	114	0.00029483	44.916821	90651.75	4.052632	1490	1.234884	1954.9	2	0	0	7947	1E+07	2E+07	968535

Table 1:relation of characteristics with a cluster

5.3 DMAs design

The results of the DMA design carried out by Gondwana gave several different designs. These are composed of 5, 6, 7 or 8 DMAs per model network. The number of sensors per network varies from 130 to 180, with an average value of 180. Concerning the "optimal" DMA design score, all the networks have high values, and none of the designs represents a complete Pareto front graph.
Figure 32 shows the diversity of DMAs designs created from Gondwana and the variety of each DMA's size.

Figure 32: DMAs design generate by Gondwana

Figure 33 shows two examples of the Pareto chart generated due to designing DMAs for two different networks.

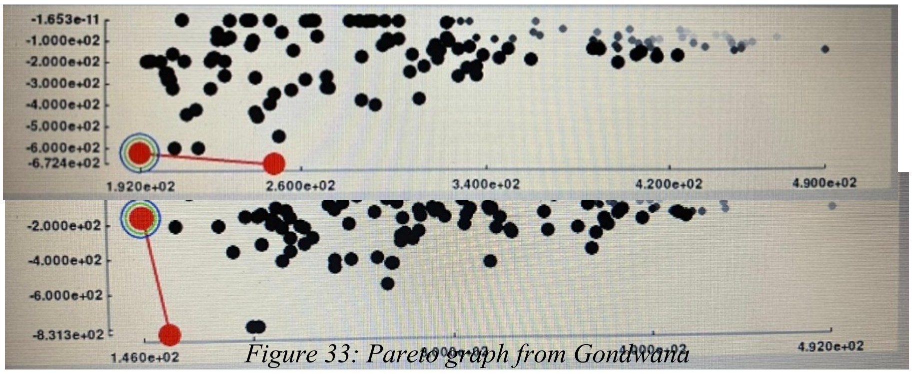


Figure 33: Pareto graph from Gondwana

5.4 Characterization and DMAs design

The created DMA designs were evaluated based on the clusters created with the characteristics. The data set is distributed as follows: 10 DMAs correspond to cluster 2, 11 networks correspond to cluster 0, 34 networks correspond to DMA 1, and 11 networks correspond to cluster 3.

The DMA designs corresponding to cluster 2 can be described as eight networks with five DMA designs and two nets with 6 DMA designs.
The DMA designs corresponding to cluster 0 are described as five DMA designs in six networks, four DMA designs in six networks and seven DMA designs in one net.
The DMA designs correspond to cluster 1: twenty-five networks with five DMAs, eight networks with six DMAs, and two with seven DMAs.
The DMA designs correspond to cluster 3: seven networks with 5 DMAs, two nets with six DMAs, one networks with seven DMAs, and one net with 8 DMAs.
From the above, it can be said that there is no clear relationship between the characteristics that can define an DMA, supported by the multiple variations that exist for a given cluster or set of characteristics.

Chapter 6 Conclusions

The main research question is formulated as follows:

1. How can a proper characterization of a WDN help successful DMA design solutions?

A proper characterization of the networks can be considered as the integration of different methodologies that allow a description of the network in detail. In this sense, recognizing which combination of them and what ranges can be established to contribute to a better understanding of an optimal DMA.

Followed by the specific research questions:

2. Which characterization methods are available in the literature that are feasible for this study?

The methods of characterization have a countless of criteria and therefore all of them could be feasible for this study. However, the grouping of the methods most common are taken as basis of this research. These are unified into five main metrics as follows: hydraulic, water quality, topological, water safety and economical.

3. How can it be generated multiple model networks defining hydraulic limitation from the demands and supply?

Multiple model networks can be generated by using optimization criteria by using genetic algorithms in which the establishment of the boundaries, constrains, objective funtions, decision variables are the back bone of its development. The hydraulics limitation were give as a constraint and decision variables in order to create new model networks with good hydraulic performance.

4. How does the current optimization practices conceptualise DMA design?

Gondwana presents two objective functions for the design of DMAs; the first objective seeks the minimum number of limits for the total design of DMAs in the network. The second objective evaluates the demand for each of the DMA designs created to make an equitable distribution of the demand for DMAs. Genetic algorithms evaluate these two objective functions, and the result is a Pareto graph.

5. What are the relationships between the characteristics in the generated networks?

The characteristics are highly correlated with each other, as can be seen in the correlation matrix and in the results of the clusters where the ranges of relationships between the characteristics were determined.

Limitations y Recommendations

Generation of the new model networks

1. The new network models were generated using Gondwana and its genetic algorithm. The base network was run for 600 generations with a very low population. Based on the above, it is recommended to implement a higher number of population and generation to obtain a higher diversity in the networks.
2. The skeletonisation method used with the networks is based on removing and joining pipes under the assigned threshold. For this case study, this threshold was not changed, but it could be a way to generate variation in the networks.

Characterization development

1. When using the features for different networks, it must be taken into account that they have different simulation times. This means that some features do not generate a single value, but rather a several operations such as summation, averaging or counting of data must be performed to find a single value.
2. The Den Haag network (base network) is taken as a perfect network for this case study. This means that it represents as close as possible the reality. Additionally, the number of nodes or pipes that may have values of 0 in the system does not affect the calculation of the characteristics.
3. For the implementation and use of the water quality and safety metrics, it is necessary to use Epanet version 2.2, otherwise, the simulation will not be performed correctly.
4. The implementation of other clustering methods is necessary to compare the results already found with the elbow method.

DMA design

1. The allocation of the ranges generated as initial values for the number of DMAs was related only to the base demand and the number of consumers. Therefore, it is necessary to perform an analysis according to the case study.
2. In order to create DMAs automatically for a certain number of nets, it is necessary to implement a Python script to facilitate the process and make it faster.
3. The number of generations that should be used for the design of DMAs should be higher than 500 to obtain better results in the design of DMAs.

Characteristics and DMAs design

1. To better analyse the relationship between the characteristics and the DMA designs, it is necessary to have a greater number of networks with DMAs. In turn, the designs generated should represent an "optimal" solution.

References

A Step-by-Step Explanation of Principal Component Analysis (PCA) | Built In. (n.d.). Retrieved April 17, 2022, from https://builtin.com/data-science/step-step-explanation-principal-component-analysis

Alvisi, S., & Franchini, M. (2010). Pipe roughness calibration in water distribution systems using grey numbers. *Journal of Hydroinformatics*, *12*(4), 424–445. https://doi.org/10.2166/HYDRO.2010.089

Alvisi, S., & Franchini, M. (2014). A Procedure for the Design of District Metered Areas in Water Distribution Systems. *Procedia Engineering*, *70*, 41–50. https://doi.org/10.1016/J.PROENG.2014.02.006

Armand, H., Stoianov, I., & Graham, N. (2018). Impact of network sectorisation on water quality management. *Journal of Hydroinformatics*, *20*(2), 424–439. https://doi.org/10.2166/HYDRO.2017.072

Ausiello, G., Firmani, D., & Laura, L. (2012). Real-time monitoring of undirected networks: Articulation points, bridges, and connected and biconnected components. *Networks*, *59*(3), 275–288. https://doi.org/10.1002/NET.21450

Awe, O. M., Okolie, S. T. A., & Fayomi, O. S. I. (2019). Optimization of Water Distribution Systems: A Review. *Journal of Physics: Conference Series*, *1378*(2). https://doi.org/10.1088/1742-6596/1378/2/022068

