Aalborg University

Identification of Randers Water Distribution Network for Optimal Control

Electronic & IT:
Control & Automation

Group:
CA9-938

Student report

December 11, 2017

AALBORG UNIVERSITY
STUDENT REPORT

Second year of MSc study
Electronic and IT
Fredrik Bajers Vej 7
DK-9220 Aalborg East, Denmark
http://www.es.aau.dk

Topic:

Optimal Control for Water Distribution

Project:

P9-project

Project time:

September 2017 - December 2017

Projectgroup:

17gr938

Participants:

Krisztian Mark Balla

Supervisors:

Tom Nørgaard Jensen
Jan Dimon Bendsten
Carsten Skovmose Kallesøe

Synopsis:

The synopsis is going to be written here...

Circulation: x
Number of pages: x
Appendix: x
Completed xx-xx-2017

Preface

The preface is going to be written here... The preface is going to be written here... The preface is going to be written here... The preface is going to be written here... The preface is going to be written here... The preface is going to be written here... The preface is going to be written here... The preface is going to be written here... The preface is going to be written here... The preface is going to be written here... The preface is going to be written here...

Aalborg University, th of December 2017

Krisztian Mark Balla
kballa16@student.aau.dk

Nomenclature

Acronyms

ANN	Artificial Neural Network
CT	Circuit Theory
D-W	Darcy-Weisbach
EPA	Environmental Protection Agency
FCV	Flow Control Valve
GIS	Geographic Information Systems
GT	Graph Theory
KCL	Kirchhoff's current law
LS	Least squares
MPC	Model Predictive Control
NN	Neural Network
PCA	Principal Component Analysis
PRV	Pressure Regulating Valve
RBF	Radial Basis Function
RBFNN	Radial Basis Function Neural Network
SCADA	Supervisory Control And Data Acquisition
SS	State-Space
TS	Training Set
OD	Opening Degree
ODE	Ordinary Differential Equation
WSS	Water Supply System
WT	Water Tank

Acronyms - Randers Network

BKV	Bunkedal Water Work
LZ	Low Zone
HBP	Hobrovej Pumping Station
HNP	Hornbæk Pumping Station
HSP	Hadsundvej Pumping Station
HZ	High Zone
OMV	Oust Mølle Water Work
TA	Water tank A in Hobrovej
TB	Water tank B in Hobrovej
TC	Water tank C in Hadsundvej
TBP	Toldbodgade Pumping Station
VSV	Vilstrup Water Work
ØSV	Østrup Skov Water Work

Symbols

Symbol	Description	Unit
A_{wt}	Cross sectional area of water tanks	$[m^2]$
a_{h2}, a_{h1}, a_{h0}	Pump constants	$[\cdot]$
c_D	Darcy-Weisbach equation coefficient	$[s^2/m]$
D	Diameter	$[m]$
d	Flow demand	$[m^3/s]$
$f(q)$	Pressure drop due to pipe resistance	$[Pa]$
h	Pressure drop due to elevation	$[Pa]$
h_l	Water level in tanks	$[m]$
h_p	Pressure head	$[m]$
h_t	Total head	$[m]$
J	Mass inertia of water pipes	$[kgm^2]$
k_v	Valve conductivity function	$[\cdot]$
l	Length (of pipes)	$[m]$
p	Absolute pressure	$[Pa]$
q	Volumetric flow	$[m^3/s]$
Re	Reynolds number	$[\cdot]$
T	Period of time	$[s]$
z	Elevation head	$[m]$
γ	Resistance parameter of pipes	$[\cdot]$
Δp	Differential pressure	$[Pa]$
ϵ	Roughness of pipes	$[m]$
$\mu(q, k_v)$	Pressure drop on valves	$[Pa]$
ω_r	Impeller rotational speed of centrifugal pumps	$[rad/s]$
τ	Elevated reservoir parameter	$[Pa/m^3]$

Constants

Symbol	Description	Unit
$g = 9.83$	Gravitational acceleration	$[m/s^2]$
$\rho = 1000$	Density of water	$[kg/m^3]$
$f_D = 0.05$	Darcy friction factor	$[kg/m^3]$

Graph theory

Symbol	Description
B	Cycle matrix
$\mathcal{E}$	Set of edges
$\mathcal{E}_\mathcal{T}$	Edges regarding the sub-graph $\mathcal{T}$
$\mathcal{E}_\mathcal{C}$	Edges regarding the sub-graph $\mathcal{C}$
I	Identity matrix
$\mathcal{G}$	Directed and connected graph
H	Incidence matrix
$\mathcal{V}$	Set of vertices
$\bar{\mathcal{V}}$	Vertices regarding non-inlet points
$\hat{\mathcal{V}}$	Vertices regarding inlet points
c	Number of pumping stations
l	Number of elevated reservoirs
m	Number of columns in the incidence matrix
n	Number of rows in the incidence matrix
$\mathcal{T}$	General sub-graph
$\mathcal{T}^*$	Tree in a graph
$\mathcal{T}^*_{span}$	Spanning tree in a graph

Glossary of mathematical notation

Matrices

Diagonal matrices are noted with $diag(\cdot)$, which maps an n-tuple to the corresponding diagonal matrix:

$$diag : \mathbb{R}^n \to \mathbb{R}^{n \times n},$$

$$diag(a_1, \ldots, a_n) := \begin{pmatrix} a_1 & & \\ & \ddots & \\ & & a_n \end{pmatrix}.$$

A positive semi-definite (resp. positive definite) matrix A is is denoted $A \succeq 0$ ($resp.$ $A \succ 0$).

Contents

1. Introduction

Due to the fast-paced technological development all over the world, the demand for industrial growth and for energy resources has seen a rapid increase. Along with the industrial growth, the sudden rise in population has made the world realize that shortage of energy sources is an actual and universally anticipated problem [1]. In order to cope with such shortage issues and to make the rapid development possible and less expensive, the world is moving towards more efficient use of resources and optimization of the infrastructure. Therefore, technological development is also moving the focus to green energy and to solutions for more optimized operation of the existing infrastructure, resulting in more and more renewable energy sources added to the grid and smarter solutions for control [2].

Water Supply Systems(WSSs) are among the sectors which make the industrial growth possible. On top of this, WSSs are one of the most vital infrastructures of modern societies in the world. In Denmark typically, such networks are operating by making pumps transport water from reservoirs through the pipe network, to the consumers. In most cases, elevated reservoirs are utilized in these WSSs, such that they can even out the demand differences for the consumers. Although elevated reservoirs are usually an integrated part of these systems, providing drinking water is a highly energy-intensive activity. For instance, in the United States alone, the drinking water and waste water systems are typically the largest energy consumers, accounting for 25 to 40 percent of a municipality's total public expenditure. [3].

Since fresh water is limited, and due to the presence of global changes such as climate change and urbanization, new trends are emerging in the water supply sector. In the past few decades, several research and case study showed that the operation of WSSs and other energy distribution networks need to be improved due to the leakages in the system, high cost of maintenance and due to high energy consumption. Companies which are operating WSSs also realized that by using proper pressure management in their networks, the effect of leakages can be reduced, thereby huge amount of fresh water and money can be saved annually[4].

In Denmark recently, the larger water suppliers have been focusing on making the water supply sector more effective through introducing a benchmarking system. This system is focusing on the environment, the security of supply and the efficiency based on user demands. Since 1980, these efficiency activities has been an important issue [5]. It has been proved that by utilizing advanced, energy- or cost-optimizing control solutions and utilizing renewable energy sources such as elevated reservoirs, the lifetime of the existing infrastructure can be extended and money or energy can be saved [1]. Therefore, there is a growing demand in industry for developing methods, leading towards more efficient WSSs.

The presented project is executed in collaboration with the company, Verdo A/S. It is in the interest of Verdo A/S to utilize an advanced model-based optimal control scheme on the WSS with several storages in their system in Randers, Denmark. For a large municipality such as Randers, the water distribution network is complex and consists of thousands of hydraulic elements. Since an online optimizing control algorithm is considered to be complex, the computational effort of such algorithms is also high. Furthermore, offline optimisation of large-scale WSSs means that any modifications in the existing network may require significant changes in the optimisation method, which leads to high costs of system maintenance [6]. In fact, modifications in a water distribution network are not unusual, as with time passing by, more and more residential areas are being covered, requiring the extension of the existing infrastructure. Therefore, a well-describing, reduced model is required, on which the capability of executing complex, online control algorithms is a good possibility.

The long-term goal of this thesis is to find a solution for implementing Model Predictive Control(MPC) on the Randers WSS. However, before the implementation of any control scheme would be possible, a reduced and well-describing, identified model is required. Therefore, as the main goal of the project, the following problem statement can be formulated:

How can the WSS in Randers be simplified and identified, with storages included in the system, such that the reduced model preserves the original non-linear behaviour and remains suitable for a plug-and-play commissionable Model Predictive Control scheme.

Part I

Analysis

2. Description of Water Supply Systems

This chapter gives a general overview of hydraulic systems and an introduction to the WSS in Randers. The typical topologies, components and structures of water supply networks are discussed. In the Randers WSS, the different zones, pumping stations and waterworks are introduced and their operation purposes are explained.

2.1 Hydraulic system overview

WSSs are designed to deliver water to consumers in terms of sufficient pressure and appropriate chemical composition. Distribution systems as such are typically transport water from one geographical place to another. In practice, there are different methods exist to achieve this water transport. One example is the use of natural advantages such as the water stored in mountains, and thereby use the potential energy of the water to provide pressure in the supply area. Good examples for this are countries like Norway where the advantages of the landscape are being utilized [7].

However, considering the fact that in Denmark typically water is tapped from the ground, such an advantage of the landscape cannot be utilized, unless the water is first pumped to high elevations. Furthermore, it is worth noting that the quality of groundwater, in general, is very good in Denmark, therefore it is used for drinking water supply purposes. After tapping the water, it goes through an aeration process at the waterworks and afterwards the pure water is pumped into the network [8]. In WSSs, pumps and valves are the elements that enable the control and thereby the proper delivery of water to the consumers or to elevated reservoirs, storing water for later use. Such a network is illustrated in the figure below

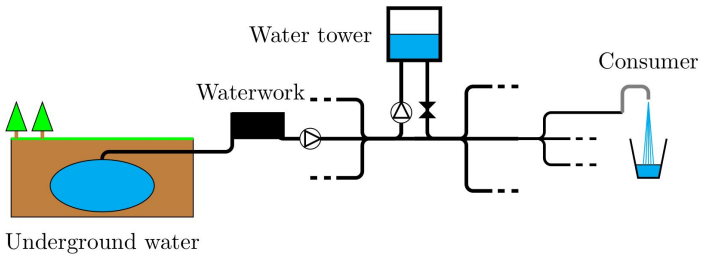


Figure 2.1: Illustration of a WSS [9].

The delivered water needs to fulfil a certain pressure criteria in order to reach consumers at higher levels. For example, in some cases the pressure has to be high enough to make it to the fourth floor of a building and still provide appropriate pressure in the water taps. Generally, in such cases booster pumps are placed in the basement of buildings, helping to supply the pressure. Too large pressure values, however increase water losses due to pipe waste [10].

Another criteria is that the flow through particular pipes need to stay within acceptable limits. A low flow rate can lead to water quality problems due to the undesirable microorganisms in the water and due to the metal and salt accumulation on the wall of the pipes [10].

As can be seen in *Figure 2.1*, typically WSSs consist of pipe, valve, reservoir, elevated reservoir(tank) and pump components. The common property of them is that they are all two-terminal components, therefore they can be characterized by the dynamic relationship

between the pressure drop across their two corresponding endpoints and the flow through them [11].

2.1.1 Pipe networks

Pipes have a major role in WSSs since they are used for carrying pressurized water. They serve as a connection between components. Normally, the pipe network can be split into different sub-parts, taking into account the physical characteristics and the attributes of the pipes. Therefore, water supply networks can consist of transmission mains, arterial mains, distribution mains and service lines as shown in the example below:

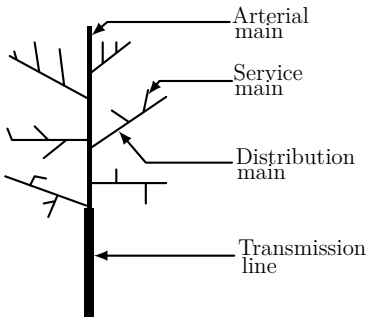


Figure 2.2: Illustration of pipe mains. Tree configuration.

Transmission mains deliver large amounts of water over long distances. Arterial and distribution mains provide intermediate steps towards delivering water to the end-users. Service lines transmit the water from the distribution mains straight to the end-users [12].

The transmission and distribution network can have a topology that is called a loop or a tree structure. *Figure 2.2* shows an example for a tree configuration. This type of configuration is most frequently used in rural areas [13]. Typically the network has only one path for the water to reach the end-users. A more frequent problem compared to looped configurations is, that on the outer parts of the system lower pressures can be experienced due to the pressure losses from long flow paths. The flow dynamics within this kind of systems therefore consist of large flows closer to the source that turn into smaller flows on the outer parts of the system. Main disadvantage of a purely tree structure system is that due to maintenance or momentary breakdowns, the system suffers disruption of service [13].

Loop networks have a configuration as shown in *Figure 2.3*.

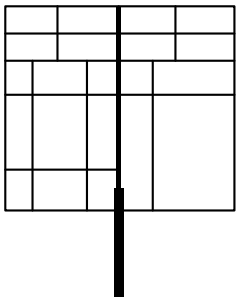


Figure 2.3: Loop configuration.

Loop networks are usually composed of smaller loops which are composed of smaller distribution mains, and larger loops that are connected to arterial or transmission mains. Elevated reservoirs are typically placed in the centre of the system due to pressure losses resulting from flows through the loop network [14]. This is reasonable because within a certain grid, the same pressure is provided by the tank, instead of providing the pressure

through long pipelines to different distances. Furthermore, in the presence of a ring structure, the large loop around the area may be used to feed an internal distribution grid or a distribution grid attached to the outer part of the loop. Loop configurations are generally associated with larger suburban and city distribution systems such as larger cities[14]. The Randers WSS falls into this category.

2.1.2 Elevated reservoirs

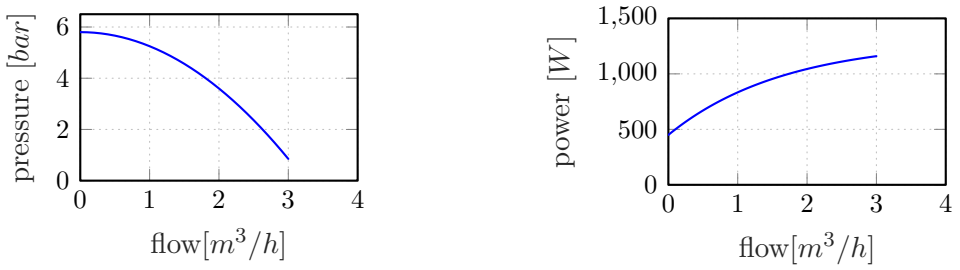
Elevated reservoirs, or tanks, are typically placed in the system to use them as buffers and level out the pressure and flow demand differences. When the demand is high, the waterworks might not be able to provide the sufficient amount of water in the network. In these cases, the elevated reservoir supplies the remaining demand. When the user consumption decreases, the system can be controlled such that the tank is being refilled to provide the required demand for the next peak time of consumption. Having such an elevated reservoir in the network, the system becomes more independent of the pump stations, as the refilled tank can itself maintain the desired pressure and flow for a limited time.

Due to the elevation of the tank, when it is filled up, the pumping stations need to provide a pressure higher than the pressure in the water tank(WT). Therefore when the tank is being emptied, the pumping stations can reduce the amount of pressure they provide to the system, since the pressure from the elevated reservoir becomes dominant. This is due to the fact that the dynamics of systems with large storages come primarily from the pressure of the tank [15]. However, it should be noted that normally the level in the tank is varying less than a meter. This means that the effect on the pump operation is limited. Due to these considerations, the dynamics of these elevated reservoirs has to be taken into account while modelling the system.

2.1.3 Pumps

Water pumps are used to increase pressure in hydraulic systems, thus making the water flow. Pumps are typically the main actuators of a WSS and they can be either flow or pressure controlled. Therefore, pumps can have controllers to produce a desired flow or pressure. This is done by changing the rotational speed of the pump. In this way, when the pump has a reference pressure or flow, simple control makes it possible to produce the desired flow or pressure respectively [16].The pressure required to make the water reach some height is the sum of the pressure required to overcome the elevation and the friction losses in the pipe network.

The most common pumps in WSSs are centrifugal pumps. Normally, the characteristics of such pumps are described by two pump curves. The two curves depict the volume flow versus the pressure and the power of the pump respectively. Normally the curves describe the characteristics for one particular speed, which is denoted the nominal speed [16]. An example of these pump curves is shown in *Figure 2.4*.



(a) Flow versus pressure difference

(b) Flow versus power consumption

Figure 2.4: Pump curves describing the performance of a centrifugal pump at nominal speed.

As can be seen, at a given flow, the pump can deliver a pressure with a maximum limit. This pressure decreases when the flow is increasing. At a certain flow and pressure value, the pump has an optimal point where the operation is the most energy efficient. Pumps are normally designed such that the optimal point lies in the operational area for the pumping application [9].

As in almost all WSSs, the flow is varying in the system, according to the flow demand from the end-consumers. Therefore, when dealing with varying flow in the system, pumps are often placed parallel at the pump stations such that they can keep their optimum points. As the flow increases, more pumps get activated to keep the pressure constant [9].

2.1.4 Valves

Valves in the WSS can be also seen as actuators along with pump elements. Unlike the pumps, valves are passive actuators in the sense that they do not consume energy. In principle, there are many types of valves existing. They can be categorized as non-return valves, control valves, shut-off valves and the combination of the two former one. Non-return valves allow waterflow only in one direction, while control valves can either adjust the flow or the pressure on their two endpoints. The former category is typically called a Flow Control Valve(FCV), while the latter is called a Pressure Reducer Valve or Pressure Regulating Valve(PRV). Shut-off valves are important components of the network since they can change the structure of the system, when for example doing maintenance or just redirecting the flow. This project deals with all three types of valves.

Valves can be controlled such that no flow passes through. In these cases the valve is closed and thereby certain parts of the system can be isolated as mentioned above. Other possibility is that the valve is fully open. In such case the pressure drop between the two endpoints is experienced because of the friction loss of the valve.

2.2 The Randers water supply network

The Randers drinking WSS is managed by Verdo A/S, which is the main supplier of drinking water and heating to the city of Randers. Verdo supplies water to approximately 46.000 customers in Randers Municipality [17]. The WSS is a complex, looped configuration with many different distribution areas. The coverage of the distribution areas is shown in *Figure 2.5* below.