Awumah, K., Goulter, I., & Bhatt, S. K. (1990). Assessment of reliability in water distribution networks using entropy based measures. *Stochastic Hydrology and Hydraulics 1990 4:4*, *4*(4), 309–320. https://doi.org/10.1007/BF01544084

Bui, X. K., Marlim, M. S., & Kang, D. (2020). Water network partitioning into district metered areas: A state-of-the-art review. *Water (Switzerland)*, *12*(4). https://doi.org/10.3390/W12041002

Bui, X. K., Marlim, M. S., & Kang, D. (2021). Optimal design of district metered areas in a water distribution network using coupled self-organizing map and community structure algorithm. *Water (Switzerland)*, *13*(6). https://doi.org/10.3390/W13060836

Cassidy, J., Barbosa, B., Damião, M., Ramalho, P., Ganhão, A., Santos, A., & Feliciano, J. (2021). Taking water efficiency to the next level: Digital tools to reduce non-revenue water. *Journal of Hydroinformatics*, *23*(3), 453–465. https://doi.org/10.2166/HYDRO.2020.072

Chang, D. E., Yoo, D. G., & Kim, J. H. (2021). A study on the practical pressure-driven hydraulic analysis method considering actual water supply characteristics of water distribution network. *Sustainability (Switzerland)*, *13*(5). https://doi.org/10.3390/su13052793

Chondronasios, A., Gonelas, K., Kanakoudis, V., Patelis, M., & Korkana, P. (2017). Optimizing DMAs' formation in a water pipe network: The water aging and the operating pressure factors. *Journal of Hydroinformatics*, *19*(6), 890–899. https://doi.org/10.2166/HYDRO.2017.156

citypopulation. (2022). *Netherlands: Administrative Division (Provinces and Municipalities) - Population Statistics, Charts and Map*. https://www.citypopulation.de/en/netherlands/admin/

Clark, R. M. (1992). Water Quality Modeling in Distribution Systems. *Journal of Environmental Science and Health. Part A: Environmental Science and Engineering and Toxicology*, *27*(5), 1329–1366. https://doi.org/10.1080/10934529209375799

Closeness centrality | Network analysis of protein interaction data. (n.d.). Retrieved April 12, 2022, from https://www.ebi.ac.uk/training/online/courses/network-analysis-of-protein-interaction-data-an-introduction/building-and-analysing-ppins/topological-ppin-analysis/closeness-centrality/

de Paola, F., Fontana, N., Galdiero, E., Giugni, M., Degli Uberti, G. S., & Vitaletti, M. (2014). Optimal design of district metered areas in water distribution networks. *Procedia Engineering*, *70*, 449–457. https://doi.org/10.1016/J.PROENG.2014.02.050

Diao, K., Berardi, L., Laucelli, D. B., Ulanicki, B., & Giustolisi, O. (2022). Topological and hydraulic metrics-based search space reduction for optimal re-sizing of water distribution networks. *Journal of Hydroinformatics*, *00*, 1. https://doi.org/10.2166/HYDRO.2022.158

Diao, K., Farmani, R., Fu, G., & Butler, D. (2014). Vulnerability Assessment Of Water Distribution Systems Using Directed And Undirected Graph Theory. *International Conference on Hydroinformatics*. https://academicworks.cuny.edu/cc_conf_hic/267

World Environmental and Water Resources Congress 2019, 573–586. https://doi.org/10.1061/9780784482353.055

Water Network Tool for Resilience (WNTR) User Manual: Version 0.2.3 | Science Inventory | US EPA. (n.d.). Retrieved April 13, 2022, from https://cfpub.epa.gov/si/si_public_record_Report.cfm?dirEntryId=349798&Lab=CESER

WNTR User Manual — WNTR 0.1.4 documentation. (n.d.). Retrieved March 31, 2022, from https://dbhart-wntr.readthedocs.io/en/metrics/

Wright, R., Stoianov, I., Parpas, P., Henderson, K., & King, J. (2014). Adaptive water distribution networks with dynamically reconfigurable topology. *Journal of Hydroinformatics*, *16*(6), 1280–1301. https://doi.org/10.2166/HYDRO.2014.086

Wu, Z., Boulos, P., de Schaetzen, W., Orr, C. H., & Moore, M. (2004). Using genetic algorithms for water distribution system optimization. *Bridging the Gap: Meeting the World's Water and Environmental Resources Challenges - Proceedings of the World Water and Environmental Resources Congress 2001*, *111*. https://doi.org/10.1061/40569(2001)409

Yazdani, A., & Jeffrey, P. (2011). Complex network analysis of water distribution systems. *Chaos*, *21*(1). https://doi.org/10.1063/1.3540339

Zarghami, S., Gunawan, I., & Schultmann, F. (2019). Integrating Topological and Hydraulic Attributes for Robustness Analysis of Water Distribution Networks. *International Journal of Industrial Engineering and Operations Management*, *01*(01), 1–11. https://doi.org/10.46254/j.ieom.20190101

Zhang, M., Gong, G., & Zeng, L. (2021). Investigation for a novel optimization design method of ground source heat pump based on hydraulic characteristics of buried pipe network. *Applied Thermal Engineering*, *182*. https://doi.org/10.1016/j.applthermaleng.2020.116069

A Step-by-Step Explanation of Principal Component Analysis (PCA) | Built In. (n.d.). Retrieved April 17, 2022, from https://builtin.com/data-science/step-step-explanation-principal-component-analysis

Alvisi, S., & Franchini, M. (2010). Pipe roughness calibration in water distribution systems using grey numbers. *Journal of Hydroinformatics*, *12*(4), 424–445. https://doi.org/10.2166/HYDRO.2010.089

Alvisi, S., & Franchini, M. (2014). A Procedure for the Design of District Metered Areas in Water Distribution Systems. *Procedia Engineering*, *70*, 41–50. https://doi.org/10.1016/J.PROENG.2014.02.006

Armand, H., Stoianov, I., & Graham, N. (2018). Impact of network sectorisation on water quality management. *Journal of Hydroinformatics*, *20*(2), 424–439. https://doi.org/10.2166/HYDRO.2017.072

Ausiello, G., Firmani, D., & Laura, L. (2012). Real-time monitoring of undirected networks: Articulation points, bridges, and connected and biconnected components. *Networks*, *59*(3), 275–288. https://doi.org/10.1002/NET.21450

Awe, O. M., Okolie, S. T. A., & Fayomi, O. S. I. (2019). Optimization of Water Distribution Systems: A Review. *Journal of Physics: Conference Series*, *1378*(2). https://doi.org/10.1088/1742-6596/1378/2/022068

Awumah, K., Goulter, I., & Bhatt, S. K. (1990). Assessment of reliability in water distribution networks using entropy based measures. *Stochastic Hydrology and Hydraulics 1990 4:4*, *4*(4), 309–320. https://doi.org/10.1007/BF01544084

Bui, X. K., Marlim, M. S., & Kang, D. (2020). Water network partitioning into district metered areas: A state-of-the-art review. *Water (Switzerland)*, *12*(4). https://doi.org/10.3390/W12041002

Bui, X. K., Marlim, M. S., & Kang, D. (2021). Optimal design of district metered areas in a water distribution network using coupled self-organizing map and community structure algorithm. *Water (Switzerland)*, *13*(6). https://doi.org/10.3390/W13060836

Cassidy, J., Barbosa, B., Damião, M., Ramalho, P., Ganhão, A., Santos, A., & Feliciano, J. (2021). Taking water efficiency to the next level: Digital tools to reduce non-revenue water. *Journal of Hydroinformatics*, *23*(3), 453–465. https://doi.org/10.2166/HYDRO.2020.072