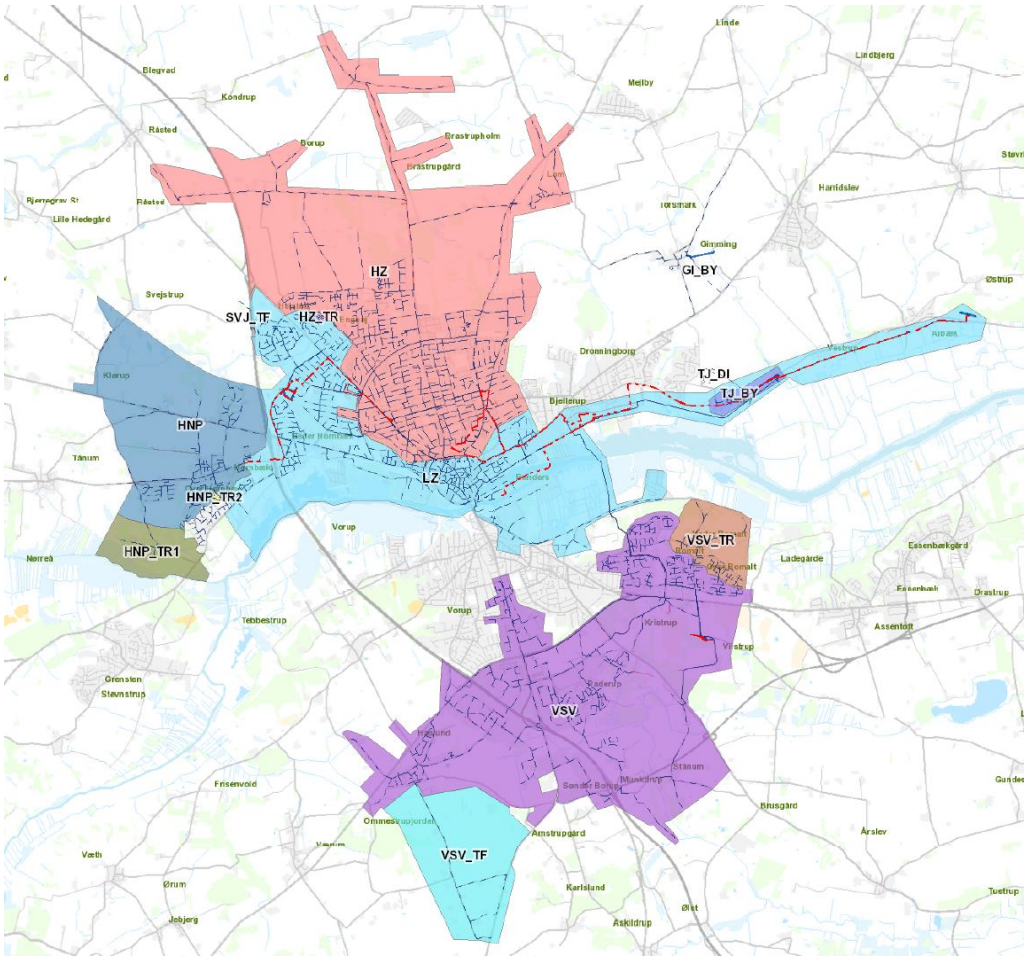


Figure 2.5: Distribution zones in the Randers WSS.

Although the network consists of many different distribution zones, the distribution in the city can be split into four main parts, regarding the different geographical properties of the different areas. Furthermore, Randers is split into two regions, Randers North and Randers South. This is due to the fact that Randers fjord divides the city into a northern and a southern part[17].

Furthermore, it is important to note that the distribution coverage illustrated in *Figure 2.5*, describes the network in its current form from the year 2013. Since 2013, additional pipelines and smaller distribution parts have been added to the network, meaning that Verdo's distribution coverage has grown during the past few years. Although the structure has slightly changed, the main water works and pumping stations are still the same, therefore the characteristics and operation of the network has not changed significantly. Due to this, the network topology is considered to be sufficiently updated for modelling and control purposes.

The distribution network which is located in the southern part of the city is called Vilstrup. This zone has its own waterwork and pumping station which allows to supply the whole

southern area by itself. The distribution area on the southern part to Randers fjord consists of this zone only, with one water work and the corresponding pumping station. This zone is shown in *Figure 2.6a* below.

(a) Vilstrup zone. (b) High Zone.

Figure 2.6: Geographical illustration of Vilstrup and High Zone distribution areas in Randers.

The only connection between Vilstrup and the northern region is through an emergency line, which is indeed used in emergency cases when the waterwork and pumping station in Vilstrup malfunctions or if there is a pollution in the WTs. In these cases waterworks from the northern parts can provide water. An emergency case is for instance when the drilled water in the water works becomes contaminated due to the intensive farming activity in this area. It is important to note that the emergency line is closed in most of the time by a valve, however this valve opens every night for approximately one hour to let the stored water out from this pipeline. It is important to let fresh water through this pipe every day, since otherwise it never gets cleaned. Besides the emergency cases, however, the drinking water distribution in Vilstrup does not rely on the waterworks and pumping stations in Randers North.

Randers North consists of three different areas, each having its own particular geographical property. The water distribution in these regions are strongly related to each other, meaning that the pump schedules rely on the water levels in different WTs or on the consumption in the different regions.

The area shown in *Figure 2.6b* above is called the High Zone(HZ) due to the high elevation level of the region. This part of the city lies approximately 55 meters above sea level, which means that high pumping effort is required to deliver the sufficient amount of pressure. Furthermore, this area has a grid structure with several loops, which is not surprising since the city center lies on this high elevation region.

The area underneath the HZ is called the Low Zone(LZ). This zone in Randers lies approximately on sea level. Therefore, the HZ and LZ have a significant elevation difference which requires special pumping solutions in this area. In order to get a visual overview of the geographical properties of the HZ and LZ, the elevation profile is shown between these two areas in *Appendix: A*, in *Figure A.1* and *Figure A.2*. The area is illustrated in *Figure 2.7*.

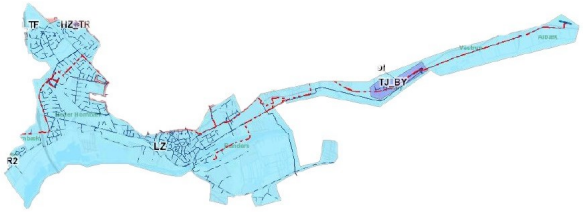


Figure 2.7: Low Zone.

The fourth main area in the Randers WSS is an area, which according to its elevation, neither belongs to the HZ, nor to the LZ. This area is called Hornbæk and shown in *Figure 2.8*.

Figure 2.8: Hornbæk zone.

The elevation in this area is around 30 meters above sea level and covers the western part of the city. This zone is connected to the rest of the distribution network through a transmission line and the properties of the structure is mainly of a tree configuration network.

2.2.1 Waterworks and pumping stations

Verdo provides drinking water by pumping water from groundwater bases and treating the water in four different water works. Due to the high quality of ground water, this water treatment is only an aeration process in all areas, except the Vilstrup zone. In the Vilstrup zone, as mentioned above, farming activity is high, therefore the water often gets contaminated.

The WSS in Randers has four waterworks and four pumping stations, located in different locations in the city. In order to draw a better picture and to introduce in a simple way the corresponding pumping and water works in the system, first they are listed and their name is given, secondly their geographical locations and properties are described. The water works and pumping stations are the following

BKV	Bunkedal Water work
ØSV	Østrup Skov Water work
VSV	Vilstrup Water work
OMV	Oust Mølle Water work

Table 2.1: Waterworks in the network.

HBP	Hobrovej Pumping Station
HSP	Hadsundvej Pumping Station
TBP	Toldbodgade Pumping Station
HNP	Hornbæk Pumping Station

Table 2.2: Pumping stations in the network.

In order to show the geographical location of the waterworks and pumping stations, an illustration of the network model is shown where each pumping station and waterwork are labelled with their names. It is also important to point out, which station belongs to the corresponding supply zones. The illustration is shown in *Figure 2.9*.

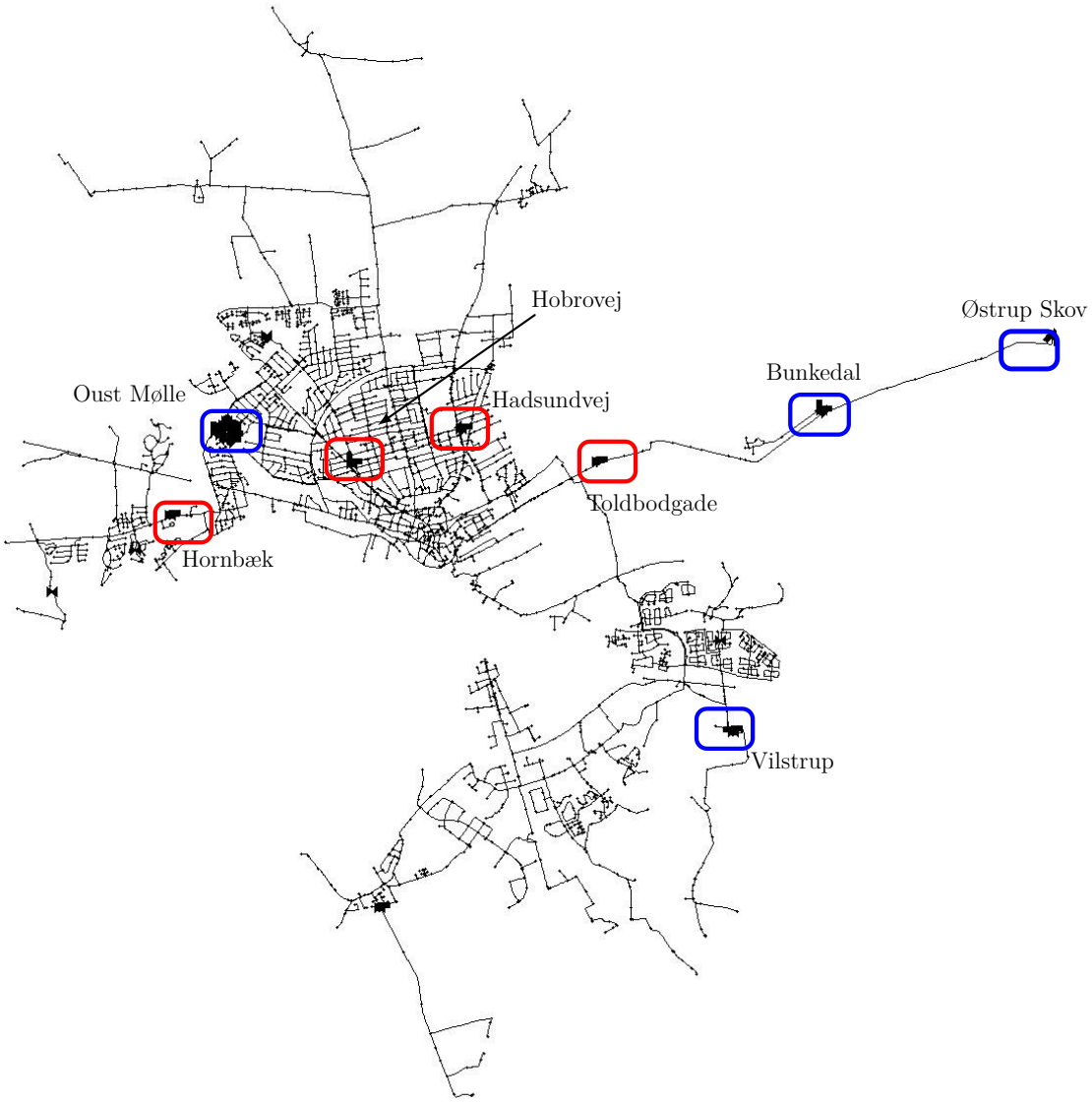


Figure 2.9: Water works(encircled in blue) and Pumping stations(encircled in red) in the network.

The two main water works in the northern region are Bunkedal and Østrup Skov. In these areas, the water quality is sufficiently good, therefore the water pumped to the surface is considered very clean and suitable for distribution. This is due to the fact that this groundwater lies under the ground such that it is protected by certain layers of the ground which makes it possible to distribute this water without any kind of cleaning process, except aeration. These protection layers has been created by Randers fjord due to glacial erosion over the past few centuries [18]. In this case, contamination by farming is not an option since these drilling stations are located very close to the fjord and therefore considered as reliable groundwater sources.

One of the drawbacks, however, is that the fresh water sources are located in the LZ, therefore the water has to be pumped from BKV, ØSV and OMV to locations with higher elevation. Since BKV, ØSV and OMV are the main sources to the HZ and LZ areas, water has to be pumped to approximately 55 meters above sea level. For this reason, the supply pumping stations Toldbodgade and Oust Mølle provide the pressurized water

to the pumping stations and corresponding elevated reservoirs in the HZ area. In the HZ, at the border of the LZ and HZ, two pumping stations are located, Hadsundvej and Hobrovej, which pump the water towards the HZ areas. The water is coming from the mentioned main supply stations in the LZ.

As can be seen in *Figure 2.9*, and as mentioned before, the network topology in the HZ is a grid. The two main pump stations, HSP and HBP provide the pressure and flow to this grid and to the LZ areas, such that they keep a balance in pressure and flow. Furthermore, there are WTs placed both at HSP and HBP. The main supply stations, OMV and TBP in the LZ operate such that the pumps are controlled according to the water level in the WTs in HBP and HSP, respectively. Since the HZ has an elevation of approximately 55 meters, the static pressure in each WT at the two pumping stations is sufficient to provide pressure in the LZ areas, without any pumping effort. Therefore HSP and HBP provide pressure to the LZ areas such that the geodesic properties are exploited. An illustration of the two pumping stations and the HZ grid is shown in *Figure 2.10*.

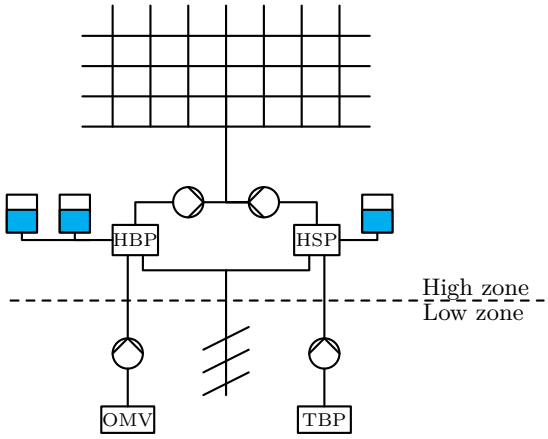


Figure 2.10: Hadsundvej and Hobrovej pumping stations with the HZ grid and LZ.

As can be seen in *Figure 2.10*, the pumping stations with the WTs are connected through the grid, and therefore the WTs are connected as well.

In order to avoid too high or too low pressure in the system, the pressure needs to be controlled in one pumping station and the flow in the other one. HBP is responsible for flow control, while HSP is responsible for the pressure control. Thereby it is avoided to provide too high flow or too low or high pressure to the end-users.

One special case of the operation which illustrates well the behaviour of the tanks located in the HZ, is when the supply pump stations, TBP and OMV are turned off. In this case, the tanks are being emptied and all the static pressure stored in the tanks is being used to supply the LZ. *Figure 2.11* shows an illustration for the pressure balance in the tanks.

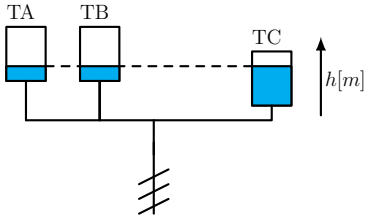


Figure 2.11: Pressure balance in the elevated reservoirs in the HZ.

It is important to note however, that it requires a long time for balancing out the levels in the WTs. Sometimes the aim is to empty the tanks in the HZ due to maintenance or

other issues and in this case, by experience, it can take 12 hours for the water level to balance out.

Furthermore, in Hornbæk, the elevation is above sea level, therefore the static pressure from the two main pumping stations are not sufficient to supply this area. Due to this, boosting is needed. The HNP pumping station in Hornbæk boosts the pressure of the water, coming from OMV.

2.2.2 Summarizing the properties of the Randers WSS

As it was described in Section *2.2: The Randers water supply network*, if normal operation is assumed, Vilstrup is an individual distribution area. Therefore, VSV is able to provide all flow demands in Randers South, without the help of the other waterworks in Randers North. Due to this consideration, the WSS in Randers South can be separated when the control of Randers North is analysed.

Taking into account the properties of the network explained in this chapter, a brief summary of the operation can be given. The control is focused on the two main pumping stations, TBP and OMV, which are both operated by controlling the flow, according to the water level in the corresponding WTs. The region called VSV in the southern part of the network is discarded from the distribution, furthermore the region called HNP is seen as a region which demands a certain amount of flow, however the control of the boosting pumps are not considered. Similarly to the water works BKV and ØSV, they treated as water sources, however in the control it it not taken into account that water comes from them. Rather, TBP main pumping station is considered as a source of the water. With these considerations, the network simplifies to the following, shown in *Figure 2.12*

Figure 2.12: The simplified network map of the Randers WSS.

The model of this system is going to be used for control purposes, which means that the model needs to be simple, but at the same time needs to have the same characteristics as of the original network. In the WSS in Randers, the control purpose is to find an optimal control scheme which is able to actuate the pumps at the two main pumping stations, such that the dynamic effect of the WTs at each station are taken into account. Challenges of control come up due to the multiple WTs in the system, actuated by multiple pumping stations. In a later chapter, the mathematical description of such systems is proposed.

3. Simulation framework in EPANET

This chapter gives an insight into the simulation model of the Randers WSS in EPANET. In order to understand better how the simulation works, the modelling of the different pumping stations, waterworks, and the consumption patterns are explained in detail. In the end, an attempt is made to split the network into different subsystems and make necessary modifications such that data is sufficiently reduced and therefore easily extractable for system identification purposes.

3.1 Model and data structure

As it was explained in Chapter *2: Description of Water Supply Systems*, the typical components of WSSs are: reservoirs, pipes, pumps and valves. For simulation purposes, it is required that the model of the real-life network consists of thousands of elements in order to accurately replicate hydraulic behaviour and the topographical layout of the system. Such models are appropriate for simulation purposes, however, online optimisation tasks are much more computationally demanding. Therefore the available data in the simulation framework will be used to create a reduced order model for control purposes.

The use of Geographic Information Systems(GIS) and Supervisory Control and Data Acquisition(SCADA) systems in the water industry resulted in an increasing amount of information about actual network topology and service that can be utilized in a model [19]. As a consequence of this, normally the simulation model of WSSs include exactly the same amount of components as in real-life.

The following data and model description strongly relies on the documentation of the Randers WSS EPANET model [20], provided by Verdo A/S, where the considerations and modelling steps with the available data from GIS are gathered. The model is mainly based on the data stored in GIS and the SCADA system, however considerations on case studies and experience have also been taken into account. It is important to note that the following results and properties of the EPANET model serve as a basis for the data processing and therefore is discussed in the report.

Initially, during the modelling, nodes were made among each different pipe elements, however this increased the calculation time significantly. Therefore the number of nodes in the network were reduced based on the fact that pipe sections with the same material and dimensions can be treated as one pipeline. In order to illustrate the complexity of the network, the number of elements in the simulation model, including the VSV region, is shown in *Table: 3.1*

Element type	**Number**
Links	4144
Nodes	4180
Tanks	6

Table 3.1: Number of WSS elements.

When we consider the simulation of the network in EPANET, it needs to be stated that certain abstractions were made in the model. These abstractions are not necessarily reflect the real-world configuration of the system, however with including them in the model, the same results can be achieved as in real-world. Therefore for instance the number of WTs in the model does not necessarily reflect the number of WTs in the real system. Water works usually have large WTs, where the aeration process and filtering is being done, however in some parts water works abstractions are used for water sources, being replaced with

nodes. Furthermore, in certain parts of the system, the modelling reflects the real-world. However, the structure of certain pumping stations are different from the real world. For the same reason, at some places in the simulation, one pump serves to simulate a whole pumping station while in the real world more pumps are placed in parallel.

3.1.1 Water consumption data

In the model, consumption data is divided into two groups: non-industrial and industrial demands. It has been chosen to keep the number of demand categories low, as it turned out that the quality of this data from GIS is not representative. This is due to the fact that single one storey buildings and two or more storey buildings have very similar consumptions. The consumption curves were defined such that they match the pumped water volumes from waterworks and pumping stations in each individual zone. Therefore, a calculation for the deviation has been made over the day such that it was possible to conclude on the uniformity of the consumption types in each zone [20]. With around 3 percent uncertainty, the consumption patterns turned out to be reliable in the model. The demand patterns over a 24 hours long period are shown in *Figure 3.1* .