Chang, D. E., Yoo, D. G., & Kim, J. H. (2021). A study on the practical pressure-driven hydraulic analysis method considering actual water supply characteristics of water distribution network. *Sustainability (Switzerland)*, *13*(5). https://doi.org/10.3390/su13052793

Chondronasios, A., Gonelas, K., Kanakoudis, V., Patelis, M., & Korkana, P. (2017). Optimizing DMAs' formation in a water pipe network: The water aging and the operating pressure factors. *Journal of Hydroinformatics*, *19*(6), 890–899. https://doi.org/10.2166/HYDRO.2017.156

citypopulation. (2022). *Netherlands: Administrative Division (Provinces and Municipalities) - Population Statistics, Charts and Map*. https://www.citypopulation.de/en/netherlands/admin/

Clark, R. M. (1992). Water Quality Modeling in Distribution Systems. *Journal of Environmental Science and Health. Part A: Environmental Science and Engineering and Toxicology*, *27*(5), 1329–1366. https://doi.org/10.1080/10934529209375799

Closeness centrality | Network analysis of protein interaction data. (n.d.). Retrieved April 12, 2022, from https://www.ebi.ac.uk/training/online/courses/network-analysis-of-protein-interaction-data-an-introduction/building-and-analysing-ppins/topological-ppin-analysis/closeness-centrality/

de Paola, F., Fontana, N., Galdiero, E., Giugni, M., Degli Uberti, G. S., & Vitaletti, M. (2014). Optimal design of district metered areas in water distribution networks. *Procedia Engineering*, *70*, 449–457. https://doi.org/10.1016/J.PROENG.2014.02.050

Diao, K., Berardi, L., Laucelli, D. B., Ulanicki, B., & Giustolisi, O. (2022). Topological and hydraulic metrics-based search space reduction for optimal re-sizing of water distribution networks. *Journal of Hydroinformatics*, *00*, 1. https://doi.org/10.2166/HYDRO.2022.158

Diao, K., Farmani, R., Fu, G., & Butler, D. (2014). Vulnerability Assessment Of Water Distribution Systems Using Directed And Undirected Graph Theory. *International Conference on Hydroinformatics*. https://academicworks.cuny.edu/cc_conf_hic/267

World Environmental and Water Resources Congress 2019, 573–586. https://doi.org/10.1061/9780784482353.055
Water Network Tool for Resilience (WNTR) User Manual: Version 0.2.3 | Science Inventory | US EPA. (n.d.). Retrieved April 13, 2022, from https://cfpub.epa.gov/si/si_public_record_Report.cfm?dirEntryId=349798&Lab=CESER
WNTR User Manual — WNTR 0.1.4 documentation. (n.d.). Retrieved March 31, 2022, from https://dbhart-wntr.readthedocs.io/en/metrics/
Wright, R., Stoianov, I., Parpas, P., Henderson, K., & King, J. (2014). Adaptive water distribution networks with dynamically reconfigurable topology. *Journal of Hydroinformatics*, *16*(6), 1280–1301. https://doi.org/10.2166/HYDRO.2014.086
Wu, Z., Boulos, P., de Schaetzen, W., Orr, C. H., & Moore, M. (2004). Using genetic algorithms for water distribution system optimization. *Bridging the Gap: Meeting the World's Water and Environmental Resources Challenges - Proceedings of the World Water and Environmental Resources Congress 2001*, *111*. https://doi.org/10.1061/40569(2001)409
Yazdani, A., & Jeffrey, P. (2011). Complex network analysis of water distribution systems. *Chaos*, *21*(1). https://doi.org/10.1063/1.3540339
Zarghami, S., Gunawan, I., & Schultmann, F. (2019). Integrating Topological and Hydraulic Attributes for Robustness Analysis of Water Distribution Networks. *International Journal of Industrial Engineering and Operations Management*, *01*(01), 1–11. https://doi.org/10.46254/j.ieom.20190101
Zhang, M., Gong, G., & Zeng, L. (2021). Investigation for a novel optimization design method of ground source heat pump based on hydraulic characteristics of buried pipe network. *Applied Thermal Engineering*, *182*. https://doi.org/10.1016/j.applthermaleng.2020.116069

A Step-by-Step Explanation of Principal Component Analysis (PCA) | Built In. (n.d.). Retrieved April 17, 2022, from https://builtin.com/data-science/step-step-explanation-principal-component-analysis

Alvisi, S., & Franchini, M. (2010). Pipe roughness calibration in water distribution systems using grey numbers. *Journal of Hydroinformatics*, *12*(4), 424–445. https://doi.org/10.2166/HYDRO.2010.089

Alvisi, S., & Franchini, M. (2014). A Procedure for the Design of District Metered Areas in Water Distribution Systems. *Procedia Engineering*, *70*, 41–50. https://doi.org/10.1016/J.PROENG.2014.02.006

Armand, H., Stoianov, I., & Graham, N. (2018). Impact of network sectorisation on water quality management. *Journal of Hydroinformatics*, *20*(2), 424–439. https://doi.org/10.2166/HYDRO.2017.072

Ausiello, G., Firmani, D., & Laura, L. (2012). Real-time monitoring of undirected networks: Articulation points, bridges, and connected and biconnected components. *Networks*, *59*(3), 275–288. https://doi.org/10.1002/NET.21450

Awe, O. M., Okolie, S. T. A., & Fayomi, O. S. I. (2019). Optimization of Water Distribution Systems: A Review. *Journal of Physics: Conference Series*, *1378*(2). https://doi.org/10.1088/1742-6596/1378/2/022068

Awumah, K., Goulter, I., & Bhatt, S. K. (1990). Assessment of reliability in water distribution networks using entropy based measures. *Stochastic Hydrology and Hydraulics 1990 4:4*, *4*(4), 309–320. https://doi.org/10.1007/BF01544084

Bui, X. K., Marlim, M. S., & Kang, D. (2020). Water network partitioning into district metered areas: A state-of-the-art review. *Water (Switzerland)*, *12*(4). https://doi.org/10.3390/W12041002

Bui, X. K., Marlim, M. S., & Kang, D. (2021). Optimal design of district metered areas in a water distribution network using coupled self-organizing map and community structure algorithm. *Water (Switzerland)*, *13*(6). https://doi.org/10.3390/W13060836

Cassidy, J., Barbosa, B., Damião, M., Ramalho, P., Ganhão, A., Santos, A., & Feliciano, J. (2021). Taking water efficiency to the next level: Digital tools to reduce non-revenue water. *Journal of Hydroinformatics*, *23*(3), 453–465. https://doi.org/10.2166/HYDRO.2020.072

Chang, D. E., Yoo, D. G., & Kim, J. H. (2021). A study on the practical pressure-driven hydraulic analysis method considering actual water supply characteristics of water distribution network. *Sustainability (Switzerland)*, *13*(5). https://doi.org/10.3390/su13052793

Chondronasios, A., Gonelas, K., Kanakoudis, V., Patelis, M., & Korkana, P. (2017). Optimizing DMAs' formation in a water pipe network: The water aging and the operating pressure factors. *Journal of Hydroinformatics*, *19*(6), 890–899. https://doi.org/10.2166/HYDRO.2017.156

citypopulation. (2022). *Netherlands: Administrative Division (Provinces and Municipalities) - Population Statistics, Charts and Map*. https://www.citypopulation.de/en/netherlands/admin/