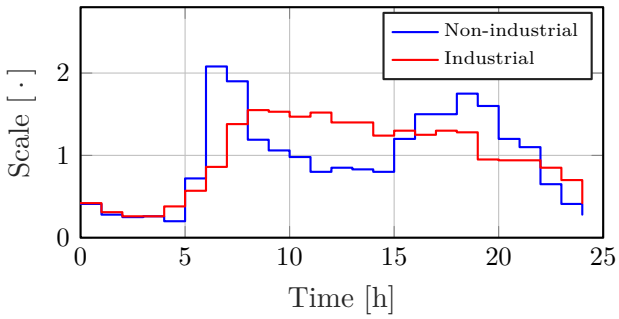


Figure 3.1: Industrial and Non-industrial demand patterns in the EPANET simulation.

In *Figure 3.1*, the y-axis shows the ratio of change compared to the base demand of the two different types of consumers. For each type of consumer, different base demands are determined. However, the variation from this base value can be described in the same manner with the presented consumption curves for users falling in the same category.

It is worth noting that the industrial consumption profile in the VSV zone slightly differs from the industries in the rest of the network. However, since the VSV zone is not considered in the further simulation model, the demand pattern is not shown here. Furthermore, there are no consumers in the corresponding supply areas who have significant effect on water consumption to influence the model calibration and simulation.

3.1.2 Control and water source abstractions

As described in Section *2.2: The Randers water supply network*, there are several different pumping stations and waterworks in the network, supplying different zones. In general, there is a possibility in EPANET to simulate the cleaning process in the waterworks, including drilling, raw water and clean water treatment. However, in the model the focus is on the correct distribution. Therefore, all waterworks are simulated with clean water reservoirs. In the EPANET simulation, pumps in the waterworks pump the water out corresponding to the real pump curves. However, there are two exceptions, where the simulation of the pumping station is different. In the OMV water work, which is responsible for filling the tanks(T1A and T1B) in HBP pumping station, the precise operation has not been taken into account. Instead, the control schedule of the pumps has been simulated as a positive demand node, which is negative according to the EPANET sign convention. The same technique was used in HBP pumping station, where the two tanks are located and which provides the distribution to the HZ. One of the reasons for this is that the controls turned out to be relatively complicated with frequency converters

and the simulation result of the pumping was not matching the real world scenario. The advantage of this is that in one of the flow controlled pumping stations, the mass balance is controlled. Furthermore, the risk that the model does not simulate correctly is reduced. It should be noted that this abstraction can be done easily in HBP and OMV, since pumps in these two pumping stations are flow controlled.

Apart from the OMV water work and HBP pumping station, all water works have been simulated with reservoirs which cannot be emptied. Water works and pumping stations, where the pumps are pressure controlled, are controlled by PRVs, as this is the typical way of controlling pressure controlled pumps in EPANET networks [21]. Such an arrangement is shown in *Figure 3.2*

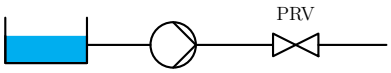


Figure 3.2: Abstraction for pressure controlled pumping stations in EPANET.

It is important to note, however, that in case of high demands, with this arrangement in the network it can happen that water works and pumping stations deliver more flow than allowed or physically possible. Therefore control rules have been incorporated in the models which prevent the pumping stations produce more flow than available in the real world.

In case of flow controlled pumping stations, such as OMV, the following abstractions are made in the EPANET simulation framework

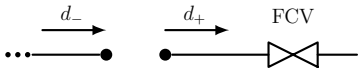


Figure 3.3: Abstraction for flow controlled pumping stations in EPANET.

In *Figure 3.3*, the node which has demand d_+ is the flow controlled pumping station. The water coming out from this node can be controlled by a FCV. In every case, the pumping stations are supplied by flow from water works. A good example for this is TBP pumping station where the water works BKV and ØSV are modelled. However, the pumping station can be modelled as two nodes, where inflow is defined on the node on the water work side and outflow is defined on the outlet side.

3.2 Model calibration and validation

The simulation data has been validated by using pressure measurements on different fire hydrants in the network for several times corresponding to different years. When the model was made, the data was not completely up to date, as these pressure measurements were carried out in years before the EPANET modelling. The major uncertainty about this data is in the arrangement of the pipe network and the pumping stations, since it has been changed over the years and old facilities have been replaced. Although the data on which the model relies is uncertain and there might be variations in pressure, the validation of the model has been carried out according to these highly uncertain pressure measurements.

3.2.1 Pipe roughness

In the model, all pipes are associated with tags, indicating dimensions, material and year information. With this information it is possible to estimate an average resistance, i.e. roughness of the pipes. During the validation process of the pipe resistances, it was chosen to consider an average roughness value, taking into account that the roughness should not be lower than a roughness of new pipes and at the same time, should not exceed a certain

upper bound. Roughnesses were upscaled at places where the pressure was too high while downscaled where the pressure was certainly too low. Correct pressure data, however is an essential information for a more detailed and precise calibration, therefore high deviations up to 5 meter heads are present in the system. *Figure 3.4* shows the pressure deviation in the system compared to the data on which the system was calibrated.

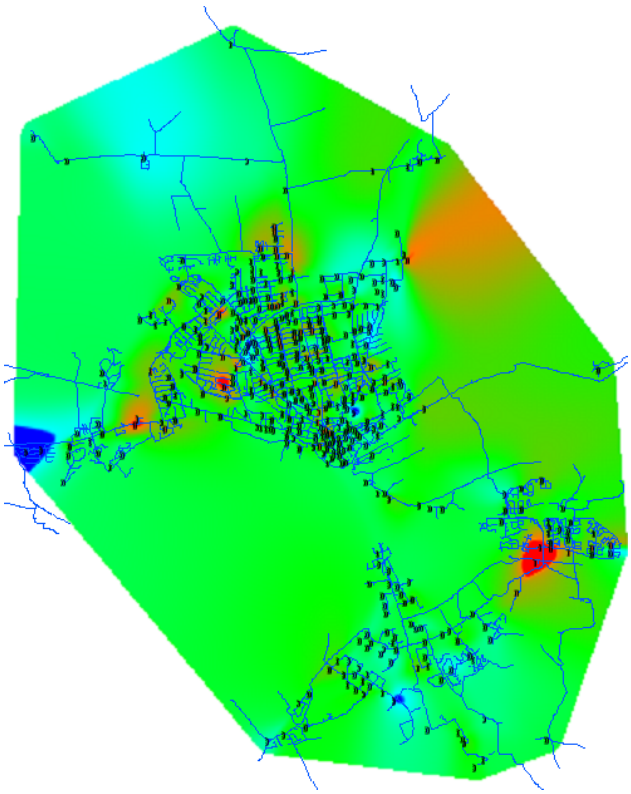
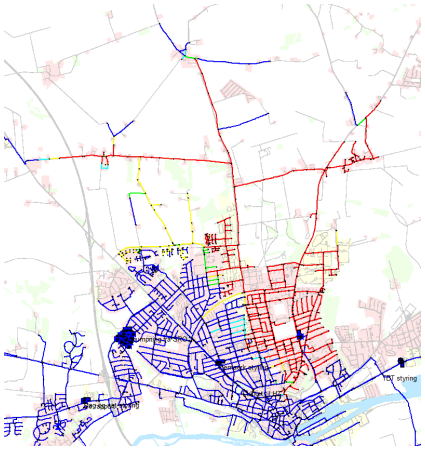


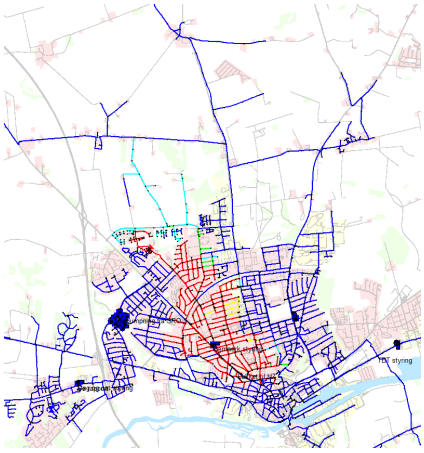
Figure 3.4: Difference calculation between observed and calculated pressures on fire hydrants[20].

3.2.2 Grid balance and supply zones

With the simulation model in EPANET there is a possibility to illustrate which pumping stations supply the different areas in the network. Simulations can be carried out for instance on supply areas in certain zones where more than one pumping station contributes to the consumption. The two zones, considered in this report are the HZ and LZ, in which the former is supplied by OMV and TBP and the latter is supplied by HBP and HSP where the tanks are placed on high elevation level. Consequently, as mentioned in Section *2.2.1: Waterworks and pumping stations*, OMV and TBP are the waterworks and pumping stations which are responsible for filling the tanks in the HZ in HBP and HSP, respectively. Using the water quality properties of EPANET, a contamination was introduced first in OMV pumping station and then TBP. By introducing a contamination into the system, it can be tracked which parts of the network gets contaminated. *Figure 3.5* shows an example for this, illustrating the distribution between the two pumping stations, OMV and TBP.



(a) HSP supply area, marked with red



(b) HBP supply area, marked with red

Figure 3.5: Supply area by HSP(on the left) and supply area by HBP(on the right)[20].

The red areas in *Figure 3.5* indicate 80-100 percent of drinking water originating from one or the other pumping station. However, the other colours(primarily yellow and green) indicate that there is a mix of water from both stations. The result in the grid is according to the control goals, as one pumping station supplies half of the region and an other the other one. This is achieved by controlling the flow in HBP and the pressure in HSP.

3.3 Modifications in the model

(in progress)

4. System Modelling

This chapter gives a mathematical description of the component modelling. Thus, the different physical and mathematical measures of hydraulic systems are introduced. The similarities to electronic networks are shown by explaining the relevant properties of graph theory. A reduced model for multi-inlet networks is first introduced, then the inclusion of tanks is considered. In the end, the EPANET-based modelling approach is introduced and the proposed models are verified on simple pipe networks by comparing them to simulation results in EPANET.

4.1 Hydraulic component modelling

In this section the mathematical relation between pressure and flow is given for each component in a WSS system, in order to show their non-linear behaviour. The purpose here is not to derive the different models, but to introduce the mathematical formalism which describes them.

Equation: (4.1) shows the dual variable which describes all two-terminal components in the network

$$\begin{pmatrix} \Delta p \\ q \end{pmatrix} = \begin{pmatrix} p_{in} - p_{out} \\ q \end{pmatrix}, \tag{4.1}$$

where

Δp	is the differential pressure across the element,	[Pa]
q	is the flow through the element,	[m^3/s]
p_{in}, p_{out}	are the absolute pressures.	[Pa]

4.1.1 Hydraulic head

As can be seen in *Equation: (4.1)*, the measure of pressure is in $[Pa]$ and the measure of volumetric flow is in $[\frac{m^3}{s}]$. In the further report, the units used for calculation are in SI, however sometimes results are shown in non-SI units. Non-SI units are considered due to the fact that EPANET uses meter head and liters instead of pascals and cubic meters for pressure and flow simulations, respectively. The unit conversion between liters and cubic meters is a constant, however meter head expresses pressure in terms of length. The relation between pressure and pressure head is explained in *Equation: (4.2)* below.

$$h_p = \frac{p}{\rho g}. \tag{4.2}$$

where

h_p	is the pressure head,	[m]
p	is the absolute pressure in pascals,	[kg/ms^2]
g	is the gravitational constant,	[m/s^2]
ρ	is the density of the fluid.	[kg/m^3]

As can be seen in *Equation: (4.2)*, if the density of the liquid is a known parameter, the conversion can be made easily between pressure and pressure head. In this project, water is considered and its density is assumed to be constant.

In general, the hydraulic head, or total head, is a measure of the potential of fluid at a specific measurement point. It relates the energy of an incompressible fluid to the height of an equivalent static column of that fluid. The different forms of energies concerning

fluids can be measured in distance, and therefore that is the reason that these terms are sometimes referred to as heads. The total hydraulic head of a fluid is composed of the pressure head and elevation head. [1]

The total head is given

$$h_t = h_p + z, \tag{4.3}$$

where

h_t	is the total head,	[m]
z	is the elevation(head).	[m]

Therefore, pressure head is a measurement of length, which is dependant on the density of the fluid but can be converted to the units of pressure. Using meters for describing pressure in the system is convenient for the reason, that pressure can be treated the same way as the elevation. In calculations, this property is exploited.

4.1.2 Pipe model

Pipes in the network are governed by the dynamic equation

$$\Delta p_i = J_i \dot{q}_i + f_i(q_i) - h_i, \tag{4.4}$$

where

J_i	is the mass inertia of the water in the pipes,
$f_i(q_i)$	is the pressure drop due to friction,
h_i	is the pressure drop due to geodesic level difference across the two terminals of pipe elements.

The dynamics of pipes due to mass inertia are discarded in this project, as it is shown in other works that the small time constant of these mass inertia dynamics are not dominant in the system, especially if there are elevated reservoirs included [9, 15]. Therefore the pressure drop across pipes can be written as

$$\Delta p_i = f_i(q_i) - h_i, \tag{4.5}$$

Additionally, the dynamics due to inertia of the liquid is not the only possible dynamics for pipes. The phenomenon, called water hammer occurs when a fluid, or gas in motion is forced to stop or change direction suddenly. In this case, a pressure wave runs through the pipe, causing vibration and possible damage in the network [23]. However, *Equation: (4.4)* models the pressure drops or equivalently, headloss, due to the elevation of the pipes and friction of the fluid. Therefore such rapid flow change is not assumed in the network.

The pressure drop due to friction across the i^{th} edge is a diagonal map where $f : \mathbb{R}^m \to \mathbb{R}^m$ is strictly increasing.[2] As it is shown in *Equation: (4.6)*, f_i describes a flow dependant pressure drop due to the hydraulic resistance such that

$$f_i(q_i) = \gamma_i |q_i| q_i, \tag{4.6}$$

where

$\gamma_i > 0$	is the resistance coefficient, the parameter of pipes.	$[\cdot]$

Equation: (4.6) is motivated by turbulent flow in the pipes, which is typical in water supply

[1] There is a third term, called the kinetic head which is discarded, since the velocity of the fluid is assumed to be constant along the cross sectional area in the whole length of pipes [22].

[2] A map $f : \mathbb{R}^m \to \mathbb{R}^m$ is strictly increasing if $\langle x - y, f(x) - f(y) \rangle > 0$ for every $x, y \in \mathbb{R}^m$ such that $x \neq y$ [24].

applications. The resistance coefficient is calculated according to the Darcy-Weisbach formula, which provides the theoretically most precise result and is the most commonly used in Europe[8, 21]. γ is given as shown in *Equation: (4.7)* below[3].

$$\gamma(q) = \frac{c_D f_D(\epsilon, Re(q), D) l}{D^5}, \tag{4.7}$$

where

f_D	is the Darcy friction factor,	$[\cdot]$
ϵ	is the roughness of the pipe,	[m]
D	is the diameter of the pipe,	[m]
Re	is the Reynolds-number,	$[\cdot]$
c_D	is a coefficient in the D-W equation,	$[s^2/m]$
l	is the length of the pipe.	[m]

In *Equation: (4.7)*, the Darcy friction factor, f_D, is dependant on the Reynolds number, which is defined by the volumetric flow in the pipes. However, at high flows Re becomes nearly constant and therefore normally this flow dependency is disregarded. Thus, f_D is considered to be constant in the further project.

The derivation of *Equation: (4.7)* is explained in more detail in appref [22]. Furthermore, in the following sections it is assumed that each f_i has a structure shown in *Equation: (4.6)*.

It is important to note here that $f_i(\cdot)$ is a homogeneous map which means that if the argument is multiplied by a positive scalar, then its value is multiplied by some power of this scalar [4]. For $f_i(q_i)$, it can be shown that

$$\gamma_i|(\alpha q_i)|(\alpha q_i) = f_i(\alpha q_i) = \alpha^2 f_i(q_i). \tag{4.8}$$

More precisely, with the given structure of $f(\cdot)$ the scaling would be $|\alpha|\alpha$, however $\alpha \geq 0$ is already assumed in *Equation: (4.8)*. This property is noted here and used later in the system description, in Section *4.4: Multi-inlet model*, where the scaling is indeed such that $\alpha \in \mathbb{R}_+$.

4.1.3 Valve model

Valves in the network are governed by the following algebraic expression

$$\Delta p_i = \mu_i(q_i, k_v) = \frac{1}{k_v(OD)^2}|q_i|q_i, \tag{4.9}$$

where

k_v	is the valve conductivity function, taking the Opening Degree(OD) of the valve in its argument [15].

$\mu_i(q_i, k_v)$ is a continuously differentiable and proper function which for $q_i = 0$ is zero and monotonically increasing.

4.1.4 Pump model

Centrifugal pumps are governed by the following expression [16]

$$\Delta p = -a_{h2}{q_i}^2 + a_{h1}\omega_r q_i + a_{h0}{\omega_r}^2 \tag{4.10}$$

[3] EPANET uses the D-W formula for calculating the resistance terms.
[4] $g(\alpha v) = \alpha^k g(v)$

where

Δp	is the differential pressure produced by the pump,	[Pa]
a_{h2},a_{h1},a_{h0}	are constants describing the pump,	[·]
ω_r	is the impeller rotational speed.	[rad/s]

The model described in *Equation: (4.10)*, works only for positive flows, therefore it is assumed that liquid cannot flow back to the pump.

4.1.5 Elevated reservoir model

In elevated reservoirs, the rate of change in the volume of the fluid inside the tank is equal to the volumetric flow at which water enters or leaves the tank. Since the pressure on the bottom is due to the cross sectional area of the tank and the amount of water in it, proportional relation can be set between the pressure and the flow in and out of the tank. The dynamics of such system can be described by a first order differential equation of the form

$$\dot{p}_i = -\tau_i\Big(\frac{p_i}{h_{l,i}}\Big)q_i \tag{4.11}$$

where

p_i	is the pressure at the node connected to the tank,	[Pa]
$h_{l,i}$	is the water level in the tank,	[m]
τ_i	is the parameter in terms of the cross sectional area and the pressure/water level ratio in the tank,	[Pa/m^3]
q_i	is the flow into the tank if $q_i < 0$ and flow out of the tank if $q_i > 0$.	[m^3/s]

As can be seen in *Equation: (4.11)*, in general, the parameter of the tank depends on the pressure and water level ratio, if the cross sectional are is not constant along the height of the tank. However, it is assumed that tanks have the same cross sectional areas in the entire height. Then *Equation: (4.11)* simplifies to

$$\dot{p}_i = -\tau_i q_i. \tag{4.12}$$

In this case the parameter of the tank is given by

$$\tau_i = \rho g \frac{1}{A_{wt,i}}, \tag{4.13}$$

where

$A_{wt,i}$	is the cross sectional area of the i^{th} tank.	[m^2]

4.2 Graph-based network modelling

Tools from Graph Theory(GT) can be utilized to model a WSS as a directed graph, similarly as in Circuit Theory(CT). In the graph the set of vertices represent sources, consumption nodes and reservoirs, while the set of edges, represent pipes, pumps and valves.

In order to track the pressure and flow in the desired part of the network, the equation system of the network has to be solved for the corresponding edges and vertices. The whole network can be described by writing up the equations for all edges in the network, based on the mathematical modelling of the different components in the system, as shown in Section *4.1: Hydraulic component modelling*. However, in case of complex systems such as water networks for large cities, these systems of equations are difficult to handle individually and typically cannot be solved explicitly if the system consists of loops. Therefore the properties of GT are not only useful for setting up relations between flow and pressure, but to make the handling of algebraic constraints easier by exploiting the properties of

linear algebra. Thereby making it convenient for implementing it in computer algorithms for iterative solving methods.