Clark, R. M. (1992). Water Quality Modeling in Distribution Systems. *Journal of Environmental Science and Health. Part A: Environmental Science and Engineering and Toxicology*, *27*(5), 1329–1366. https://doi.org/10.1080/10934529209375799

Closeness centrality | Network analysis of protein interaction data. (n.d.). Retrieved April 12, 2022, from https://www.ebi.ac.uk/training/online/courses/network-analysis-of-protein-interaction-data-an-introduction/building-and-analysing-ppins/topological-ppin-analysis/closeness-centrality/

de Paola, F., Fontana, N., Galdiero, E., Giugni, M., Degli Uberti, G. S., & Vitaletti, M. (2014). Optimal design of district metered areas in water distribution networks. *Procedia Engineering*, *70*, 449–457. https://doi.org/10.1016/J.PROENG.2014.02.050

Diao, K., Berardi, L., Laucelli, D. B., Ulanicki, B., & Giustolisi, O. (2022). Topological and hydraulic metrics-based search space reduction for optimal re-sizing of water distribution networks. *Journal of Hydroinformatics*, *00*, 1. https://doi.org/10.2166/HYDRO.2022.158

Diao, K., Farmani, R., Fu, G., & Butler, D. (2014). Vulnerability Assessment Of Water Distribution Systems Using Directed And Undirected Graph Theory. *International Conference on Hydroinformatics*. https://academicworks.cuny.edu/cc_conf_hic/267

Vargas, K., & Saldarriaga, J. (2019). Potential district metered area identification in water distribution networks using hydraulic criteria and the box covering algorithm. *World Environmental and Water Resources Congress 2019: Hydraulics, Waterways, and Water Distribution Systems Analysis - Selected Papers from the World Environmental and Water Resources Congress 2019*, 573–586. https://doi.org/10.1061/9780784482353.055

Water Network Tool for Resilience (WNTR) User Manual: Version 0.2.3 | Science Inventory | US EPA. (n.d.). Retrieved April 13, 2022, from https://cfpub.epa.gov/si/si_public_record_Report.cfm?dirEntryId=349798&Lab=CESER

WNTR User Manual — WNTR 0.1.4 documentation. (n.d.). Retrieved March 31, 2022, from https://dbhart-wntr.readthedocs.io/en/metrics/

Wright, R., Stoianov, I., Parpas, P., Henderson, K., & King, J. (2014). Adaptive water distribution networks with dynamically reconfigurable topology. *Journal of Hydroinformatics*, *16*(6), 1280–1301. https://doi.org/10.2166/HYDRO.2014.086

Wu, Z., Boulos, P., de Schaetzen, W., Orr, C. H., & Moore, M. (2004). Using genetic algorithms for water distribution system optimization. *Bridging the Gap: Meeting the World's Water and Environmental Resources Challenges - Proceedings of the World Water and Environmental Resources Congress 2001*, *111*. https://doi.org/10.1061/40569(2001)409

Yazdani, A., & Jeffrey, P. (2011). Complex network analysis of water distribution systems. *Chaos*, *21*(1). https://doi.org/10.1063/1.3540339

Zarghami, S., Gunawan, I., & Schultmann, F. (2019). Integrating Topological and Hydraulic Attributes for Robustness Analysis of Water Distribution Networks. *International Journal of Industrial Engineering and Operations Management*, *01*(01), 1–11. https://doi.org/10.46254/j.ieom.20190101

Zhang, M., Gong, G., & Zeng, L. (2021). Investigation for a novel optimization design method of ground source heat pump based on hydraulic characteristics of buried pipe network. *Applied Thermal Engineering*, *182*. https://doi.org/10.1016/j.applthermaleng.2020.116069

A Step-by-Step Explanation of Principal Component Analysis (PCA) | Built In. (n.d.). Retrieved April 17, 2022, from https://builtin.com/data-science/step-step-explanation-principal-component-analysis

Alvisi, S., & Franchini, M. (2010). Pipe roughness calibration in water distribution systems using grey numbers. *Journal of Hydroinformatics*, *12*(4), 424–445. https://doi.org/10.2166/HYDRO.2010.089

Alvisi, S., & Franchini, M. (2014). A Procedure for the Design of District Metered Areas in Water Distribution Systems. *Procedia Engineering*, *70*, 41–50. https://doi.org/10.1016/J.PROENG.2014.02.006

Armand, H., Stoianov, I., & Graham, N. (2018). Impact of network sectorisation on water quality management. *Journal of Hydroinformatics*, *20*(2), 424–439. https://doi.org/10.2166/HYDRO.2017.072

Ausiello, G., Firmani, D., & Laura, L. (2012). Real-time monitoring of undirected networks: Articulation points, bridges, and connected and biconnected components. *Networks*, *59*(3), 275–288. https://doi.org/10.1002/NET.21450

Awe, O. M., Okolie, S. T. A., & Fayomi, O. S. I. (2019). Optimization of Water Distribution Systems: A Review. *Journal of Physics: Conference Series*, *1378*(2). https://doi.org/10.1088/1742-6596/1378/2/022068

Awumah, K., Goulter, I., & Bhatt, S. K. (1990). Assessment of reliability in water distribution networks using entropy based measures. *Stochastic Hydrology and Hydraulics 1990 4:4*, *4*(4), 309–320. https://doi.org/10.1007/BF01544084

Bui, X. K., Marlim, M. S., & Kang, D. (2020). Water network partitioning into district metered areas: A state-of-the-art review. *Water (Switzerland)*, *12*(4). https://doi.org/10.3390/W12041002

Bui, X. K., Marlim, M. S., & Kang, D. (2021). Optimal design of district metered areas in a water distribution network using coupled self-organizing map and community structure algorithm. *Water (Switzerland)*, *13*(6). https://doi.org/10.3390/W13060836

Cassidy, J., Barbosa, B., Damião, M., Ramalho, P., Ganhão, A., Santos, A., & Feliciano, J. (2021). Taking water efficiency to the next level: Digital tools to reduce non-revenue water. *Journal of Hydroinformatics*, *23*(3), 453–465. https://doi.org/10.2166/HYDRO.2020.072

Chang, D. E., Yoo, D. G., & Kim, J. H. (2021). A study on the practical pressure-driven hydraulic analysis method considering actual water supply characteristics of water distribution network. *Sustainability (Switzerland)*, *13*(5). https://doi.org/10.3390/su13052793

Chondronasios, A., Gonelas, K., Kanakoudis, V., Patelis, M., & Korkana, P. (2017). Optimizing DMAs' formation in a water pipe network: The water aging and the operating pressure factors. *Journal of Hydroinformatics*, *19*(6), 890–899. https://doi.org/10.2166/HYDRO.2017.156

citypopulation. (2022). *Netherlands: Administrative Division (Provinces and Municipalities) - Population Statistics, Charts and Map*. https://www.citypopulation.de/en/netherlands/admin/

Clark, R. M. (1992). Water Quality Modeling in Distribution Systems. *Journal of Environmental Science and Health. Part A: Environmental Science and Engineering and Toxicology*, *27*(5), 1329–1366. https://doi.org/10.1080/10934529209375799

Closeness centrality | Network analysis of protein interaction data. (n.d.). Retrieved April 12, 2022, from https://www.ebi.ac.uk/training/online/courses/network-analysis-of-protein-interaction-data-an-introduction/building-and-analysing-ppins/topological-ppin-analysis/closeness-centrality/

de Paola, F., Fontana, N., Galdiero, E., Giugni, M., Degli Uberti, G. S., & Vitaletti, M. (2014). Optimal design of district metered areas in water distribution networks. *Procedia Engineering*, *70*, 449–457. https://doi.org/10.1016/J.PROENG.2014.02.050