WSSs can be described by a directed and connected graph, such that [25]:

$$\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}, \tag{4.14}$$

where

$\mathcal{G}$	is a directed and connected graph,
$\mathcal{V}$	is the set of vertices, where $\mathcal{V} = \{v_1, ..., v_n\}$,
$\mathcal{E}$	is the set of edges, where $\mathcal{E} = \{e_1, ..., e_m\}$.

4.2.1 Incidence matrix

The incidence matrix, H, of a connected graph, $\mathcal{G}$, is a matrix where the number of rows and columns correspond to the number of vertices and edges, respectively. Therefore $H \in \mathbb{R}^{n \times m}$. In case of hydraulic networks, edges are directed in order to keep track of the direction of the flow in the system.

$$H_{i,j} = \begin{cases} 1 & \text{if the } j^{th} \text{ edge is incident out of the } i^{th} \text{ vertex.} \\ -1 & \text{if the } j^{th} \text{ edge is incident into the } i^{th} \text{ vertex.} \\ 0 & \text{if the } j^{th} \text{ edge is not connected to the } i^{th} \text{ vertex.} \end{cases} \tag{4.15}$$

It is worth mentioning that the reduced incidence matrix can be obtained by removing any arbitrary row from H. Therefore H always has $(n-1)$ row rank. This statement is induced by the mass conservation in the network and explained in the following section, Section *4.3: Kirchhoff's and Ohm's law for hydraulic networks.*

4.2.2 Cycle matrix

Purely tree structure of a WSS is not common when complex distribution systems are considered. However, trees can be arbitrarily chosen from the underlying graph of the network.[5] A tree, $\mathcal{T}^*$, of a graph is a connected sub-graph where any two vertices are connected by exactly one path [26]. Therefore trees in the network can be represented as follows

$$\mathcal{T}^* = \{\mathcal{V}_{\mathcal{T}^*}, \mathcal{E}_{\mathcal{T}^*}\}. \tag{4.16}$$

A special case of connected tree sub-graphs is the spanning tree of the network. A spanning tree contains all the vertices of $\mathcal{G}$ and has no cycles, since it is a tree. A spanning tree of the network therefore can be represented as

$$\mathcal{T}^*_{span} = \{\mathcal{V}, \mathcal{E}_{\mathcal{T}^*}\}. \tag{4.17}$$

In order to obtain a spanning tree, an edge has to be removed from each cycle. The removed edges are $\mathcal{G} - \mathcal{T}^*$, and called the chords of $\mathcal{T}^*$ with respect to $\mathcal{G}$. By adding a chord to $\mathcal{T}^*$, a cycle is created which is called a fundamental cycle. A graph is conformed by as many fundamental cycles as the number of chords [26].

The set of fundamental cycles correspond to the fundamental cycle matrix, B, such that the number of rows and columns are defined by the number of chords and edges, respectively. The cycle matrix of the system is given by

$$B_{i,j} = \begin{cases} 1 & \text{if the } j^{th} \text{ edge belongs to the } i^{th} \text{ cycle and their directions agree.} \\ -1 & \text{if the } j^{th} \text{ edge belongs to the } i^{th} \text{ cycle and their directions are opposite.} \\ 0 & \text{if the } j^{th} \text{ edge does not belong to the } i^{th} \text{ cycle.} \end{cases} \tag{4.18}$$

[5] Recall that a tree with n vertices has $n-1$ edges [26].

4.3 Kirchhoff's and Ohm's law for hydraulic networks

In this project, the hydraulic system is considered to be an open network with pipes, valves, pumps and the storage tanks, where water is able to enter and leave the network at a subset of the vertices. For such system, Kirchhoff's vertex law, or equivalently Kirchhoff's current law(KCL), corresponds to conservation of mass in each vertex and described by

$$Hq = d, \tag{4.19}$$

where

$d \in \mathbb{R}^n$	is the vector of nodal demands, with $d_i > 0$ when demand flow is into vertex i and $d_i < 0$ when demand flow is out of vertex i.	$[\mathrm{m}^3/\mathrm{s}]$

Nodal demands are seen as the end-user consumption, which means that water is taken out from the network. The mass conservation corresponds to the fact that what is consumed from the system must also be produced. Due to mass conservation, there can be only $(n-1)$ independent nodal demands in the network

$$d_n = -\sum_{i=1}^{n-1} d_i. \tag{4.20}$$

As a matter of fact, *Equation: (4.20)* is not an additional constraint since it follows from *Equation: (4.19)*. This can be sown by using the knowledge that 1_n is the left kernel[6] of H.

In the further report, a distinction is made between inlet and non-inlet nodes. It is assumed that the demand at non-inlet nodes fulfil the following constraint

$$d_i \leq 0. \tag{4.21}$$

It is worth noting however, that in closed hydraulic networks the vertex law becomes

$$Hq = 0. \tag{4.22}$$

Ohm's law for hydraulic networks is expressed with the incidence matrix, when H^T is applied to the vector of absolute pressures, p. Important to point out that the description below in *Equation: (4.23)* is valid if edges of the underlying graph are considered as only pipe elements.

$$\Delta p = H^T p = f(q) - H^T h. \tag{4.23}$$

In *Equation: (4.23)*, the differential pressure is described across each edge in the network, taking into account the pressure loss due to friction, $f(q)$, and the pressure drop due to geodesic level differences, where $h \in \mathbb{R}^n$ is the vector of geodesic levels at each vertex expressed in units of potential, i.e. pressure. It is noted that the pressure loss, $f(q)$, the absolute pressure, p and the geodesic level h can be both considered in units of meter. As mentioned in the previous section, handling Ohm's law for hydraulic systems in meter is convenient, since the elevation is already in meters.

4.4 Multi-inlet model

The system is a water network supplied from more than one pumping stations and the distribution is to several end-users. In the underlying graph therefore the nodes are pipe

[6] The kernel of matrix $A \in \mathbb{R}^{m \times n)}$ is $\{x \in \mathbb{R}^n | Ax = 0\}$.

connections, with possible water demand from the end-users, and the edges are only pipes.

The aim of the modelling here is to obtain a reduced order network model which is able to capture the dependence of the measured output pressures, on the flows and pressures at the inlets. The inclusion of storage tanks is the next step of the model development, therefore it is described in a following section, in Section *4.5: Inclusion of elevated reservoirs*.

It is assumed that the inlet pressures and demands are measured. Furthermore, pressure measurement is available in certain parts of the remaining network, at the end-users. Considering generality, the model is described for c inlets, however it should be noted that regarding the Randers WSS, two inlet vertices are taken into account.

In order to put the system into a form which can handle the measured pressure dependencies on the control inputs, the underlying graph of the network is first partitioned. The n vertices of the graph are separated into two sets

$$\mathcal{V} = \{\bar{\mathcal{V}}, \hat{\mathcal{V}}\}, \tag{4.24}$$

where

$\hat{\mathcal{V}} = \{\hat{v}_1, ..., \hat{v}_c\}$ represents the vertices corresponding to the inlet points,

and

$\bar{\mathcal{V}} = \{\bar{v}_1, ..., \bar{v}_{n-c}\}$ represents the remaining vertices in the graph.

The partitioning for the m edges of the graph is being chosen such that

$$\mathcal{E} = \{\mathcal{E}_\mathcal{T}, \mathcal{E}_\mathcal{C}\}, \tag{4.25}$$

where

$\mathcal{E}_\mathcal{T} = \{e_{\mathcal{T},1}, ..., e_{\mathcal{T},n-c}\}$

and

$\mathcal{E}_\mathcal{C} = \{e_{\mathcal{C},1}, ..., e_{\mathcal{C},m-n+c}\}$.

The subsets regarding edges and the partitioning is chosen such that the sub-matrix, which maps edges in $\mathcal{E}_\mathcal{T}$ to vertices in $\bar{\mathcal{V}}$, is invertible. It is worth mentioning that such partitioning is always possible for connected graphs.

Therefore, the incidence matrix can be split into four sub-matrices, as shown in *Equation: (4.26)* below

$$H = \begin{pmatrix} \bar{H}_\mathcal{T} & \bar{H}_\mathcal{C} \\ \hat{H}_\mathcal{T} & \hat{H}_\mathcal{C} \end{pmatrix}, \tag{4.26}$$

where

$\bar{H}_\mathcal{T} \in \mathbb{R}^{(n-c)\times(n-c)}$ is the sub-matrix, mapping edges in $\mathcal{E}_\mathcal{T}$ to vertices in $\bar{\mathcal{V}}$,
$\bar{H}_\mathcal{C} \in \mathbb{R}^{(n-c)\times(m-n+c)}$ is the sub-matrix, mapping edges in $\mathcal{E}_\mathcal{C}$ to vertices in $\bar{\mathcal{V}}$,
$\hat{H}_\mathcal{T} \in \mathbb{R}^{c\times(n-c)}$ is the sub-matrix, mapping edges in $\mathcal{E}_\mathcal{T}$ to vertices in $\hat{\mathcal{V}}$,
$\hat{H}_\mathcal{C} \in \mathbb{R}^{c\times(m-n+c)}$ is the sub-matrix, mapping edges in $\mathcal{E}_\mathcal{C}$ to vertices in $\hat{\mathcal{V}}$.

It is worth noting that the only requirement for the edge partitioning is $\bar{H}_\mathcal{T}$ being invertible[7]. Furthermore, the set $\mathcal{T} = \{\mathcal{V}, \mathcal{E}_\mathcal{T}\}$ is not necessarily a tree of the underlying graph, it can be any form of graph that fulfils the invertibility requirements. The set here, $\mathcal{T}$, is not connected due to the requirement of $\mathcal{E}_\mathcal{T} \geq (n-1)$. For the multi-inlet case, $c > 1$, therefore $\mathcal{E}_\mathcal{T} = (n-c)$. However, one special case is given when $c = 1$, meaning that the network has only one inlet. In this case, $\mathcal{T}$ is indeed a spanning tree.

With the chosen partitioning, Kirchhoff's vertex law in *Equation: (4.19)* can be rewritten

[7] $\exists\{\mathcal{V}, \mathcal{E}\} : \bar{H}_\mathcal{T}^{-1} \because rank(H) = (n-1)$ [26]

as

$$\bar{d} = \bar{H}_{\mathcal{T}} q_{\mathcal{T}} + \bar{H}_{\mathcal{C}} q_{\mathcal{C}}, \tag{4.27a}$$

$$\hat{d} = \hat{H}_{\mathcal{T}} q_{\mathcal{T}} + \hat{H}_{\mathcal{C}} q_{\mathcal{C}}, \tag{4.27b}$$

and Ohm's law in *Equation: (4.23)*, separating the pressure drops due to hydraulic resistance

$$f_{\mathcal{T}}(q_{\mathcal{T}}) = \bar{H}_{\mathcal{T}}^T(\bar{p} + \bar{h}) + \hat{H}_{\mathcal{T}}^T(\hat{p} + \hat{h}), \tag{4.28a}$$

$$f_{\mathcal{C}}(q_{\mathcal{C}}) = \bar{H}_{\mathcal{C}}^T(\bar{p} + \bar{h}) + \hat{H}_{\mathcal{C}}^T(\hat{p} + \hat{h}). \tag{4.28b}$$

Writing up *Equation: (4.28a)* and *Equation: (4.28b)* in matrix form, the following yields

$$\begin{pmatrix} f_{\mathcal{T}}(q_{\mathcal{T}}) \\ f_{\mathcal{C}}(q_{\mathcal{C}}) \end{pmatrix} = \underbrace{\begin{pmatrix} \bar{H}_{\mathcal{T}}^T & \hat{H}_{\mathcal{T}}^T \\ \bar{H}_{\mathcal{C}}^T & \hat{H}_{\mathcal{C}}^T \end{pmatrix}}_{\begin{pmatrix} \bar{H}^T & \hat{H}^T \end{pmatrix}} \begin{pmatrix} \bar{p} + \bar{h} \\ \hat{p} + \hat{h} \end{pmatrix} \tag{4.29}$$

As it is shown in *Equation: (4.29)*, the transposed incidence matrices can be written up as the two sub-matrices partitioned according to inlet and non-inlet nodes.

Now, defining a matrix Γ, in which the partitioning of the edges are the same as for the incidence matrix, H. Γ is defined as follows

$$\Gamma = \begin{pmatrix} -\bar{H}_{\mathcal{C}}^T \bar{H}_{\mathcal{T}}^{-T} & I \end{pmatrix} \tag{4.30}$$

It should be noted that the expressions in matrix Γ are of the same structure as the structure of a partitioned cycle matrix [26]. However, as mentioned, the set $\mathcal{T}$ does not define a spanning tree when $c > 1$, therefore matrix Γ is not a cycle matrix corresponding to any spanning trees. Multiplying H with Γ from the left-hand side

$$\Gamma H^T = \begin{pmatrix} -\bar{H}_{\mathcal{C}}^T \bar{H}_{\mathcal{T}}^{-T} & I \end{pmatrix} \begin{pmatrix} \bar{H}_{\mathcal{T}}^T & \hat{H}_{\mathcal{T}}^T \\ \bar{H}_{\mathcal{C}}^T & \hat{H}_{\mathcal{C}}^T \end{pmatrix} = \begin{pmatrix} 0 & -\bar{H}_{\mathcal{C}}^T \bar{H}_{\mathcal{T}}^{-T} \hat{H}_{\mathcal{T}}^T + \hat{H}_{\mathcal{C}}^T \end{pmatrix}. \tag{4.31}$$

Multiplying with Γ from the left in *Equation: (4.29)*

$$\begin{pmatrix} -\bar{H}_{\mathcal{C}}^T \bar{H}_{\mathcal{T}}^{-T} & I \end{pmatrix} \begin{pmatrix} f_{\mathcal{T}}(q_{\mathcal{T}}) \\ f_{\mathcal{C}}(q_{\mathcal{C}}) \end{pmatrix} = \begin{pmatrix} -\bar{H}_{\mathcal{C}}^T \bar{H}_{\mathcal{T}}^{-T} & I \end{pmatrix} \begin{pmatrix} \bar{H}_{\mathcal{T}}^T & \hat{H}_{\mathcal{T}}^T \\ \bar{H}_{\mathcal{C}}^T & \hat{H}_{\mathcal{C}}^T \end{pmatrix} \begin{pmatrix} \bar{p} + \bar{h} \\ \hat{p} + \hat{h} \end{pmatrix} \tag{4.32}$$

induces the following expression

$$f_{\mathcal{C}}(q_{\mathcal{C}}) - \bar{H}_{\mathcal{C}}^T \bar{H}_{\mathcal{T}}^{-T} f_{\mathcal{T}}(q_{\mathcal{T}}) = (\hat{H}_{\mathcal{C}}^T - \bar{H}_{\mathcal{C}}^T \bar{H}_{\mathcal{T}}^{-T} \hat{H}_{\mathcal{T}}^T)(\hat{p} + \hat{h}). \tag{4.33}$$

From *Equation: (4.27a)*, the vector $q_{\mathcal{T}}$, of flows in edges $\mathcal{E}_{\mathcal{T}}$ can be expressed

$$q_{\mathcal{T}} = -\bar{H}_{\mathcal{T}}^{-1} \bar{H}_{\mathcal{C}} q_{\mathcal{C}} + \bar{H}_{\mathcal{T}}^{-1} \bar{d}. \tag{4.34}$$

Therefore using *Equation: (4.34)*, *Equation: (4.33)* can be rewritten

$$f_{\mathcal{C}}(q_{\mathcal{C}}) - \bar{H}_{\mathcal{C}}^T \bar{H}_{\mathcal{T}}^{-T} f_{\mathcal{T}}(-\bar{H}_{\mathcal{T}}^{-1} \bar{H}_{\mathcal{C}} q_{\mathcal{C}} + \bar{H}_{\mathcal{T}}^{-1} \bar{d}) = (\hat{H}_{\mathcal{C}}^T - \bar{H}_{\mathcal{C}}^T \bar{H}_{\mathcal{T}}^{-T} \hat{H}_{\mathcal{T}}^T)(\hat{p} + \hat{h}). \tag{4.35}$$

Now expressing the vertex demands at non-inlet vertices, $\bar{d}$, such that

$$\bar{d} = -v\sigma \tag{4.36}$$

where

$\bar{d} \in \mathbb{R}^{n-c}$	is the vector of nodal demands in non-inlet vertices,
$\sigma \in \mathbb{R}_+$	is the total demand in the network, representing the total consumption of the end-users,
$v \in \mathbb{R}_{n-c}$	represents the distribution vector of nodal demands among the non-inlet vertices with the property $\sum_i v_i = 1$ and $v_i \in (0;1)$.

Furthermore, introducing a vector, $a_\mathcal{C}$, such that

$$q_\mathcal{C} = a_\mathcal{C}\sigma. \tag{4.37}$$

It is worth mentioning that such an $a_\mathcal{C}$ can always be defined in this manner as long as $\sigma \neq 0$.

Having $\bar{d}$ and $q_\mathcal{C}$ introduced as the linear function of the total demand, σ, *Equation: (4.35)* can be expressed such that

$$\begin{aligned} &f_\mathcal{C}(q_\mathcal{C}) - \bar{H}_\mathcal{C}^T \bar{H}_\mathcal{T}^{-T} f_\mathcal{T}(-\bar{H}_\mathcal{T}^{-1}\bar{H}_\mathcal{C} q_\mathcal{C} + \bar{H}_\mathcal{T}^{-1}\bar{d}) = \\ &f_\mathcal{C}(a_\mathcal{C}\sigma) - \bar{H}_\mathcal{C}^T \bar{H}_\mathcal{T}^{-T} f_\mathcal{T}(-\bar{H}_\mathcal{T}^{-1}\bar{H}_\mathcal{C} a_\mathcal{C}\sigma - \bar{H}_\mathcal{T}^{-1} v\sigma) = \\ &f_\mathcal{C}(a_\mathcal{C})\sigma^2 - \bar{H}_\mathcal{C}^T \bar{H}_\mathcal{T}^{-T} f_\mathcal{T}(-\bar{H}_\mathcal{T}^{-1}\bar{H}_\mathcal{C} a_\mathcal{C} - \bar{H}_\mathcal{T}^{-1} v)\sigma^2. \end{aligned} \tag{4.38}$$

where the latter equality is due to the homogeneity property of the pressure drops due to frictions, previously explained in Section *4.1.2: Pipe model*.

Defining a function $F_v : \mathbb{R}^{m-n+c} \to \mathbb{R}^{m-n+c}$, parametrized with v such that it takes $a_\mathcal{C}$ as input, the following expression can be formed

$$F_v(a_\mathcal{C}) = f_\mathcal{C}(a_\mathcal{C}) - \bar{H}_\mathcal{C}^T \bar{H}_\mathcal{T}^{-T} f_\mathcal{T}(-\bar{H}_\mathcal{T}^{-1}\bar{H}_\mathcal{C} a_\mathcal{C} - \bar{H}_\mathcal{T}^{-1} v) \tag{4.39}$$

Furthermore, $F_v(\cdot)$ equals to the following, according to *Equation: (4.35)*

$$F_v(a_\mathcal{C}) = \frac{1}{\sigma^2}(\hat{H}_\mathcal{C}^T - \bar{H}_\mathcal{C}^T \bar{H}_\mathcal{T}^{-T} \hat{H}_\mathcal{T}^T)(\hat{p} + \hat{h}). \tag{4.40}$$

An algebraic expression for a_c can be found iff $\exists F_v^{-1}(\cdot)$. It can be shown, however that $\exists F_v^{-1}(\cdot)$ by showing that $F_v(\cdot)$ is a homeomorphism[8], which is done in [24].

Equation: (4.40) being a homeomorphism in $q_\mathcal{C}$ means that the solution $q_\mathcal{C}$ is unique for every given values of $\hat{p}$, $\hat{h}$ and σ.

As a result of using the inverse mapping of F_v, a unique expression can be obtained for $a_\mathcal{C}$

$$a_\mathcal{C} = F_v^{-1} A_1(\hat{p} + \hat{h}), \tag{4.41}$$

where

$A_1 = \hat{H}_\mathcal{C}^T - \bar{H}_\mathcal{C}^T \bar{H}_\mathcal{T}^{-T} \hat{H}_\mathcal{T}^T \in \mathbb{R}^{(m-n+c \times c)}$.

A_1 has a non-trivial kernel, and for every unique value of $\frac{1}{\sigma^2} A_1(\hat{p} + \hat{h})$, there is a unique solution for $a_\mathcal{C}$.

The main objective of writing up $a_\mathcal{C}$ is to show that it can be expressed in terms of $v, \sigma(t), \hat{p}(t)$ and $\hat{h}$, where $\hat{h}$ and $\sigma(t)$ are assumed to be known signals and parameters, $v(t)$ is an unknown parameter and $\hat{p}(t)$ is the control signal. The difficulty about the constraint

[8] Two functions are homeomorphic if they can be formed into each other by continuous, invertible mapping [27]. However, here invertibility is a sufficient condition.

on $a_{\mathcal{C}}$ in *Equation: (4.41)* is that its structure is unknown. Equivalently, the solution for $a_{\mathcal{C}}$ cannot be expressed analytically but there exists a unique numerical solution for it.

Furthermore, assuming that $(\hat{p} + \hat{h}) \neq 0 \in \ ker(A_1)$, then $a_{\mathcal{C}}$, *Equation: (4.40)* can be simplified such that

$$a_{\mathcal{C}} = F_v^{-1}(0). \tag{4.42}$$

Equation: (4.42) shows, that in the special case when the input vertices are chosen such that the product $A_1(\hat{p} + \hat{h}) = 0$, then $a_{\mathcal{C}}$ becomes only dependent on the parameter v.

Now, using the equations for Ohm's law in *Equation: (4.28a)* and the vector $q_{\mathcal{T}}$, of flows in edges $\mathcal{E}_{\mathcal{T}}$ in *Equation: (4.34)*, the vector $\bar{p}$ of pressures at non-inlet vertices is expressed

$$\begin{aligned} \bar{p} =& \bar{H}_{\mathcal{T}}^{-T} f_{\mathcal{T}}(-\bar{H}_{\mathcal{T}}^{-1}\bar{H}_{\mathcal{C}} q_{\mathcal{C}} + \bar{H}_{\mathcal{T}}^{-1}\bar{d}) - \bar{H}_{\mathcal{T}}^{-T}\hat{H}_{\mathcal{T}}^{T}(\hat{p} + \hat{h}) - \bar{h} \\ =& \bar{H}_{\mathcal{T}}^{-T} f_{\mathcal{T}}(-\bar{H}_{\mathcal{T}}^{-1}\bar{H}_{\mathcal{C}} a_c + \bar{H}_{\mathcal{T}}^{-1} v)\sigma^2 - \bar{H}_{\mathcal{T}}^{-T}\hat{H}_{\mathcal{T}}^{T}(\hat{p} + \hat{h}) - \bar{h} \end{aligned} \tag{4.43}$$

As shown in *Equation: (4.43)*, the output vector which consists of the pressures in the non-inlet vertices can be written up in terms of $\sigma(t)$, $\hat{p}(t)$ time-varying signals, in terms of $\hat{h}$ and $\bar{h}$ constants and in terms of the parameter $a_{\mathcal{C}}$ and v. In the non-general case, as shown in *Equation: (4.42)*, a_c is a parameter which is governed by the behaviour of the total demand distribution among the non-inlet vertices. In case vector v is constant, thereby time-invariant, which means that the distribution of nodal demands are the same in all vertices in the network at all time, the output pressure in the i^{th} non-inlet vertices can be written as follows:

$$\bar{p}_i(t) = \alpha_i \sigma^2(t) + \sum_j \beta_{ij}\hat{p}_j(t) + \gamma_i \tag{4.44}$$

where

$$\begin{aligned} \alpha_i &= (\bar{H}_{\mathcal{T}}^{-T})_i f_{\mathcal{T}}(-\bar{H}_{\mathcal{T}}^{-1}\bar{H}_{\mathcal{C}} a_{\mathcal{C}} + \bar{H}_{\mathcal{T}}^{-1} v) \\ \beta_{ij} &= -(\bar{H}_{\mathcal{T}}^{-T}\hat{H}_{\mathcal{T}}^{T})_{ij} \\ \gamma_i &= -(\bar{H}_{\mathcal{T}}^{-T}\hat{H}_{\mathcal{T}}^{T})_i \hat{h} - \bar{h}_i \end{aligned}$$

However in WSSs, the above-mentioned consideration for v is unrealistic, meaning that the distribution of nodal demands in the non-inlet vertices should depend on time, as the end-user water consumption is not the same in every hour. This consumption behaviour of the end-users, however, is assumed to be periodic, which is a fair assumption, taking into account that the daily consumption shows approximately the same trends every day.

Therefore the demand in non-inlet vertices, described in *Equation: (4.36)* can be rewritten such that

$$\bar{d}(t) = -v(t)\sigma(t) \tag{4.45}$$

where
$v(t + T) = v(t)$,
$\sigma(t + T) = \sigma(t)$,
and T is the length of the period.

If the non-inlet demands are time-varying, but periodic behaviour is assumed and on top of this, the input vertices are arranged such that *Equation: (4.42)* is fulfilled, *Equation: (4.46)* can be rewritten as follows

$$\bar{p}_i(t) = \alpha_i(t) \sigma^2(t) + \sum_j \beta_{ij}\hat{p}_j(t) + \gamma_i, \tag{4.46}$$

where α_i is also a time-varying parameter of the model.

4.4.1 Simulation example

In order to illustrate and to show how the implementation of the model described in Section *4.4: Multi-inlet model* works, it is first tested on a simple, two-source, two-loop pipe network. This simple pipe system serves as an example to point out the differences and possible simplifications which were done regarding the modelling in Matlab and in EPANET. In this case, simulation results are steady-state values of pressures and flows, and the models are excited by some arbitrary pressure inputs.

The underlying graph of the example network is shown in *Figure 4.1*

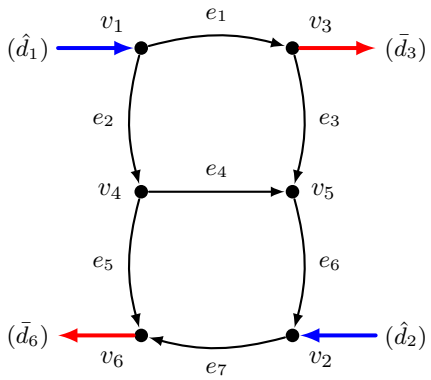


Figure 4.1: Graph of a simple multi-inlet network.

In *Figure 4.1*, arrows illustrate the in-and outflows such that input flows are present in v_1 and v_2, and user-consumption is defined only in v_3 and v_6. The diameters are the same for all pipes, however regarding the length, there are two types of pipes in this network. The corresponding parameters of the pipes, pumps, the elevation profile and the hourly demand variation is described in detail and can be found in (appref).

The partitioning of edges in the network is chosen such that

$$\mathcal{E}_\mathcal{T} = \{e_1, e_5, e_3, e_4\} \equiv \{e_{\mathcal{T},1}, e_{\mathcal{T},2}, e_{\mathcal{T},3}, e_{\mathcal{T},4}\}, \tag{4.47}$$

and

$$\mathcal{E}_\mathcal{C} = \{e_2, e_6, e_7\} \equiv \{e_{\mathcal{C},1}, e_{\mathcal{C},2}, e_{\mathcal{C},3}\}, \tag{4.48}$$

The corresponding vectors describing the pressures and flows in vertices and edges, respectively, furthermore the elevation and distribution profiles are given such that

$$p(t) = \begin{pmatrix} \bar{p}_3(t) \\ \bar{p}_4(t) \\ \bar{p}_5(t) \\ \bar{p}_6(t) \\ \hat{p}_1(t) \\ \hat{p}_2(t) \end{pmatrix}, \quad d(t) = \begin{pmatrix} \bar{d}_3(t) \\ 0 \\ 0 \\ \bar{d}_6(t) \\ \hat{d}_1(t) \\ \hat{d}_2(t) \end{pmatrix}, \quad h = \begin{pmatrix} \bar{h}_3 \\ \bar{h}_4 \\ \bar{h}_5 \\ \bar{h}_6 \\ 0 \\ 0 \end{pmatrix}, \quad v = \begin{pmatrix} v_3 \\ 0 \\ 0 \\ v_6 \end{pmatrix}. \tag{4.49}$$

From this arrangement of the edges and nodes, it follows that the sub-matrix, $\bar{H}_\mathcal{T}$, which maps non-inlet vertices to edges in set $\mathcal{E}_\mathcal{T}$, is invertible.

Since there are two nodes in the system with demand, the total flow leaving the network can be written as

$$\sigma(t) = -\bar{d}_3(t) - \bar{d}_6(t). \tag{4.50}$$

Having the input pressures and the parameters of the network, the output pressures, i.e. the pressures in the non-inlet vertices can be calculated. However, first recall *Equation: (4.35)*,

$$g(q_{\mathcal{C}}) = f_{\mathcal{C}}(q_{\mathcal{C}}) - A_1(\hat{p} + \hat{h}) + A_2^T f_{\mathcal{T}}(A_2 q_{\mathcal{C}} + A_3 \bar{d}) = 0. \tag{4.51}$$

where

$$A_1 = \hat{H}_{\mathcal{C}}^T - \bar{H}_{\mathcal{C}}^T \bar{H}_{\mathcal{T}}^{-T} \hat{H}_{\mathcal{T}}^T,$$
$$A_2 = -\bar{H}_{\mathcal{T}}^{-1} \bar{H}_{\mathcal{C}},$$
$$A_3 = \bar{H}_{\mathcal{T}}^{-1}.$$

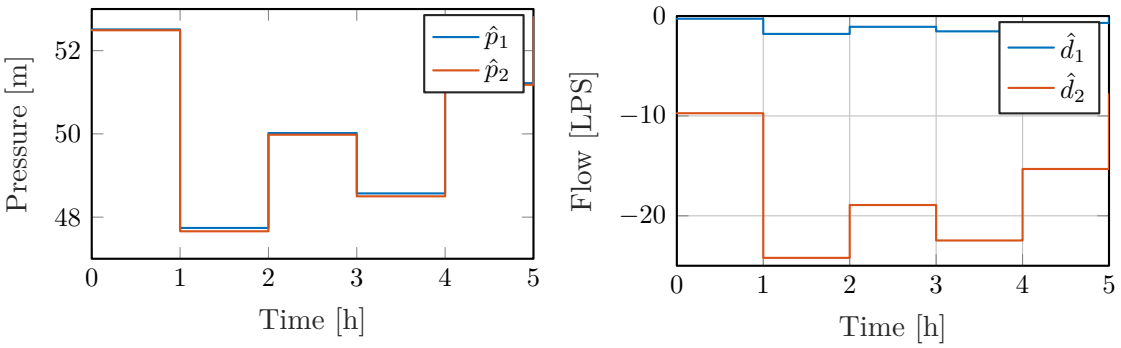
On account of non-linearity in *Equation: (4.51)*, it is not possible to solve the network analysis problem analytically. Instead, iterative numerical solution methods are used. As $g(q_{\mathcal{C}})$ is differentiable with respect to $q_{\mathcal{C}}$, gradient-based root finding algorithms such as Newton-Raphson method can be used. With iterative methods, initial values of flows are repeatedly adjusted until the difference between two successive iterates is within an acceptable tolerance. Furthermore, we know from the homogeneity and monotonicity property of $g(q_{\mathcal{C}})$ that the function is a homeomorphism in $q_{\mathcal{C}}$, therefore its root is unique. Using gradient-based searching methods, and squaring $g(q_{\mathcal{C}})$

$$2\frac{\partial g^T(q_{\mathcal{C}})}{\partial q_{\mathcal{C}}} g(q_{\mathcal{C}}) = 0, \tag{4.52}$$

the unique solution can be found. Furthermore, if $g^T(q_{\mathcal{C}})g(q_{\mathcal{C}})$ is convex, the solution is the global minimum. By solving *Equation: (4.51)* for $a_{\mathcal{C}}$, the non-inlet pressures and all flows in the network can be calculated in terms of the input pressures and the total demand in the network. In order to obtain these values, the previously-derived output equation in *Equation: (4.46)* can be used.

In the simulation, the most simple case is considered, when the total flow demand in the network varies, however the distribution among vertices remains the same. In this case, v is a constant vector, and the base demands in all vertices are the same.

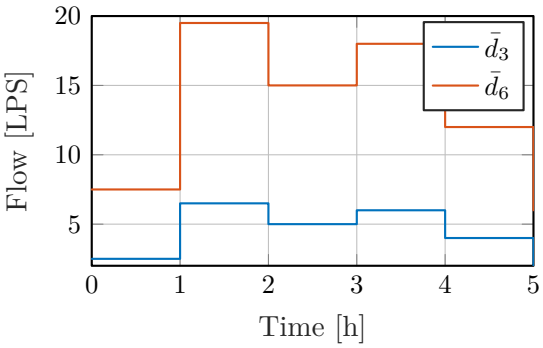
The variation curve for the input pressures are set the same in both EPANET and in the simulator.



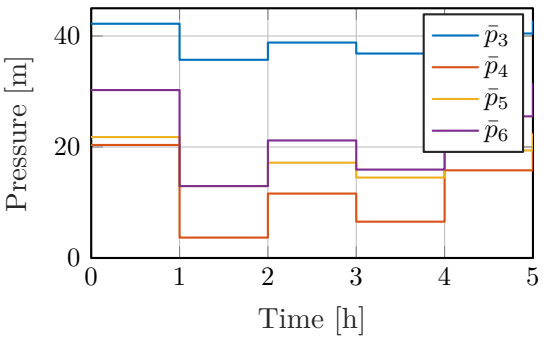
(a) Inlet pressures, ($\hat{p}$)

(b) Inlet demands, ($\hat{d}$)

Figure 4.2: Signals describing the input pressures(left) and flows(right) of the pumping stations.

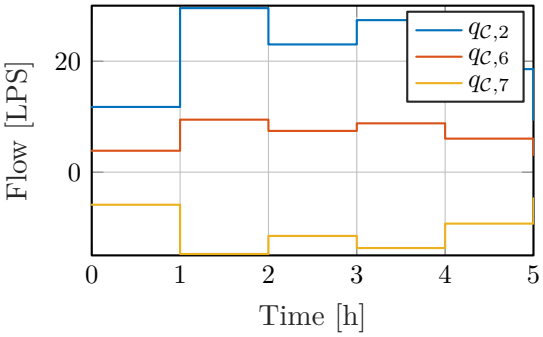


(a) Non-inlet demands, ($\bar{d}$)

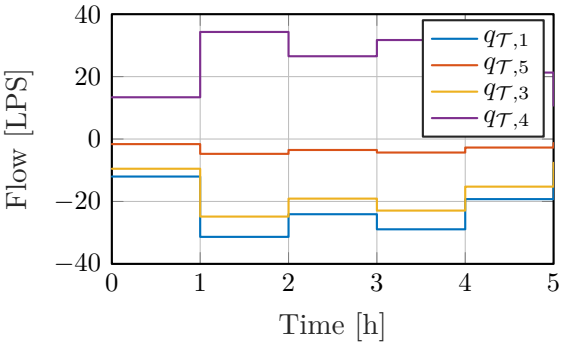


(b) Non-inlet pressures, ($\bar{p}$)

Figure 4.3: Signals describing the demand flows by the end-users(left) and output pressures(right) in the network.



(a) Flows in $\mathcal{C}$, ($q_{\mathcal{C}}$)



(b) Flows in $\mathcal{T}$, ($q_{\mathcal{T}}$)

Figure 4.4: Signals describing the flows in all pipes in the network.

4.5 Inclusion of elevated reservoirs

When a WT is being attached to an existing pipe network, certain properties of the previous system must be modified. Regarding the underlying graph of the network, when a WT is attached, the vertex to which the tank is connected becomes a vertex with a demand. However, the demand flow which describes the filling and emptying process of the tank, in this case, is not directly related to any user consumption profile, since flow can go into and come out of the tank. Therefore, the constraint on demand flows, described in *Equation: (4.21)* is not true in case of elevated reservoirs, as the demand can be both positive or negative. For this reason, demands on the reservoirs are treated as inputs and added to the nodes marked with hats $\hat{p}$ and $\hat{d}$. The nodes on which demand of elevated reservoirs are defined are selected from the input flows of the pumps such that

$$\hat{d} = F\hat{d}_t + G\hat{d}_c, \tag{4.53a}$$

where

$\hat{d}_t \in \mathbb{R}^{(l\times 1)}$ is the vector including the nodal demands of the tanks,
$\hat{d}_c \in \mathbb{R}^{(c\times 1)}$ is the the vector including the nodal demands of the pump inputs,
$F^T \in \mathbb{R}^{(l\times c+l)}$ is a mapping which selects the nodes belonging to tanks,
$G^T \in \mathbb{R}^{(c\times c+l)}$ is a mapping which selects the nodes belonging to pump inputs.

The same partitioning is done for the input pressures and elevation, regarding pumps and tanks

$$\hat{h} = F\hat{h}_t + G\hat{h}_c, \tag{4.53b}$$

$$\hat{p} = F\hat{p}_t + G\hat{p}_c, \tag{4.53c}$$

where

$\hat{h}_t \in \mathbb{R}^{(l\times 1)}$ is the vector including the elevation of the tanks,
$\hat{h}_c \in \mathbb{R}^{(c\times 1)}$ is the the vector including the elevation of the pump stations,
$\hat{p}_t \in \mathbb{R}^{(l\times 1)}$ is the vector including the absolute pressures in the tanks,
$\hat{p}_c \in \mathbb{R}^{(c\times 1)}$ is the the vector including the absolute pressures of the pump inputs.

With the inclusion of tanks, the network is not only constrained by the static pressure and flow relation, as in case of the model without tanks, but also governed by the dynamic equation describing the WT. These dynamics act as integrators on the flows which go in or out of the tank, $\hat{d}_t$. In terms of $\hat{d}_t$, the dynamics of the tanks set the pressure contribution, $\hat{p}_t$, as an input to the distribution system. The block diagram of such system is shown in *Figure 4.5.*

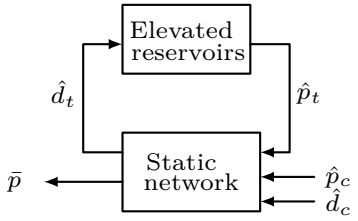


Figure 4.5: Block diagram of the system with WTs.

where the 'Elevated reservoirs' block represents the subsystem with dynamics, and the 'Static network' subsystem is governed by the algebraic flow and pressure relations which

describe the pipe network.