Diao, K., Berardi, L., Laucelli, D. B., Ulanicki, B., & Giustolisi, O. (2022). Topological and hydraulic metrics-based search space reduction for optimal re-sizing of water distribution networks. *Journal of Hydroinformatics*, *00*, 1. https://doi.org/10.2166/HYDRO.2022.158

Diao, K., Farmani, R., Fu, G., & Butler, D. (2014). Vulnerability Assessment Of Water Distribution Systems Using Directed And Undirected Graph Theory. *International Conference on Hydroinformatics*. https://academicworks.cuny.edu/cc_conf_hic/267

Vargas, K., & Saldarriaga, J. (2019). Potential district metered area identification in water distribution networks using hydraulic criteria and the box covering algorithm. *World Environmental and Water Resources Congress 2019: Hydraulics, Waterways, and Water Distribution Systems Analysis - Selected Papers from the World Environmental and Water Resources Congress 2019*, 573–586. https://doi.org/10.1061/9780784482353.055

Water Network Tool for Resilience (WNTR) User Manual: Version 0.2.3 | Science Inventory | US EPA. (n.d.). Retrieved April 13, 2022, from https://cfpub.epa.gov/si/si_public_record_Report.cfm?dirEntryId=349798&Lab=CESER

WNTR User Manual — WNTR 0.1.4 documentation. (n.d.). Retrieved March 31, 2022, from https://dbhart-wntr.readthedocs.io/en/metrics/

Wright, R., Stoianov, I., Parpas, P., Henderson, K., & King, J. (2014). Adaptive water distribution networks with dynamically reconfigurable topology. *Journal of Hydroinformatics*, *16*(6), 1280–1301. https://doi.org/10.2166/HYDRO.2014.086

Wu, Z., Boulos, P., de Schaetzen, W., Orr, C. H., & Moore, M. (2004). Using genetic algorithms for water distribution system optimization. *Bridging the Gap: Meeting the World's Water and Environmental Resources Challenges - Proceedings of the World Water and Environmental Resources Congress 2001*, *111*. https://doi.org/10.1061/40569(2001)409

Yazdani, A., & Jeffrey, P. (2011). Complex network analysis of water distribution systems. *Chaos*, *21*(1). https://doi.org/10.1063/1.3540339

Zarghami, S., Gunawan, I., & Schultmann, F. (2019). Integrating Topological and Hydraulic Attributes for Robustness Analysis of Water Distribution Networks. *International Journal of Industrial Engineering and Operations Management*, *01*(01), 1–11. https://doi.org/10.46254/j.ieom.20190101

Zhang, M., Gong, G., & Zeng, L. (2021). Investigation for a novel optimization design method of ground source heat pump based on hydraulic characteristics of buried pipe network. *Applied Thermal Engineering*, *182*. https://doi.org/10.1016/j.applthermaleng.2020.116069

A Step-by-Step Explanation of Principal Component Analysis (PCA) | Built In. (n.d.). Retrieved April 17, 2022, from https://builtin.com/data-science/step-step-explanation-principal-component-analysis

Alvisi, S., & Franchini, M. (2010). Pipe roughness calibration in water distribution systems using grey numbers. *Journal of Hydroinformatics*, *12*(4), 424–445. https://doi.org/10.2166/HYDRO.2010.089

Alvisi, S., & Franchini, M. (2014). A Procedure for the Design of District Metered Areas in Water Distribution Systems. *Procedia Engineering*, *70*, 41–50. https://doi.org/10.1016/J.PROENG.2014.02.006

Armand, H., Stoianov, I., & Graham, N. (2018). Impact of network sectorisation on water quality management. *Journal of Hydroinformatics*, *20*(2), 424–439. https://doi.org/10.2166/HYDRO.2017.072

Ausiello, G., Firmani, D., & Laura, L. (2012). Real-time monitoring of undirected networks: Articulation points, bridges, and connected and biconnected components. *Networks*, *59*(3), 275–288. https://doi.org/10.1002/NET.21450

Awe, O. M., Okolie, S. T. A., & Fayomi, O. S. I. (2019). Optimization of Water Distribution Systems: A Review. *Journal of Physics: Conference Series*, *1378*(2). https://doi.org/10.1088/1742-6596/1378/2/022068

Awumah, K., Goulter, I., & Bhatt, S. K. (1990). Assessment of reliability in water distribution networks using entropy based measures. *Stochastic Hydrology and Hydraulics 1990 4:4*, *4*(4), 309–320. https://doi.org/10.1007/BF01544084

Bui, X. K., Marlim, M. S., & Kang, D. (2020). Water network partitioning into district metered areas: A state-of-the-art review. *Water (Switzerland)*, *12*(4). https://doi.org/10.3390/W12041002

Bui, X. K., Marlim, M. S., & Kang, D. (2021). Optimal design of district metered areas in a water distribution network using coupled self-organizing map and community structure algorithm. *Water (Switzerland)*, *13*(6). https://doi.org/10.3390/W13060836

Cassidy, J., Barbosa, B., Damião, M., Ramalho, P., Ganhão, A., Santos, A., & Feliciano, J. (2021). Taking water efficiency to the next level: Digital tools to reduce non-revenue water. *Journal of Hydroinformatics*, *23*(3), 453–465. https://doi.org/10.2166/HYDRO.2020.072

Chang, D. E., Yoo, D. G., & Kim, J. H. (2021). A study on the practical pressure-driven hydraulic analysis method considering actual water supply characteristics of water distribution network. *Sustainability (Switzerland)*, *13*(5). https://doi.org/10.3390/su13052793

Chondronasios, A., Gonelas, K., Kanakoudis, V., Patelis, M., & Korkana, P. (2017). Optimizing DMAs' formation in a water pipe network: The water aging and the operating pressure factors. *Journal of Hydroinformatics*, *19*(6), 890–899. https://doi.org/10.2166/HYDRO.2017.156

citypopulation. (2022). *Netherlands: Administrative Division (Provinces and Municipalities) - Population Statistics, Charts and Map*. https://www.citypopulation.de/en/netherlands/admin/

Clark, R. M. (1992). Water Quality Modeling in Distribution Systems. *Journal of Environmental Science and Health. Part A: Environmental Science and Engineering and Toxicology*, *27*(5), 1329–1366. https://doi.org/10.1080/10934529209375799

Closeness centrality | Network analysis of protein interaction data. (n.d.). Retrieved April 12, 2022, from https://www.ebi.ac.uk/training/online/courses/network-analysis-of-protein-interaction-data-an-introduction/building-and-analysing-ppins/topological-ppin-analysis/closeness-centrality/

de Paola, F., Fontana, N., Galdiero, E., Giugni, M., Degli Uberti, G. S., & Vitaletti, M. (2014). Optimal design of district metered areas in water distribution networks. *Procedia Engineering*, *70*, 449–457. https://doi.org/10.1016/J.PROENG.2014.02.050

Diao, K., Berardi, L., Laucelli, D. B., Ulanicki, B., & Giustolisi, O. (2022). Topological and hydraulic metrics-based search space reduction for optimal re-sizing of water distribution networks. *Journal of Hydroinformatics*, *00*, 1. https://doi.org/10.2166/HYDRO.2022.158