When elevated reservoirs are introduced in the network, the system usually operates such that the control inputs are the flows, $\hat{d}$, as this is the most practical and robust way to control the network. In fact, this is the case regarding the Randers WSS, as the two pumping stations which are filling up the tanks are controlled by flow, as it was explained in Section *2.2.1: Waterworks and pumping stations.*

This modelling approach is different, however, from what is handled in case of a network without a tank, since instead of using the pressures, $\hat{p}$, the inlet flows, $\hat{d}_c$, are considered as inputs. Furthermore, regarding the dynamics, the filling and emptying of a tank is dependent on how much flow is delivered by the pumps. This inlet flow is provided by multiple pumping stations, however, a distinction has to be made between the presence of single or multiple WTs in the network. In the former case, the flows delivered by the pumps are filling or emptying the one and only tank in the network. In the latter case, however, the model needs to be able to handle the flow distribution among the different WTs, meaning that different pumping stations can have different filling or emptying effects on the different WTs. Therefore, in the following, the two different approaches are presented.

4.6 Multi-inlet, single-WT model

Recalling the dynamic equation of one tank, in *Equation: (4.12)*

$$\dot{\hat{p}}_t = -\tau \hat{d}_t, \tag{4.54}$$

it is seen that the flow, $\hat{d}_t$, needs to be expressed in terms of the inputs, $\hat{d}_c$. In order to express the demand regarding a tank, the whole network is treated as one node, as illustrated in *Figure 4.6.*

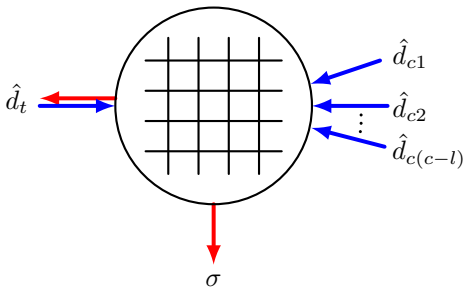


Figure 4.6: Mass balance on the network with one WT.

According to the mass-balance in the network, and using the corresponding partitioning as shown in *Figure 4.6*, the balance equation on all demand flows can be written in the form of

$$1^T d = 1^T \begin{pmatrix} \bar{d} \\ \hat{d} \end{pmatrix} = 1^T \begin{pmatrix} -v\sigma \\ \hat{d}_c \\ \hat{d}_t \end{pmatrix} = 0 \tag{4.55}$$

Now, expressing $\hat{d}_t$ from *Equation: (4.55)* yields

$$\hat{d}_t = \sigma - 1^T \hat{d}_c, \tag{4.56}$$

where it is shown, that if the control flows, $\hat{d}_c$, are higher than the total non-inlet demand, σ, the tank is being filled, and the tank is emptied if $\hat{d}_c$ is lower than σ. Thus, the system dynamics can be reformulated such that

$$\dot{\hat{p}}_t = -\tau(\sigma - 1^T \hat{d}_c). \tag{4.57}$$

It is important to point out that this first order differential equation in *Equation: (4.57)*, describes the mass-balance model of a single tank system, where the input flows equal to the overall demand in the network and the rate of change in storage in the tank. The result is not surprising, since according to *Equation: (4.57)*, the water level, or pressure, fluctuates according to the in- and outflow of water, which relates to the pumping effort and demand in the network.

In the single-WT modelling case, the dynamics can be described by a one-state, scalar equation, however the model is restricted to only one WT. As shown before, we know that this is not the case in the Randers WSS. Therefore, a more general description of the network is required, without restriction on the number of elevated reservoirs.

4.7 Multi-inlet, multi-WT model

In case there are more than one WTs, the model based on the mass-balance equation, proposed in *Equation: (4.55)* is not applicable due to the fact that the pressures delivered by the tanks have to be balanced, similarly as in a simple connected-volume system. Therefore, a model framework is required which can handle the dynamics of multiple tanks, constrained by the static network and driven by the input flows of the pumping stations. In order to derive such a model, the partitioning of the network is reconsidered. As described previously, the incidence matrix is partitioned such that

$$H = \begin{pmatrix} \bar{H}_{\mathcal{T}} & \bar{H}_{\mathcal{C}} \\ \hat{H}_{\mathcal{T}} & \hat{H}_{\mathcal{C}} \end{pmatrix}, \tag{4.58}$$

where $\bar{H}_{\mathcal{T}}$ is invertible. This partitioning rule on the edges is kept, however in the following model description let us slightly modify the notation and partitioning of the vertices. Now let

$$\mathcal{V} = \{\bar{\mathcal{V}}, \hat{\mathcal{V}}\}, \tag{4.59}$$

where

$\hat{\mathcal{V}} = \{\hat{v}_1, ..., \hat{v}_l\}$ represents the vertices corresponding to the points where elevated reservoirs are connected,

and

$\bar{\mathcal{V}} = \{\bar{v}_1, ..., \bar{v}_{n-l}\}$ represents the remaining vertices in the graph.

We originally defined $\bar{d}$ to describe non-inlet demands in the network and $\hat{d}$ to describe inlet-flows of the pumping stations, along with the demands of the WTs. With the redefined notation and partitioning of vertices, we make sure that the vertices related to tanks are in the set $\hat{\mathcal{V}}$ and all other points in the network are in $\bar{\mathcal{V}}$. So now, $\hat{d}$ describes flows regarding the WTs. In order to express the controlled inlet and non-inlet points in the network, the vectors related to $\bar{\mathcal{V}}$, $\bar{p}$ and $\bar{d}$, are further partitioned such that

$$\bar{p} = K\bar{p}_{\mathcal{K}} + D\bar{p}_{\mathcal{D}}, \tag{4.60a}$$

$$\bar{d} = K\bar{d}_{\mathcal{K}} + D\bar{d}_{\mathcal{D}}, \tag{4.60b}$$

where

$\bar{p}_{\mathcal{K}}$	is the vector including the inlet pressures ,
$\bar{p}_{\mathcal{D}}$	is the vector including the non-inlet pressures,
$\bar{d}_{\mathcal{K}}$	is the vector including the controlled inlet flows,
$d_{\mathcal{D}} = -v_{\mathcal{D}}\sigma$	is the vector including the non-inlet demands,
$K^T \in \mathbb{R}^{(c\times n-l)}$	is a mapping which selects the nodes belonging to the inlets,
$D^T \in \mathbb{R}^{(n-l-c\times n-l)}$	is a mapping which selects the remaining vertices.

With this partitioning of vertices, similarly to the model presented in Section *4.4: Multi-inlet model*, Kirchhoff's and Ohm's law can be formulated. Recalling Kirchhoff's vertex law

$$\begin{pmatrix} \bar{H}_{\mathcal{T}}^T & \hat{H}_{\mathcal{T}}^T \\ \bar{H}_{\mathcal{C}}^T & \hat{H}_{\mathcal{C}}^T \end{pmatrix} \begin{pmatrix} q_{\mathcal{T}} \\ q_{\mathcal{C}} \end{pmatrix} = \begin{pmatrix} \bar{d} \\ \hat{d} \end{pmatrix}, \tag{4.61}$$

and recalling *Equation: (4.29)*, Ohm's law is given

$$\begin{pmatrix} f_{\mathcal{T}}(q_{\mathcal{T}}) \\ f_{\mathcal{C}}(q_{\mathcal{C}}) \end{pmatrix} = \begin{pmatrix} \bar{H}_{\mathcal{T}}^T & \hat{H}_{\mathcal{T}}^T \\ \bar{H}_{\mathcal{C}}^T & \hat{H}_{\mathcal{C}}^T \end{pmatrix} \begin{pmatrix} \bar{p} + \bar{h} \\ \hat{p} + \hat{h} \end{pmatrix}. \tag{4.62}$$

In order to get an expression for $\hat{d}$, describing the flows regarding the WTs, lets define matrix Ω as follows

$$\Omega = \begin{pmatrix} -\hat{H}_{\mathcal{T}}\bar{H}_{\mathcal{T}}^{-1} & I \end{pmatrix}. \tag{4.63}$$

Multiplying *Equation: (4.61)* with Ω from the left-hand side, $\hat{d}$ can be expressed in the form of

$$\hat{d} = (\hat{H}_{\mathcal{C}} - \hat{H}_{\mathcal{T}}\bar{H}_{\mathcal{T}}^{-1}\bar{H}_{\mathcal{C}})q_{\mathcal{C}} + \hat{H}_{\mathcal{T}}\bar{H}_{\mathcal{T}}^{-1}\bar{d}. \tag{4.64}$$

As shown in *Equation: (4.64)*, the vertex law describes the dependencies of the WT demands on $\bar{d}$, which now consists of the pump inlet flows and the non-inlet demands in the rest of the network, and also shows the dependencies on $q_{\mathcal{C}}$. Recalling and rewriting the equation governing elevated reservoirs in matrix form, we can write the following

$$\Lambda\dot{\hat{p}} = -\hat{d}, \tag{4.65}$$

where $\Lambda = diag(\frac{1}{\tau_1}, ..., \frac{1}{\tau_l}) \in \mathbb{R}^{l\times l} \succ 0$.

Inserting $\hat{d}$ in *Equation: (4.64)*, into *Equation: (4.65)* leads to

$$\Lambda\dot{\hat{p}} = -(\hat{H}_{\mathcal{C}} - \hat{H}_{\mathcal{T}}\bar{H}_{\mathcal{T}}^{-1}\bar{H}_{\mathcal{C}})q_{\mathcal{C}} - \hat{H}_{\mathcal{T}}\bar{H}_{\mathcal{T}}^{-1}\bar{d}. \tag{4.66}$$

Now, with the partitioning of $\bar{d}$, the inlet and non-inlet demands are inserted into *Equation: (4.66)*, resulting in the governing expression of the WT dynamics

$$\Lambda\dot{\hat{p}} = -(\hat{H}_{\mathcal{C}} - \hat{H}_{\mathcal{T}}\bar{H}_{\mathcal{T}}^{-1}\bar{H}_{\mathcal{C}})q_{\mathcal{C}} - \hat{H}_{\mathcal{T}}\bar{H}_{\mathcal{T}}^{-1}K\bar{d}_{\mathcal{K}} + \hat{H}_{\mathcal{T}}\bar{H}_{\mathcal{T}}^{-1}Dv_{\mathcal{D}}\sigma. \tag{4.67}$$

As shown in *Equation: (4.67)*, the dynamics are now in terms of the inputs, $\bar{d}_{\mathcal{K}}$, the overall demand in the network, σ and $q_{\mathcal{C}}$. Recalling *Equation: (4.35)* in Section *4.4: Multi-inlet model* and using the abbreviations for the corresponding matrices, Ohm's law has been reformulated such that

$$f_{\mathcal{C}}(q_{\mathcal{C}}) - A_1(\hat{p} + \hat{h}) + A_2^T f_{\mathcal{T}}(A_2 q_{\mathcal{C}} + A_3\bar{d}) = 0. \tag{4.68}$$

where

$$A_1 = \hat{H}_{\mathcal{C}}^T - \bar{H}_{\mathcal{C}}^T \bar{H}_{\mathcal{T}}^{-T} \hat{H}_{\mathcal{T}}^T,$$
$$A_2 = -\bar{H}_{\mathcal{T}}^{-1} \bar{H}_{\mathcal{C}},$$
$$A_3 = \bar{H}_{\mathcal{T}}^{-1}.$$

Now, using again that $\bar{d} = K\bar{d}_{\mathcal{K}} + D\bar{d}_{\mathcal{D}}$, the following implicit expression is formulated which enables us to calculate $q_{\mathcal{C}}$

$$f_{\mathcal{C}}(q_{\mathcal{C}}) - A_1(\hat{p} + \hat{h}) + A_2^T f_{\mathcal{T}}(A_2 q_{\mathcal{C}} + A_3 K\bar{d}_{\mathcal{K}} - A_3 D v_{\mathcal{D}}\sigma) = 0. \tag{4.69}$$

Along with the implicit expression for $q_{\mathcal{C}}$ in *Equation: (4.69)*, *Equation: (4.67)* describes the dynamics of the system, taking into account the static constraints by the demands. Therefore, the structure of the overall Multi-inlet, multi-WT model dynamics can be summarized as follows

$$\begin{cases} \Lambda\dot{\hat{p}}(t) = -(\hat{H}_{\mathcal{C}} - \hat{H}_{\mathcal{T}}\bar{H}_{\mathcal{T}}^{-1}\bar{H}_{\mathcal{C}})q_{\mathcal{C}}(t) - \hat{H}_{\mathcal{T}}\bar{H}_{\mathcal{T}}^{-1}K\bar{d}_{\mathcal{K}}(t) + \hat{H}_{\mathcal{T}}\bar{H}_{\mathcal{T}}^{-1}Dv_{\mathcal{D}}(t)\sigma(t), \\ f_{\mathcal{C}}(q_{\mathcal{C}}(t)) - A_1(\hat{p}(t) + \hat{h}) + A_2^T f_{\mathcal{T}}(A_2 q_{\mathcal{C}}(t) + A_3 K\bar{d}_{\mathcal{K}}(t) - A_3 D v_{\mathcal{D}}(t)\sigma(t)) = 0. \end{cases} \tag{4.70}$$

Equation: (4.70) is a system of first-order Ordinary Differential Equations(ODE) with a constraint, specifying how the system evolves in time, given initial values of the states $\hat{p}$, the inputs $\bar{d}_{\mathcal{K}}$, and parametrized by the time-varying parameter $v_{\mathcal{D}}$. The constraint which is given by the algebraic equations regarding the static network, defines the manifold on which the trajectory of the ODE solution can be found.

The states of the system are the pressures in the WTs. Therefore when the model equations are solved for $q_{\mathcal{C}}$, initial information of the pressure values, $\hat{p}$, is required. Equivalently, the initial water levels in the tanks have to be known, as for any ODE. Furthermore, the initial value should be inside the set described by the algebraic constraint.

In order to calculate the corresponding pressure inputs, $\bar{p}_{\mathcal{K}}$, first $\bar{p}$ is expressed from Ohm's law as proposed in Section *4.4: Multi-inlet model*

$$\bar{p} = \bar{H}_{\mathcal{T}}^{-T} f_{\mathcal{T}}(-\bar{H}_{\mathcal{T}}^{-1}\bar{H}_{\mathcal{C}}q_{\mathcal{C}} + \bar{H}_{\mathcal{T}}^{-1}\bar{d}) - \bar{H}_{\mathcal{T}}^{-T}\hat{H}_{\mathcal{T}}^T(\hat{p} + \hat{h}) - \bar{h}. \tag{4.71}$$

Now using that $\bar{p}_{\mathcal{K}} = K^T\bar{p}$, the pressure input equation becomes

$$\bar{p}_{\mathcal{K}} = K^T\bar{H}_{\mathcal{T}}^{-T} f_{\mathcal{T}}(A_2 q_{\mathcal{C}} + A_3 K\bar{d}_{\mathcal{K}} - A_3 D v_{\mathcal{D}}\sigma) - K^T\bar{H}_{\mathcal{T}}^{-T}\hat{H}_{\mathcal{T}}^T(\hat{p} + \hat{h}) - K^T\bar{h}. \tag{4.72}$$

Furthermore, using that $\bar{p}_{\mathcal{D}} = D^T\bar{p}$, the pressure output equation becomes

$$\bar{p}_{\mathcal{D}} = D^T\bar{H}_{\mathcal{T}}^{-T} f_{\mathcal{T}}(A_2 q_{\mathcal{C}} + A_3 K\bar{d}_{\mathcal{K}} - A_3 D v_{\mathcal{D}}\sigma) - D^T\bar{H}_{\mathcal{T}}^{-T}\hat{H}_{\mathcal{T}}^T(\hat{p} + \hat{h}) - D^T\bar{h}. \tag{4.73}$$

Equivalently, *Equation: (4.70)* and *Equation: (4.73)* can be described in a more abstract form, considering a non-linear State-Space(SS) system structure such that

$$\begin{cases} \dot{\hat{p}} = g_{v_{\mathcal{D}}}(\bar{d}_{\mathcal{K}}, \sigma, q_{\mathcal{C}}) \\ \bar{p}_{\mathcal{K}} = h_{v_{\mathcal{D}}}(\hat{p}, \bar{d}_{\mathcal{K}}, \sigma, q_{\mathcal{C}}) \end{cases} \tag{4.74}$$
$$s.t. \quad q_{\mathcal{C}} = q_{\mathcal{C}}((\hat{p} + \hat{h}), \bar{d}_{\mathcal{K}}, \sigma)$$

with $\hat{p} \in \mathbb{R}^l$ the state vector, $\bar{d}_{\mathcal{K}} \in \mathbb{R}^c$ the input vector, $\sigma \in \mathbb{R}_+$ the measurable disturbance, $v_{\mathcal{D}} \in \mathbb{R}^{(n-l-c)}$ the time-varying parameter and $\bar{p}_{\mathcal{K}} \in \mathbb{R}^c$ the output vector.

Although the ODE describing the dynamics is linear, the output and constraint equations

are non-linear. Therefore the system is treated as a non-linear problem since the input-output relation is non-linear.

It is worth noting that this model with the flow inputs and the dynamics of the WTs results in a scalar dynamic equation in case of one tank. Therefore, the proposed restructuring of the model here is suitable for describing one tank systems as well, and capable of replacing the mass-balance based system dynamics, proposed in Section *4.6: Multi-inlet, single-WT model*. Therefore, the single-WT network model, presented in Section *4.6: Multi-inlet, single-WT model* is a special case of the Multi-inlet, Multi-WT model framework derived here.

4.8 Model comparison

In this chapter, three different model arrangements have been proposed and two of them have been derived for control purposes. The main goal was to develop a model of a WSS with elevated reservoirs, thereby the Multi-inlet model described in the first column in *Table: 4.1*, was a baseline model to extend towards inclusion of WTs. Therefore, as a summary of the network modelling, the different approaches and system structures are compared. It should be noted again, that notation has been redefined through the modelling, therefore symbols for the Multi-inlet and Multi-inlet, single-WT models are different from the one describing the Multi-inlet, multi-WT model. With this in mind, the properties of the three network models are summed up in the following table

	Multi-inlet model	**Multi-inlet/single-WT, based on mass-balance**	**Multi-inlet/multi-WT, general model**
Description	A model, describing a system without WTs, therefore describing a static network with time-varying parameters.	A model, describing the system dynamics with only one WT, relying on the mass-balance in the system, constrained by the static network. The parameter of the system is also time-varying.	A model describing the system dynamics with multiple WTs, constrained by the static network with time-varying parameter.
Input	$\hat{p}$ - Pressures of the multiple pumping stations.	$\hat{d}_c$ - Flows of the multiple pumping stations.	$\bar{d}_{\mathcal{K}}$ - Flows of the multiple pumping stations.
States	-	$\hat{p}_t$ - The pressure in one tank.	$\hat{p}$ - The pressures in multiple tanks.
Output	$\bar{p}$ - Pressures in the non-inlet points.	$\bar{p}$ - Pressures in the non-inlet points	$\bar{p}$ - Pressures in the non-inlet points.
Advantages	The network is without dynamics, therefore all solutions are steady-state values.	The network is controlled with flow, which makes the system more robust towards pressure errors in the WT. The dynamics are governed by a scalar equation, relying on the input flows.	The control can handle the pressure balance among multiple WTs.
Governing equations	$\bar{p} = h_v(\hat{p}, \sigma, q_{\mathcal{C}})$ $q_{\mathcal{C}} = q_{\mathcal{C}}(\hat{p}, \sigma)$	$\dot{\hat{p}}_t = e(\sigma, \hat{d}_c)$ + static network equations	$\begin{cases} \dot{\hat{p}} = g_{v_{\mathcal{D}}}(\bar{d}_{\mathcal{K}}, \sigma, q_{\mathcal{C}}) \\ \bar{p}_{\mathcal{K}} = h_{v_{\mathcal{D}}}(\hat{p}, \bar{d}_{\mathcal{K}}, \sigma, q_{\mathcal{C}}) \end{cases}$ $q_{\mathcal{C}} = q_{\mathcal{C}}((\hat{p}+\hat{h}), \bar{d}_{\mathcal{K}}, \sigma)$
Section reference	Section *4.4: Multi-inlet model*	Section *4.6: Multi-inlet, single-WT model*	Section *4.7: Multi-inlet, multi-WT model*

Table 4.1: Summary of the network models.

In the next chapters, the Multi-inlet, multi-WT network model is discussed, as it is the suitable model for describing the water network in Randers.

Part II

Identification

5. System identification

In this chapter, the Multi-inlet, Multi-WT model, which has been derived based on first principles in Chapter 4: System Modelling, is reformulated such that it is suitable for system identification. First, the structure of the model is discussed, then a Neural Network(NN) based identification method is presented. In the first case, the identification is carried out on a simple example, then on the Randers WSS EPANET model. In the second case, the identification is carried out based on measurements from the real-world network.

5.1 Model structure of the Multi-inlet, Multi-WT system

In Chapter *4: System Modelling*, the model of a multi-inlet WSS with the extension of multiple WTs has been derived. The presented model has been given by a non-linear SS representation, which consists of the equations describing the dynamics, the outputs and the constraints on the network. As the model derivation is based on first principles, some insight into the structure and the input-output relation is available.

5.1.1 Output equation

The model needs to be identified in order to utilize it for control purposes. Therefore, let us recall the governing equations. The output vector $\bar{p}_{\mathcal{K}}$ of the inlet pressures is given in discrete form such that

$$\bar{p}_{\mathcal{K},k} = K^T \bar{H}_{\mathcal{T}}^{-T} f_{\mathcal{T}}(A_2 q_{\mathcal{C},k} + A_3 K \bar{d}_{\mathcal{K},k} - A_3 D v_{\mathcal{D}} \sigma_k) - K^T \bar{H}_{\mathcal{T}}^{-T} \hat{H}_{\mathcal{T}}^T (\hat{p}_k + \hat{h}) - K^T \bar{h}, \quad (5.1)$$

where

$$A_2 = -\bar{H}_{\mathcal{T}}^{-1} \bar{H}_{\mathcal{C}},$$
$$A_3 = \bar{H}_{\mathcal{T}}^{-1},$$

and furthermore, $\bar{d}_{\mathcal{K},k}$ inlet flows, σ_k total demand, $v_{\mathcal{D}}$ distribution parameter, $q_{\mathcal{C},k}$ flows in set $\mathcal{C}$, $\hat{p}_k$ pressures in the WTs, $\hat{h}$ elevations of the WTs and $K^T \bar{h} = \bar{h}_{\mathcal{K}}$ the elevations of the pumping stations. The corresponding values of the pressures and flows in the network are evaluated at each time step k.

Additionally, let us recall the constraint on $q_{\mathcal{C}}$, and rewrite it in discrete-time form such that

$$f_{\mathcal{C}}(q_{\mathcal{C},k}) - A_1(\hat{p}_k + \hat{h}) + A_2^T f_{\mathcal{T}}(A_2 q_{\mathcal{C},k} + A_3 K \bar{d}_{\mathcal{K},k} - A_3 D v_{\mathcal{D}} \sigma_k) = 0. \quad (5.2)$$

It is important to point out that the model in Chapter *4: System Modelling* has been derived in a general manner, taking into account that $v_{\mathcal{D}}$ is a time-varying distribution parameter of the demands. However, in the further description, let us restrict ourselves and assume that $v_{\mathcal{D}}$ is constant. From the technical point of view, the identification becomes less complex, as in this case $v_{\mathcal{D}}$ is a linear constant parameter. Furthermore, there is no need to introduce $v_{\mathcal{D}}$ as a time-varying parameter on the EPANET data, however on the real measurement data, the possible demand variations might be experienced. In *Equation: (5.2)* and *Equation: (5.1)* the assumption of $v_{\mathcal{D}}$ being constant is already taken into account, therefore $v_{\mathcal{D}}$ does not have any time index.

The constraint on $q_{\mathcal{C}}$ in *Equation: (5.2)* is given by an implicit expression for which an analytical soulution has not been derived. The explicit solution for the constraint has a structure like *Equation: (5.3)*, however we do not know how it looks exactly.

$$q_{\mathcal{C},k} = q_{\mathcal{C}}((\hat{p}_k + \hat{h}), \bar{d}_{\mathcal{K},k}, \sigma_k). \tag{5.3}$$

It is shown in *Equation: (5.3)* that the $q_{\mathcal{C}}$ flows in the network depend on the same physical measures, i.e. the same variables as the outputs $\bar{p}_{\mathcal{K}}$. By substituting *Equation: (5.3)* into *Equation: (5.1)*, we get the following output equation

$$\begin{aligned}\bar{p}_{\mathcal{K},k} =& K^T \bar{H}_{\mathcal{T}}^{-T} f_{\mathcal{T}}(A_2 q_{\mathcal{C}}((\hat{p}_k + \hat{h}), \bar{d}_{\mathcal{K},k}, \sigma_k) + A_3 K \bar{d}_{\mathcal{K},k} - A_3 D v_{\mathcal{D}} \sigma_k) \\ & - K^T \bar{H}_{\mathcal{T}}^{-T} \hat{H}_{\mathcal{T}}^T (\hat{p}_k + \hat{h}) - \bar{h}_{\mathcal{K}}.\end{aligned} \tag{5.4}$$

In *Equation: (5.4)*, the pressure head in the pumping stations $\bar{p}_{\mathcal{K},k}$ is given by the expression on the right-hand side. Let us write *Equation: (5.4)* in a form where the non-linear expression on the right-hand side is replaced with a non-linear function $\tilde{f}_1(\cdot)$, which has an unknown structure but has the same variables in the argument. Thus, a reformulated output equation can be given such that

$$\bar{p}_{\mathcal{K},k} = \tilde{f}_1((\hat{p}_k + \hat{h}), \bar{d}_{\mathcal{K},k}, \sigma_k) + \tilde{a}_{\mathcal{K}}(\hat{p}_k + \hat{h}) - \bar{h}_{\mathcal{K}}, \tag{5.5}$$

The output equation described in *Equation: (5.5)* is a mapping defined by the non-linear function $\tilde{f}_1$ and the linear term, which maps the input set, $u_k = (\hat{p}_k + \hat{h}\ \bar{d}_{\mathcal{K},k}\ \sigma_k)^T$ to the outputs $\bar{p}_{\mathcal{K},k}$. Typically, the pressure is measured in the WTs not the total head, therefore the elevation $\bar{h}_{\mathcal{K}}$ is not added to $\bar{p}_{\mathcal{K},k}$ on the left-hand side in *Equation: (5.5)*. In the input set, the total consumption can be calculated according to the mass-balance in the whole network such that

$$\sigma_k = 1^T \hat{d}_k + 1^T \bar{d}_{\mathcal{K},k}. \tag{5.6}$$

In *Equation: (5.6)*, we assume that the flows in the WTs are measured, as the demand flows $\bar{d}_{\mathcal{D},k}$ in the network are not measurable.

5.1.2 State equation

The state equation is a first-order system of ODEs, which has been formulated on the pressures $\hat{p}$ in the WTs. In order to give a discretization for the approximate solution of the ODEs, Euler-method is used. The Euler-method is the simplest Runge-Kutta method, which provides an acceptable precision for our problem[28]. Thus, the discretized state equation yields as follows

$$\Lambda \frac{1}{T_s}(\hat{p}_{k+1} - \hat{p}_k) = -(\hat{H}_{\mathcal{C}} - \hat{H}_{\mathcal{T}} \bar{H}_{\mathcal{T}}^{-1} \bar{H}_{\mathcal{C}}) q_{\mathcal{C},k} - \hat{H}_{\mathcal{T}} \bar{H}_{\mathcal{T}}^{-1} K \bar{d}_{\mathcal{K},k} + \hat{H}_{\mathcal{T}} \bar{H}_{\mathcal{T}}^{-1} D v_{\mathcal{D}} \sigma_k. \tag{5.7}$$

where

T_s is the sampling time. [h]

Substituting the expression for $q_{\mathcal{C}}$ into *Equation: (5.7)*, and expressing the predicted values of the WT pressures on the left-hand side, the following yields

$$\begin{aligned}\hat{p}_{k+1} =& T_s \Lambda^{-1} \big(- (\hat{H}_{\mathcal{C}} - \hat{H}_{\mathcal{T}} \bar{H}_{\mathcal{T}}^{-1} \bar{H}_{\mathcal{C}}) q_{\mathcal{C},k}((\hat{p}_k + \hat{h}), \bar{d}_{\mathcal{K},k}, \sigma_k) \\ & - \hat{H}_{\mathcal{T}} \bar{H}_{\mathcal{T}}^{-1} K \bar{d}_{\mathcal{K},k} + \hat{H}_{\mathcal{T}} \bar{H}_{\mathcal{T}}^{-1} D v_{\mathcal{D}} \sigma_k \big) + \hat{p}_k.\end{aligned} \tag{5.8}$$

The state equation in *Equation: (5.7)* describes the relation between the the flows $q_{\mathcal{C},k}$, the total head in the WTs $(\hat{p}_k + \hat{h})$ and the total consumption σ_k. However, by substituting the $q_{\mathcal{C},k}$ flows with their non-linear expression, the structure of the state equation is not a

linear combination of the corresponding signals anymore. Therefore, let us write *Equation: (5.8)* in a form where the non-linear terms are described by a non-linear $\tilde{f}_2(\cdot)$ function with unknown structure. Thus, a reformulated state equation can be given such that

$$\hat{p}_{k+1} = \tilde{f}_2((\hat{p}_k + \hat{h}), \bar{d}_{\mathcal{K},k}, \sigma_k) + \tilde{a}_2 \bar{d}_{\mathcal{K},k} + \tilde{a}_3 \sigma_k + \hat{p}_k, \tag{5.9}$$

where $\tilde{f}_2$ is a non-linear function, $\tilde{a}_2$ and $\tilde{a}_3$ are parameters of the inlet flows $\bar{d}_{\mathcal{K},k}$ and total consumption σ_k.

5.2 RBFNN model of the Multi-inlet,Multi-WT system

As a result of substituting the constraints on the flows $q_{\mathcal{C},k}$, the system description has been reduced to a non-linear State Space model with state equation given in *Equation: (5.9)* and output equation in *Equation: (5.5)*. The complete identification model is summarized in *Equation: (5.10)*.

$$\begin{cases} \hat{p}_{k+1} = \tilde{f}_2((\hat{p}_k + \hat{h}), \bar{d}_{\mathcal{K},k}, \sigma_k) + \tilde{a}_2 \bar{d}_{\mathcal{K},k} + \tilde{a}_3 \sigma_k + \hat{p}_k, \\ \bar{p}_{\mathcal{K},k} = \tilde{f}_1((\hat{p}_k + \hat{h}), \bar{d}_{\mathcal{K},k}, \sigma_k) + \tilde{a}_{\mathcal{K}}(\hat{p}_k + \hat{h}) - \bar{h}_{\mathcal{K}}. \end{cases} \tag{5.10}$$

The main goal of the system identification is to find a realization of the functions $\tilde{f}_1(\cdot)$ and $\tilde{f}_2(\cdot)$, furthermore to find the parameters $\tilde{a}_1$, $\tilde{a}_2$ and $\tilde{a}_3$. Therefore, the parameters need to be identified, such that the model is able to reproduce the approximate of the state derivatives $(\hat{p}_{k+1} - \hat{p}_k)$ and the total inlet head $\bar{p}_{\mathcal{K},k} + \bar{h}_{\mathcal{K}}$ from any input set $u_k = (\hat{p}_k + \hat{h}\ \bar{d}_{\mathcal{K},k}\ \sigma_k)^T$ within the operating regions where we are mapping from.

The identification model shown in *Equation: (5.10)* is an abstraction of the first principle model derived in Chapter *4: System Modelling*. For the constraint on $q_{\mathcal{C},k}$, existence has been implied, however exact structure has not been given. I.e. we know that the relationship exists but an analytical first principle solution for $q_{\mathcal{C},k}$ has not been derived. Therefore, by substituting the implicit expression of the constraint into the state and output equations, some of the insights on the structure of the model are lost. Thus, it is crucial to put a structure on the non-linear functions $\tilde{f}_1(\cdot)$ and $\tilde{f}_2(\cdot)$ in *Equation: (5.10)*.

From a practical point of view, it is beneficial to describe the system by a linear-in-the-parameters model. By restricting ourselves such that the structure of both functions $\tilde{f}_1(\cdot)$ and $\tilde{f}_2(\cdot)$ are linear in the parameters, the parameters of the model can be estimated by simple linear optimization methods, such as Least Squares(LS). For any linear optimization, only the parameters have to enter linearly, as the inputs can depend on any non-linear way on the input data sets.

By making the restriction on $\tilde{f}_1(\cdot)$ and $\tilde{f}_2(\cdot)$, the two non-linear terms will be approximated by some non-linear functions in both the state and output equations of the model. The tools for carrying out such identification procedure leads to the discussion of Radial Basis Functions(RBFs) and neural networks, which are introduced in *Appendix: E*. In *Appendix: E*, the main properties of such networks are explained in detail, therefore during the construction of the identification model, these properties are used and applied.

5.2.1 Output RBFNN

The output equation described in *Equation: (5.10)* is going to be approximated with RBFs in the form as shown in *Equation: (5.11)*, for two pressure inlets.

$$\bar{p}_{\mathcal{K}1,k} = \sum_{i=1}^{M} w_{\mathcal{K}1,i} \phi_i(u_k) + \sum_{j=1}^{l} \tilde{a}_{\mathcal{K}1,j}(\hat{p}_{j,k} + \hat{h}_j) + b_{\mathcal{K}1} - \bar{h}_{\mathcal{K}1}, \tag{5.11a}$$

$$\bar{p}_{\mathcal{K}2,k} = \sum_{i=1}^{M} w_{\mathcal{K}2,i} \phi_i(u_k) + \sum_{j=1}^{l} \tilde{a}_{\mathcal{K}2,j}(\hat{p}_{j,k} + \hat{h}_j) + b_{\mathcal{K}2} - \bar{h}_{\mathcal{K}2}, \tag{5.11b}$$

where $\bar{p}_{\mathcal{K},k} = (\bar{p}_{\mathcal{K}1,k}\ \bar{p}_{\mathcal{K}2,k})^T$ is the inlet pressure vector, $w_{\mathcal{K}1,i}$ is the output weight of the i^{th} RBF neuron $\phi_i(u_k)$, $\tilde{a}_{\mathcal{K}} = (\tilde{a}_{\mathcal{K}1}\ \tilde{a}_{\mathcal{K}2})^T$ is the parameter vector of the linear terms, $b_{\mathcal{K}} = (b_{\mathcal{K}1}\ b_{\mathcal{K}2})^T$ is the bias vector and $h_{\mathcal{K}} = (h_{\mathcal{K}1}\ h_{\mathcal{K}2})^T$ is a constant vector which describes the elevation of the pumping stations. Furthermore, l is the number of WTs and M is the number of RBFs with which the non-linear terms are approximated.

The NN model for the inlet pressure $\bar{p}_{\mathcal{K}1,k}$ is shown in *Figure 5.1*, with the corresponding parameters and weighs.

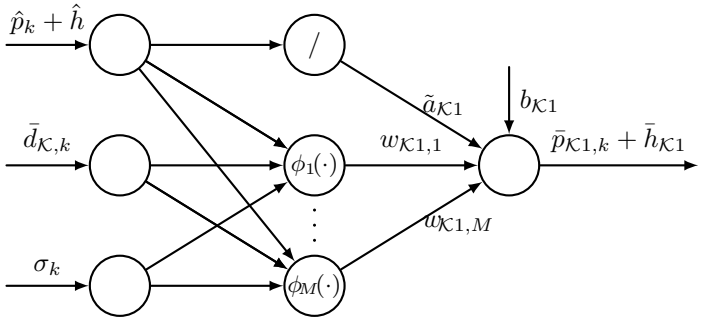


Figure 5.1: NN identification model of the total inlet head $\bar{p}_{\mathcal{K}1,k} + \bar{h}_{\mathcal{K}1}$.

In *Figure 5.1*, the first layer consists of the first layer neurons with the input vectors. The hidden layer consists of the set of RBFs and one linear neuron. The linear neuron defines a skip-layer connection in the model. This extra connection allows us to introduce the linear terms in our identification model, thereby letting $\hat{p}_k + \hat{h}$ directly influence the inlet pressures.

Furthermore, the output layer consists of the output neuron which computes $\bar{p}_{\mathcal{K}1,k}$ by summing the weighted RBFs, the weighted linear signals, the output bias and the elevation constant. For convenience, while identifying the model, the elevation constants are added to the measured pressure inlets $\bar{p}_{\mathcal{K}1,k}$. Thus, the identification for the total inlet head vector $\bar{p}_{\mathcal{K}1,k} + \bar{h}_{\mathcal{K}1}$ is carried out according to *Equation: (5.12)*

$$\bar{p}_{\mathcal{K}1,k} + \bar{h}_{\mathcal{K}1} = \theta_{\mathcal{K}1}^T \chi(u_k), \tag{5.12}$$

where the parameter vector $\theta_{\mathcal{K}1}$ and regression vector χ is given such that

$$\theta_{\mathcal{K}1} = \begin{pmatrix} w_{\mathcal{K}1,1} \\ \vdots \\ w_{\mathcal{K}1,M} \\ \tilde{a}_{\mathcal{K}1,1} \\ \vdots \\ \tilde{a}_{\mathcal{K}1,l} \\ b_{\mathcal{K}1} \end{pmatrix}, \quad \chi = \begin{pmatrix} \phi_1(u_k) \\ \vdots \\ \phi_M(u_k) \\ \hat{p}_{1,k} + \hat{h}_1 \\ \vdots \\ \hat{p}_{l,k} + \hat{h}_1 \\ 1 \end{pmatrix}. \tag{5.13}$$

With a given input measurement set u_k and inlet pressure measurements $\bar{p}_{\mathcal{K}1,k}$, the optimal solution for the parameters is calculated by using LS method, as *Equation: (5.12)* is a linear matrix equation. Thus, the parameter vector is given by *Equation: (5.14)*

$$\theta_{\mathcal{K}1}^T = (\bar{p}_{\mathcal{K}1,k} + \bar{h}_{\mathcal{K}1})\chi^\dagger(u_k), \tag{5.14}$$

where the Moore-Penrose pseudoinverse of the regression vector is computed.

In case of c pumping stations, the inlet pressure vector $\bar{p}_{\mathcal{K},k}$ is given as shown in *Equation: (5.15)*

$$\begin{pmatrix} \bar{p}_{\mathcal{K}1,k} + \bar{h}_{\mathcal{K}1} \\ \vdots \\ \bar{p}_{\mathcal{K}c,k} + \bar{h}_{\mathcal{K}c} \end{pmatrix} = \begin{pmatrix} \theta_{\mathcal{K}1} & \dots & \theta_{\mathcal{K}c} \end{pmatrix}^T \chi(u_k). \tag{5.15}$$

5.2.2 State RBFNN

The state equation, described in *Equation: (5.10)* is going to be approximated with RBFs in the same way as it has been done for the inlet pressures. Therefore, the approximate of the state derivative $\hat{p}_{k+1} - \hat{p}_k$ of Tank 1 and Tank 2 is given in the form as shown in *Equation: (5.16)*

$$\hat{p}_{\mathcal{S}1,k+1} - \hat{p}_{\mathcal{S}1,k} = \sum_{i=1}^{N} w_{\mathcal{S}1,i}\phi_i(u_k) + \sum_{j=1}^{c} \tilde{a}_{\mathcal{S}1_A,j}\bar{d}_{\mathcal{K}j,k} + \tilde{a}_{\mathcal{S}1_B}\sigma_k + b_{\mathcal{S}1}, \tag{5.16a}$$

$$\hat{p}_{\mathcal{S}2,k+1} - \hat{p}_{\mathcal{S}2,k} = \sum_{i=1}^{N} w_{\mathcal{S}2,i}\phi_i(u_k) + \sum_{j=1}^{c} \tilde{a}_{\mathcal{S}2_A,j}\bar{d}_{\mathcal{K}j,k} + \tilde{a}_{\mathcal{S}2_B}\sigma_k + b_{\mathcal{S}2}, \tag{5.16b}$$

where $w_{\mathcal{S}1,i}$ is the weight of the i^{th} RBF, $\tilde{a}_{\mathcal{S}1_A,j}$ is the parameter of the j^{th} inlet flow and $\tilde{a}_{\mathcal{S}1_B}$ is the parameter of the linear total demand. Furthermore, N is the number of RBFs with which the non-linear terms are approximated and c is the number of pumping stations.