Diao, K., Farmani, R., Fu, G., & Butler, D. (2014). Vulnerability Assessment Of Water Distribution Systems Using Directed And Undirected Graph Theory. *International Conference on Hydroinformatics*. https://academicworks.cuny.edu/cc_conf_hic/267

Vargas, K., & Saldarriaga, J. (2019). Potential district metered area identification in water distribution networks using hydraulic criteria and the box covering algorithm. *World Environmental and Water Resources Congress 2019: Hydraulics, Waterways, and Water Distribution Systems Analysis - Selected Papers from the World Environmental and Water Resources Congress 2019*, 573–586. https://doi.org/10.1061/9780784482353.055

Water Network Tool for Resilience (WNTR) User Manual: Version 0.2.3 | Science Inventory | US EPA. (n.d.). Retrieved April 13, 2022, from https://cfpub.epa.gov/si/si_public_record_Report.cfm?dirEntryId=349798&Lab=CESER

WNTR User Manual — WNTR 0.1.4 documentation. (n.d.). Retrieved March 31, 2022, from https://dbhart-wntr.readthedocs.io/en/metrics/

Wright, R., Stoianov, I., Parpas, P., Henderson, K., & King, J. (2014). Adaptive water distribution networks with dynamically reconfigurable topology. *Journal of Hydroinformatics*, *16*(6), 1280–1301. https://doi.org/10.2166/HYDRO.2014.086

Wu, Z., Boulos, P., de Schaetzen, W., Orr, C. H., & Moore, M. (2004). Using genetic algorithms for water distribution system optimization. *Bridging the Gap: Meeting the World's Water and Environmental Resources Challenges - Proceedings of the World Water and Environmental Resources Congress 2001*, *111*. https://doi.org/10.1061/40569(2001)409

Yazdani, A., & Jeffrey, P. (2011). Complex network analysis of water distribution systems. *Chaos*, *21*(1). https://doi.org/10.1063/1.3540339

Zarghami, S., Gunawan, I., & Schultmann, F. (2019). Integrating Topological and Hydraulic Attributes for Robustness Analysis of Water Distribution Networks. *International Journal of Industrial Engineering and Operations Management*, *01*(01), 1–11. https://doi.org/10.46254/j.ieom.20190101

Zhang, M., Gong, G., & Zeng, L. (2021). Investigation for a novel optimization design method of ground source heat pump based on hydraulic characteristics of buried pipe network. *Applied Thermal Engineering*, *182*. https://doi.org/10.1016/j.applthermaleng.2020.116069

A Step-by-Step Explanation of Principal Component Analysis (PCA) | Built In. (n.d.). Retrieved April 17, 2022, from https://builtin.com/data-science/step-step-explanation-principal-component-analysis

Alvisi, S., & Franchini, M. (2010). Pipe roughness calibration in water distribution systems using grey numbers. *Journal of Hydroinformatics*, *12*(4), 424–445. https://doi.org/10.2166/HYDRO.2010.089

Alvisi, S., & Franchini, M. (2014). A Procedure for the Design of District Metered Areas in Water Distribution Systems. *Procedia Engineering*, *70*, 41–50. https://doi.org/10.1016/J.PROENG.2014.02.006

Armand, H., Stoianov, I., & Graham, N. (2018). Impact of network sectorisation on water quality management. *Journal of Hydroinformatics*, *20*(2), 424–439. https://doi.org/10.2166/HYDRO.2017.072

Ausiello, G., Firmani, D., & Laura, L. (2012). Real-time monitoring of undirected networks: Articulation points, bridges, and connected and biconnected components. *Networks*, *59*(3), 275–288. https://doi.org/10.1002/NET.21450

Awe, O. M., Okolie, S. T. A., & Fayomi, O. S. I. (2019). Optimization of Water Distribution Systems: A Review. *Journal of Physics: Conference Series*, *1378*(2). https://doi.org/10.1088/1742-6596/1378/2/022068

Awumah, K., Goulter, I., & Bhatt, S. K. (1990). Assessment of reliability in water distribution networks using entropy based measures. *Stochastic Hydrology and Hydraulics 1990 4:4*, *4*(4), 309–320. https://doi.org/10.1007/BF01544084

Bui, X. K., Marlim, M. S., & Kang, D. (2020). Water network partitioning into district metered areas: A state-of-the-art review. *Water (Switzerland)*, *12*(4). https://doi.org/10.3390/W12041002

Bui, X. K., Marlim, M. S., & Kang, D. (2021). Optimal design of district metered areas in a water distribution network using coupled self-organizing map and community structure algorithm. *Water (Switzerland)*, *13*(6). https://doi.org/10.3390/W13060836

Cassidy, J., Barbosa, B., Damião, M., Ramalho, P., Ganhão, A., Santos, A., & Feliciano, J. (2021). Taking water efficiency to the next level: Digital tools to reduce non-revenue water. *Journal of Hydroinformatics*, *23*(3), 453–465. https://doi.org/10.2166/HYDRO.2020.072

Chang, D. E., Yoo, D. G., & Kim, J. H. (2021). A study on the practical pressure-driven hydraulic analysis method considering actual water supply characteristics of water distribution network. *Sustainability (Switzerland)*, *13*(5). https://doi.org/10.3390/su13052793

Chondronasios, A., Gonelas, K., Kanakoudis, V., Patelis, M., & Korkana, P. (2017). Optimizing DMAs' formation in a water pipe network: The water aging and the operating pressure factors. *Journal of Hydroinformatics*, *19*(6), 890–899. https://doi.org/10.2166/HYDRO.2017.156

citypopulation. (2022). *Netherlands: Administrative Division (Provinces and Municipalities) - Population Statistics, Charts and Map*. https://www.citypopulation.de/en/netherlands/admin/

Clark, R. M. (1992). Water Quality Modeling in Distribution Systems. *Journal of Environmental Science and Health. Part A: Environmental Science and Engineering and Toxicology*, *27*(5), 1329–1366. https://doi.org/10.1080/10934529209375799

Closeness centrality | Network analysis of protein interaction data. (n.d.). Retrieved April 12, 2022, from https://www.ebi.ac.uk/training/online/courses/network-analysis-of-protein-interaction-data-an-introduction/building-and-analysing-ppins/topological-ppin-analysis/closeness-centrality/

de Paola, F., Fontana, N., Galdiero, E., Giugni, M., Degli Uberti, G. S., & Vitaletti, M. (2014). Optimal design of district metered areas in water distribution networks. *Procedia Engineering*, *70*, 449–457. https://doi.org/10.1016/J.PROENG.2014.02.050

Diao, K., Berardi, L., Laucelli, D. B., Ulanicki, B., & Giustolisi, O. (2022). Topological and hydraulic metrics-based search space reduction for optimal re-sizing of water distribution networks. *Journal of Hydroinformatics*, *00*, 1. https://doi.org/10.2166/HYDRO.2022.158

Diao, K., Farmani, R., Fu, G., & Butler, D. (2014). Vulnerability Assessment Of Water Distribution Systems Using Directed And Undirected Graph Theory. *International Conference on Hydroinformatics*. https://academicworks.cuny.edu/cc_conf_hic/267

Dien, J. (2012). Applying Principal Components Analysis to Event-Related Potentials: A Tutorial. *Https://Doi.Org/10.1080/87565641.2012.697503*, *37*(6), 497–517. https://doi.org/10.1080/87565641.2012.697503

DTK HYDRONET. (2019, December 6). *What are District Metered Areas (DMA)?* https://www.dtkhydronet.com/post/what-are-district-metered-areas-dma

Fu, G., Kapelan, Z., & Reed, P. (2012). Reducing the Complexity of Multiobjective Water Distribution System Optimization through Global Sensitivity Analysis. *Journal of Water Resources Planning and Management*, *138*(3), 196–207. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000171

Galdiero, E., de Paola, F., Fontana, N., Giugni, M., & Savic, D. (2016). Decision support system for the optimal design of district metered areas. *Journal of Hydroinformatics*, *18*(1), 49–61. https://doi.org/10.2166/HYDRO.2015.023

Giustolisi, O., Ridolfi, L., & Simone, A. (2019). Tailoring Centrality Metrics for Water Distribution Networks. *Water Resources Research*, *55*(3), 2348–2369. https://doi.org/10.1029/2018WR023966

Gondwana - KWR. (n.d.). Retrieved April 13, 2022, from https://www.kwrwater.nl/en/tools-producten/gondwana/

Gonelas, K., Chondronasios, A., Kanakoudis, V., Patelis, M., & Korkana, P. (2017). Forming DMAs in a water distribution network considering the operating pressure and the chlorine residual concentration as the design parameters. *Journal of Hydroinformatics*, *19*(6), 900–910. https://doi.org/10.2166/HYDRO.2017.157

Gupta, I., Gupta, A., & Khanna, P. (1999). Genetic algorithm for optimization of water distribution systems. *Environmental Modelling & Software*, *14*(5), 437–446. https://doi.org/10.1016/S1364-8152(98)00089-9

Hajebi, S., Roshani, E., Cardozo, N., Barrett, S., Clarke, A., & Clarke, S. (2016). Water distribution network sectorisation using graph theory and many-objective optimisation. *Journal of Hydroinformatics*, *18*(1), 77–95. https://doi.org/10.2166/HYDRO.2015.144

Hajibabaei, M., Nazif, S., & Sitzenfrei, R. (2019). Improving the performance of water distribution networks based on the value index in the system dynamics framework. *Water (Switzerland)*, *11*(12). https://doi.org/10.3390/W11122445

Imran, M., Siddiqui, M. K., Abunamous, A. A. E., Adi, D., Rafique, S. H., & Baig, A. Q. (2018). Eccentricity Based Topological Indices of an Oxide Network. *Mathematics 2018, Vol. 6, Page 126*, *6*(7), 126. https://doi.org/10.3390/MATH6070126

Jayaram, N., & Srinivasan, K. (2008). Performance-based optimal design and rehabilitation of water distribution networks using life cycle costing. *Water Resources Research*, *44*(1), 1417. https://doi.org/10.1029/2006WR005316

Johnson, R. P., Blackschleger, V., Boccelli, D. L., & Lee, Y. (2008). Water security initiative field study: improving confidence in a distribution system model. *World Environmental and Water Resources Congress 2008: Ahupua'a - Proceedings of the World Environmental and Water Resources Congress 2008*, *316*. https://doi.org/10.1061/40976(316)509

Jolly, M. D., Lothes, A. D., Sebastian Bryson, L., & Ormsbee, L. (2014). Research Database of Water Distribution System Models. *Journal of Water Resources Planning and Management*, *140*(4), 410–416. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000352

Kapelan, Z. S., Savic, D. A., & Walters, G. A. (2005). Multiobjective design of water distribution systems under uncertainty. *Water Resources Research*, *41*(11), 1–15. https://doi.org/10.1029/2004WR003787

Kara, S., Ethem Karadirek, I., Muhammetoglu, A., & Muhammetoglu, H. (2016). Hydraulic Modeling of a Water Distribution Network in a Tourism Area with Highly Varying Characteristics. *Procedia Engineering*, *162*. https://doi.org/10.1016/j.proeng.2016.11.096

Kourbasis, N., Patelis, M., … S. T.-E. S., & 2020, undefined. (2020). Optimizing water age and pressure in drinking water distribution networks. *Mdpi.Com*, *2020*, 51. https://doi.org/10.3390/environsciproc2020002051

Krnc, M., Sereni, J.-S., Skrekovski, R., Yilma, Z., Eccentricity, Z. Y., & Yilma, Z. B. (2018). *Eccentricity of Networks with Structural Constraints Eccentricity of Networks with Structural Constraints *. 2020*(4). https://hal.archives-ouvertes.fr/hal-01385481v2

Maciej Serda. (2013). Synteza i aktywność biologiczna nowych analogów tiosemikarbazonowych chelatorów żelaza. *Uniwersytet Śląski*, 343–354. https://doi.org/10.2/JQUERY.MIN.JS

Marinescu, D. C. (2017). Managing Complexity of Large-Scale Cyber-Physical Systems. *Complex Systems and Clouds*, 65–112. https://doi.org/10.1016/B978-0-12-804041-6.00003-7

Mean path length definition - Math Insight. (n.d.). Retrieved April 14, 2022, from https://mathinsight.org/definition/network_mean_path_length

Meng, F., Fu, G., Farmani, R., Sweetapple, C., & Butler, D. (2018). Topological attributes of network resilience: A study in water distribution systems. *Water Research*, *143*, 376–386. https://doi.org/10.1016/J.WATRES.2018.06.048

Mikulecky, P. (2011). Ambient intelligence in water management. *Proceedings - 2011 7th International Conference on Intelligent Environments, IE 2011*, 302–305. https://doi.org/10.1109/IE.2011.55

Mohammed, H., Tornyeviadzi, H. M., & Seidu, R. (2021). Modelling the impact of water temperature, pipe, and hydraulic conditions on water quality in water distribution networks. *Water Practice and Technology*, *16*(2). https://doi.org/10.2166/wpt.2021.002

Newman, M. (2010). Networks: An Introduction. *Networks: An Introduction*, 1–784. https://doi.org/10.1093/ACPROF:OSO/9780199206650.001.0001

Node Degree and Strength. (2016). *Fundamentals of Brain Network Analysis*, 115–136. https://doi.org/10.1016/B978-0-12-407908-3.00004-2

Ostfeld, A. (2015). Water distribution networks. *Studies in Computational Intelligence*, *565*, 101. https://doi.org/10.1007/978-3-662-44160-2_4

Ostfeld, A., Kogan, D., & Shamir, U. (2002). Reliability simulation of water distribution systems – single and multiquality. *Urban Water*, *4*(1), 53–61. https://doi.org/10.1016/S1462-0758(01)00055-3

Perelman, L., & Ostfeld, A. (2005). Water distribution systems optimal design using cross entropy. *GECCO 2005 - Genetic and Evolutionary Computation Conference*, 647–648. https://doi.org/10.1145/1068009.1068117

Price, E., & Ostfeld, A. (2015). Graph Theory Modeling Approach for Optimal Operation of Water Distribution Systems. *Journal of Hydraulic Engineering*, *142*(3), 04015061. https://doi.org/10.1061/(ASCE)HY.1943-7900.0001099

Priyanka Jawale, M., Iyer, V., Aavani, M. P., Professor, A., & of Civil Engineering, D. (2018). Review on Analysis of Water Distribution Network using EPANET for Normal Condition. *IJSRD-International Journal for Scientific Research & Development|*, *6*, 2321–0613. www.ijsrd.com

python - Joining two networkx graphs on a single edge - Stack Overflow. (n.d.). Retrieved April 14, 2022, from https://stackoverflow.com/questions/49427320/joining-two-networkx-graphs-on-a-single-edge

Quuen Mary. (2008). *The link density as a function of the number of nodes and the number of... | Download Scientific Diagram*. https://www.researchgate.net/figure/The-link-density-as-a-function-of-the-number-of-nodes-and-the-number-of-links-in_fig1_220519946

Ratnayaka, D. D., Brandt, M. J., & Johnson, K. M. (2009). Pipeline Design and Construction. *Water Supply*, 561–598. https://doi.org/10.1016/B978-0-7506-6843-9.00023-8

Savic·, A. D., Waiters·, G. A., Atkinson·, R. M., & Smitht, M. R. (1999). *Genetic Algorithm Optimization of Large Water Distribution System Expansion*.

Sg, P. G. (2016). Managing the water distribution network with a Smart Water Grid. *Smart Water 2016 1:1*, *1*(1), 1–13. https://doi.org/10.1186/S40713-016-0004-4

Shamsaei, H., Jaafar, O., Ezlin, N., & Basri, A. (2013). *Effects Residence Time to Water Quality in Large Water Distribution Systems*. *5*, 449–457. https://doi.org/10.4236/eng.2013.54054
Shin, S., Lee, S., Judi, D. R., Parvania, M., Goharian, E., McPherson, T., & Burian, S. J. (2018). A systematic review of quantitative resilience measures for water infrastructure systems. In *Water (Switzerland)* (Vol. 10, Issue 2). https://doi.org/10.3390/w10020164
statista. (2022). *The Hague: total population 2019*. https://www.statista.com/statistics/753256/total-population-of-the-hague/
Suribabu, C. R. (2017). Resilience-based optimal design of water distribution network. *Applied Water Science*, *7*(7). https://doi.org/10.1007/s13201-017-0560-2
Syakur, M. A., Khotimah, B. K., Rochman, E. M. S., & Satoto, B. D. (2018). Integration K-Means Clustering Method and Elbow Method For Identification of The Best Customer Profile Cluster. *IOP Conference Series: Materials Science and Engineering*, *336*(1), 012017. https://doi.org/10.1088/1757-899X/336/1/012017
Thienen, P. van, & Montiel, F. (2014). Flow Analysis And Leak Detection With The CFPD Method In The Paris Drinking Water Distribution System. *International Conference on Hydroinformatics*. https://academicworks.cuny.edu/cc_conf_hic/45
Tian, L., Bashan, A., Shi, D. N., & Liu, Y. Y. (2017). Articulation points in complex networks. *Nature Communications 2017 8:1*, *8*(1), 1–9. https://doi.org/10.1038/ncomms14223
Todini, E. (2000). Looped water distribution networks design using a resilience index based heuristic approach. *Urban Water*, *2*(2), 115–122. https://doi.org/10.1016/S1462-0758(00)00049-2
Torres, J. M., Duenas-Osorio, L., Asce, M., Li, Q., & Yazdani, A. (2016). *Exploring Topological Effects on Water Distribution System Performance Using Graph Theory and Statistical Models*. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000709
Touchette, P. E., MacDonald, R. F., & Langer, S. N. (1985). A scatter plot for identifying stimulus control of problem behavior. *Journal of Applied Behavior Analysis*, *18*(4), 343–351. https://doi.org/10.1901/JABA.1985.18-343
Tsitsifli, S., & Kanakoudis, V. (2021). Assessing the impact of dmas and the use of boosters on chlorination in a water distribution network in Greece. *Water (Switzerland)*, *13*(16). https://doi.org/10.3390/W13162141
Tzatchkov, V. G., Alcocer-Yamanaka, V. H., & Bourguett Ortíz, V. (2007). Graph theory based algorithms for water distribution network sectorization projects. *8th Annual Water Distribution Systems Analysis Symposium 2006*, 172. https://doi.org/10.1061/40941(247)172
Udhane, D., Kataria, N., Sankhe, S., Pore, T., Motegaonkar, S., & Students, U. G. (2018). Hydraulic Modeling and Simulation of Smart Water Distribution Network. *International Research Journal of Engineering and Technology*. www.irjet.net
Ulusoy, A. J., Stoianov, I., & Chazerain, A. (2018). Hydraulically informed graph theoretic measure of link criticality for the resilience analysis of water distribution networks. *Applied Network Science*, *3*(1). https://doi.org/10.1007/s41109-018-0079-y
van Laarhoven, K., Vertommen, I., van Thienen, P., de Kater, H., Schaap, P., Gardien, D., van den Boom, M., & de Groot, R. (n.d.). *OPTIMAL DESIGN OF VIRTUAL DMA'S WITH GONDWANA*.
van Thienen, P., & Vertommen, I. (2015). Gondwana: A generic optimization tool for drinking water distribution systems design and operation. *Procedia Engineering*, *119*(1), 1212–1220. https://doi.org/10.1016/J.PROENG.2015.08.978
Vargas, K., & Saldarriaga, J. (2019). Potential district metered area identification in water distribution networks using hydraulic criteria and the box covering algorithm. *World*

Environmental and Water Resources Congress 2019: Hydraulics, Waterways, and Water Distribution Systems Analysis - Selected Papers from the World Environmental and Water Resources Congress 2019, 573–586. https://doi.org/10.1061/9780784482353.055

Water Network Tool for Resilience (WNTR) User Manual: Version 0.2.3 | Science Inventory | US EPA. (n.d.). Retrieved April 13, 2022, from https://cfpub.epa.gov/si/si_public_record_Report.cfm?dirEntryId=349798&Lab=CESER

WNTR User Manual — WNTR 0.1.4 documentation. (n.d.). Retrieved March 31, 2022, from https://dbhart-wntr.readthedocs.io/en/metrics/

Wright, R., Stoianov, I., Parpas, P., Henderson, K., & King, J. (2014). Adaptive water distribution networks with dynamically reconfigurable topology. *Journal of Hydroinformatics*, *16*(6), 1280–1301. https://doi.org/10.2166/HYDRO.2014.086

Wu, Z., Boulos, P., de Schaetzen, W., Orr, C. H., & Moore, M. (2004). Using genetic algorithms for water distribution system optimization. *Bridging the Gap: Meeting the World's Water and Environmental Resources Challenges - Proceedings of the World Water and Environmental Resources Congress 2001*, *111*. https://doi.org/10.1061/40569(2001)409

Yazdani, A., & Jeffrey, P. (2011). Complex network analysis of water distribution systems. *Chaos*, *21*(1). https://doi.org/10.1063/1.3540339

Zarghami, S., Gunawan, I., & Schultmann, F. (2019). Integrating Topological and Hydraulic Attributes for Robustness Analysis of Water Distribution Networks. *International Journal of Industrial Engineering and Operations Management*, *01*(01), 1–11. https://doi.org/10.46254/j.ieom.20190101

Zhang, M., Gong, G., & Zeng, L. (2021). Investigation for a novel optimization design method of ground source heat pump based on hydraulic characteristics of buried pipe network. *Applied Thermal Engineering*, *182*. https://doi.org/10.1016/j.applthermaleng.2020.116069

Appendices

Appendix A. - Research Ethics Declaration Form

As from 2019, the IHE Delft MSc Students have to sign, submit and discuss with their mentors/supervisors a ‘Research Ethics Declaration Form’ (this form will be provided to all MSc students in October 2019). The Academic Board will specify who will endorse or approve this declaration.
The ‘Research Ethics Declaration Form’ aims to encourage all IHE Delft MSc Students to reflect on the potential ethics issues in their research proposal, and later in their thesis. **All MSc students need to read the** Netherlands Code of Conduct for Research Integrity 2018 before signing this declaration.