The NN model for the state derivative approximate of one tank $\hat{p}_{\mathcal{S}1,k+1} - \hat{p}_{\mathcal{S}1,k}$, with the corresponding weight and parameters is shown in *Figure 5.2*.

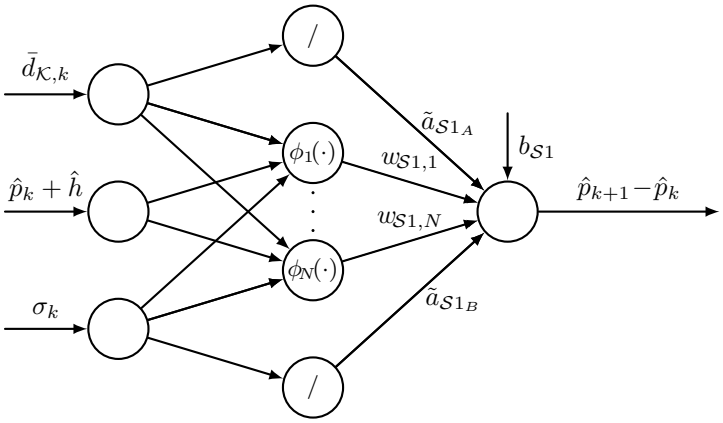


Figure 5.2: NN model of the state derivative approximate $\hat{p}_{\mathcal{S}1,k+1} - \hat{p}_{\mathcal{S}1,k}$.

The identification of the state derivative approximate is carried out according to *Equation: (5.17)*

$$\hat{p}_{\mathcal{S}1,k+1} - \hat{p}_{\mathcal{S}1,k} = \theta_{\mathcal{S}1}\chi(u_k), \tag{5.17}$$

where the parameter vector $\theta_{\mathcal{S}1}$ and regression vector χ is given such that

$$\theta_{\mathcal{S}1} = \begin{pmatrix} w_{\mathcal{S}1,1} \\ \vdots \\ w_{\mathcal{S}1,N} \\ \tilde{a}_{\mathcal{S}1_A,1} \\ \vdots \\ \tilde{a}_{\mathcal{S}1_A,c} \\ \tilde{a}_{\mathcal{S}1_B} \\ b_{\mathcal{S}1} \end{pmatrix}, \quad \chi = \begin{pmatrix} \phi_1(u_k) \\ \vdots \\ \phi_N(u_k) \\ \bar{d}_{\mathcal{K}1,k} \\ \vdots \\ \bar{d}_{\mathcal{K}c,k} \\ \sigma_k \\ 1 \end{pmatrix}. \tag{5.18}$$

The parameter matrix is computed equivalently as for the output equation with the Moore-Penrose pseudoinverse of the regression vector, shown in *Equation: (5.19)*.

$$\theta_{\mathcal{S}1}^T = (\hat{p}_{\mathcal{S}1,k+1} - \hat{p}_{\mathcal{S}1,k})\chi^\dagger(u_k). \tag{5.19}$$

In case of l WTs, the state derivative approximates $\hat{p}_{\mathcal{S}2,k+1} - \hat{p}_{\mathcal{S}2,k}$ are given as shown in *Equation: (5.20)*

$$\begin{pmatrix} \hat{p}_{\mathcal{S}1,k+1} - \hat{p}_{\mathcal{S}1,k} \\ \vdots \\ \hat{p}_{\mathcal{S}l,k+1} - \hat{p}_{\mathcal{S}l,k} \end{pmatrix} = \begin{pmatrix} \theta_{\mathcal{S}1} & \dots & \theta_{\mathcal{S}l} \end{pmatrix}^T \chi(u_k). \tag{5.20}$$

5.3 Identification of the Randers WSS

Bibliography

[1] T. Younos and T. Parece, *Sustainable Water Management in Urban Environments*. The Handbook of Environmental Chemistry, Springer International Publishing, 2016.

[2] Y. X. J.N. Tsitsiklis, "Pricing of fluctuations in electricity markets," *European Journal of Operational Research*, 2015.

[3] B. Appelbaum, "Water & sustainability (volume 4): Us electricity consumption for water supply & treatment—the next half century," 2002.

[4] N. R. Council *et al.*, *Public Water Supply Distribution Systems: Assessing and Reducing Risks–First Report*. National Academies Press, 2005.

[5] Miljøministeriet and Geus, *Water Supply in Denmark*. Danish action plan for promotion of eco-efficient technologies - Danish Lessons, 2009.

[6] M. Brdys and B. Ulanicki, *Operational Control of Water Systems: Structures, Algorithms, and Applications*. Prentice Hall, 1994.

[7] M. Beniston, *Environmental Change in Mountains and Uplands*. Key Issues in Environmental Change, Taylor & Francis, 2016.

[8] P. Swamee and A. Sharma, *Design of Water Supply Pipe Networks*. Wiley, 2008.

[9] K. Hoe, "Modeling and control of water supply network," *Danmarks Tekniske Universitet*, 2005.

[10] T. Walski and I. Haestad Methods, *Advanced water distribution modeling and management*. No. vb. 1, Haestead Press, 2003.

[11] M. Mølgaard and B. G. Pétursson, *Energy Optimization of Water Distribution Networks*. 2015. Master thesis, Aalborg University.

[12] N. Grigg, *Water, Wastewater, and Stormwater Infrastructure Management, Second Edition*. Taylor & Francis, 2012.

[13] L. Mays, "Water transmission and distribution," *American Water Works Association, Denver*, 2010.

[14] N. Council, D. Studies, W. Board, and C. Risks, *Drinking Water Distribution Systems: Assessing and Reducing Risks*. National Academies Press, 2007.

[15] D. Andersen, K. Balla, N. Christensen, I. Bolinaga, and S. Krogh, *Optimal Control for Water Distribution*. Aalborg Universitet, 2017.

[16] C. Kallesøe, "Fault detection and isolation in centrifugal pumps," *Department of Control Engineering, Aalborg Universitet*, 2005.

[17] Verdo. Last visited 17-09-2017, https://verdo.dk/da/vand/forsyningskunde-/om-verdo-vand, Title = Om Verdo Vand.

[18] G. S. (U.S.), *Water-supply Paper*. Geological Survey Water-supply Paper, U.S. Department of the Interior, Geological Survey, 1909.

[19] L. Johnson, *Geographic Information Systems in Water Resources Engineering*. CRC Press, 2016.

[20] Verdo A/S, *Ledningnetmodel for Verdo's forsyningsområde*. Verdo, 2010.

[21] E. Agency, *Epanet 2 Users Manual*. CreateSpace Independent Publishing Platform, 2016.

[22] D. Chen, *Sustainable Water Technologies*. Green Chemistry and Chemical Engineering, CRC Press, 2016.

[23] *Water Transmission and Distribution*. Principles and practices of water supply operations series, American Water Works Association, 2011.

[24] T. N. Jensen, C. S. Kallesøe, and R. Wisniewski, "Adaptive reference control for pressure management in water networks," *European Control Conference(ECC)*, vol. 3, July 2015.

[25] W. Borutzky, *Bond Graph Methodology*. Springer London, 2010.

[26] N. Deo, *Graph Theory with Applications to Engineering and Computer Science*. Dover Publications, 2017.

[27] S. Krantz, *Handbook of Complex Variables*. Birkhäuser Boston, 2012.

[28] C. Chicone, *Ordinary Differential Equations with Applications*. Texts in Applied Mathematics, Springer New York, 2006.

[29] Google Earth, *Randers, Denmark*. "56.4595314, 10.0375639", 05.10.2017.

[30] O. Nelles, *Nonlinear System Identification: From Classical Approaches to Neural Networks and Fuzzy Models*. Springer Berlin Heidelberg, 2013.

[31] M. Norgaard, O. Ravn, N. Poulsen, and L. Hansen, *Neural Networks for Modelling and Control of Dynamic Systems: A Practitioner's Handbook*. Advanced Textbooks in Control and Signal Processing, Springer London, 2003.

[32] T. Ando, S. Konishi, and S. Imoto, "Nonlinear regression model via regularized radial basis function networks," vol. 138, pp. 3616–3633, 03 2008.

Part III

Appendices

A. Elevation Profile from HZ to LZ

Figure A.1: Elevation profile along the High and Low Zones 1 [29].

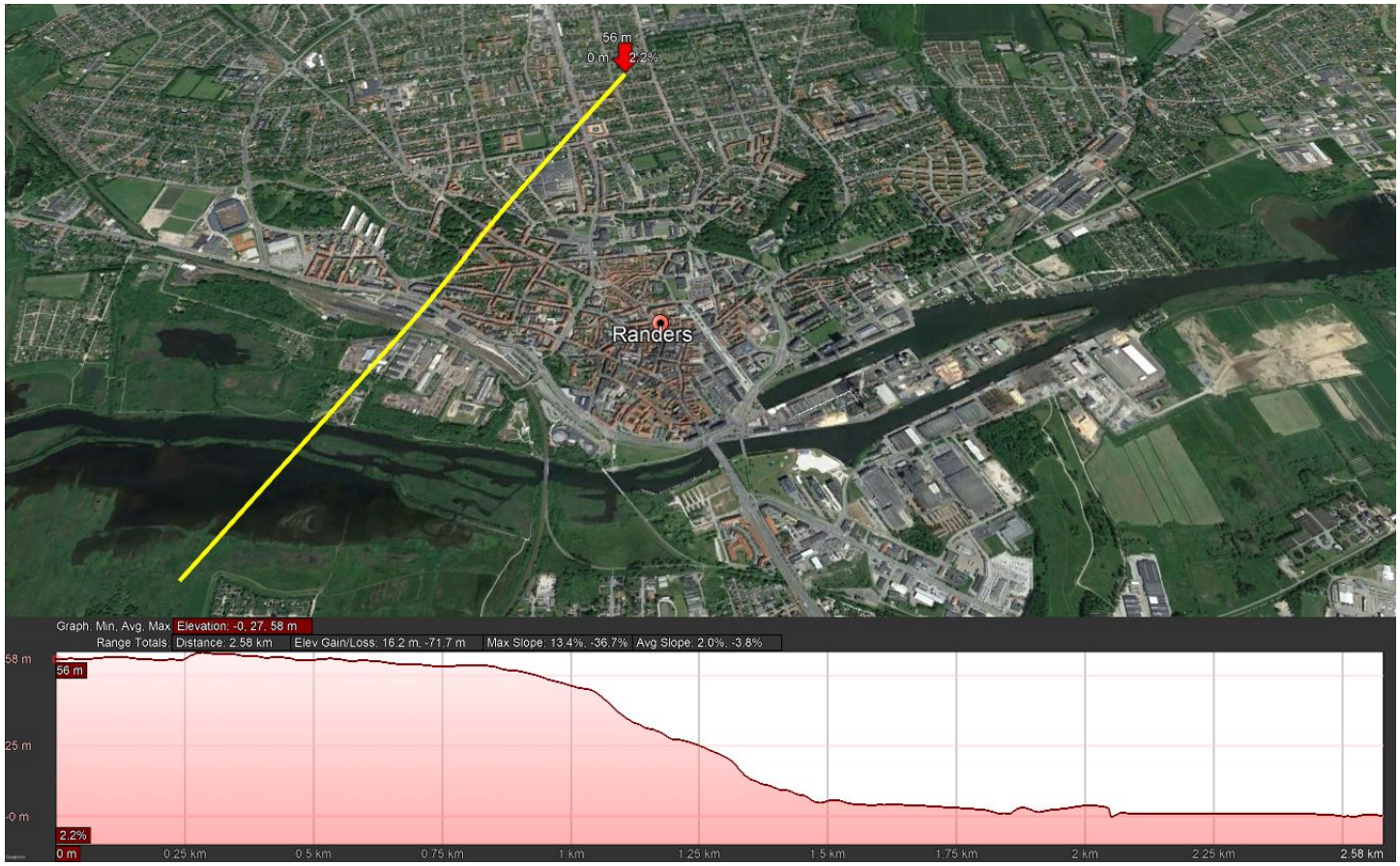


Figure A.2: Elevation profile along the High and Low Zones 2 [29].

B. Schematic drawing of the Randers WSS

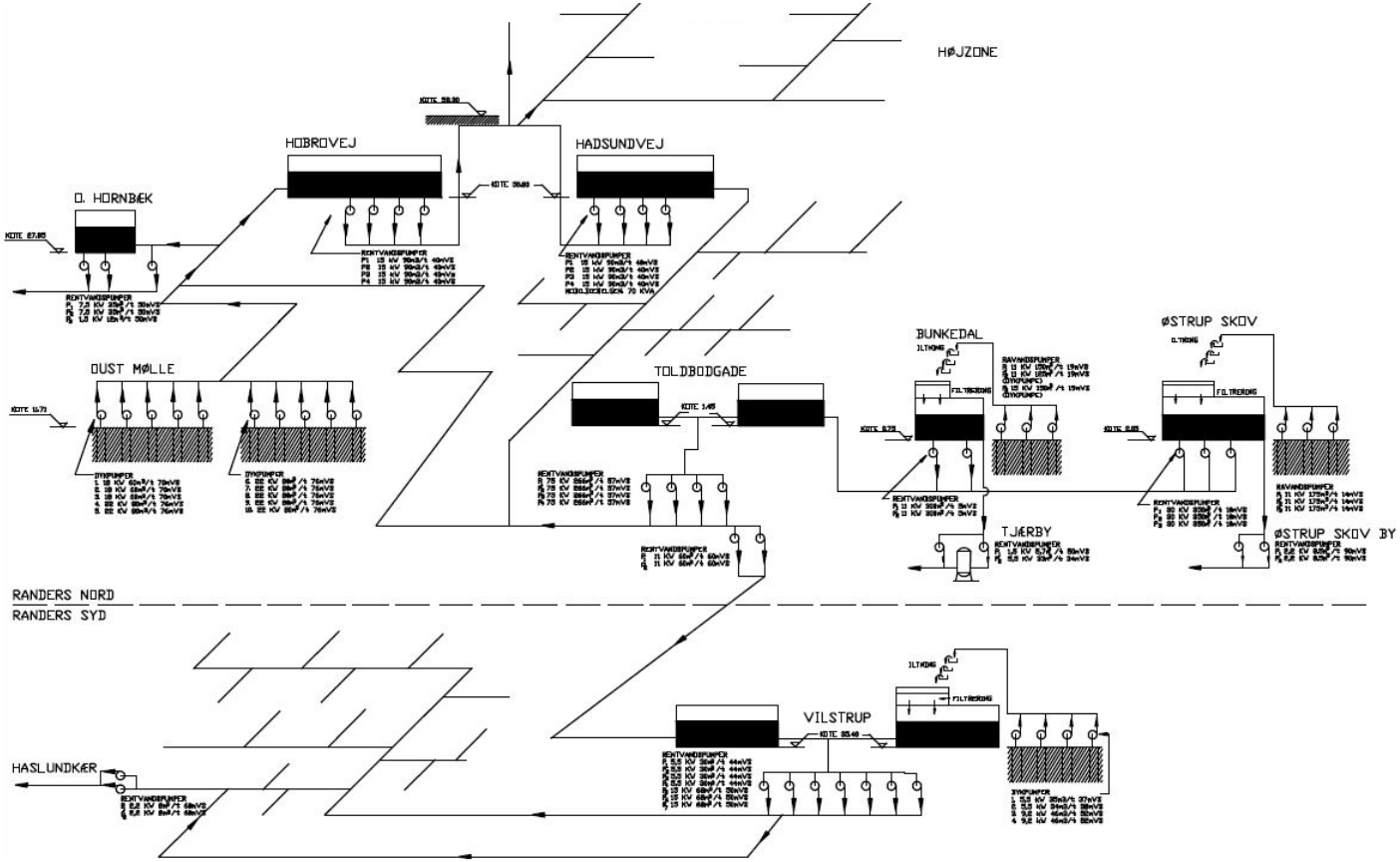


Figure B.1: Schematic drawing of the WSS in Randers [17].

C. Assumption List

In this part of the appendix, the different assumptions are collected, which were applied throughout the project. In order to ease the reading, each assumption has a section reference, which points to the place where it is used.

Number	**Assumptions**	**Section reference**
1	The fluid in the network is water.	Section *4.1.1: Hydraulic head*
2	All pipes in the system are filled up fully with water at all time.	Section *4.1.2: Pipe model*
3	The pipes have a cylindrical structure and the cross section, $A(x)$, is constant for every $x \in [0, L]$.	Section *4.1.2: Pipe model*
4	The flow of water is uniformly distributed along the cross sectional area of the pipe and the flow is turbulent.	Section *4.1.2: Pipe model*
5	At high flows, the Reynolds number is assumed to be constant. Therefore the Darcy friction factor, f_D is assumed to be constant.	Section *4.1.2: Pipe model*
6	Pumps in the network are of the type centrifugal.	Section *4.1.4: Pump model*
7	Tanks in the network have constant diameters. Equivalently, walls of the tanks are vertical.	Section *4.1.5: Elevated reservoir model*
8	The expressions which describe the pressure drop across components are continuously differentiable.	Section *4.4: Multi-inlet model*
9	The demand distribution parameter of the Multi-inlet, multi-WT model $v_{\mathcal{D}}$ is assumed to be constant. System identification is carried out with constant $v_{\mathcal{D}}$ parameters.	Section *5.1: Model structure of the Multi-inlet, Multi-WT system*
10	The identification of unknown-structured non-linear terms in the system model is carried out on linear-in-the-parameter models.	Section *5.1: Model structure of the Multi-inlet, Multi-WT system*

Table C.1: List of assumptions

D. Overview of system identification

In this part of the appendix, the basic steps of system identification are described. These principles are applied throughout the identification design.

D.1 Tasks in non-linear system identification

Modelling and identifying a non-linear system is a challenging task because non-linear processes do not share properties such as the superposition law in linear systems. As for any identification method, the goal is to find a model which is capable to represent the behaviour of a process as closely as possible. The quality of the model is typically measured in terms of the error between the output of the process and the model. In *Figure D.1* an illustration is shown for a system identification arrangement

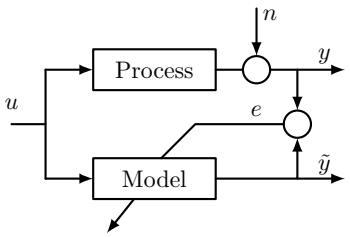


Figure D.1: Block diagram of system identification [30].

The process and model are fed with the same input signals, and their outputs are compared. This comparison gives an error signal e, which can be utilized for adapting the model. It should be noted, that in most cases the measurements on the output are disturbed by noise n. In order to carry out a succesful system identification, some major steps need to be performed, either by user interaction or using algorithms which can automatize these steps. *Figure D.2* shows the system identification loop with the major steps

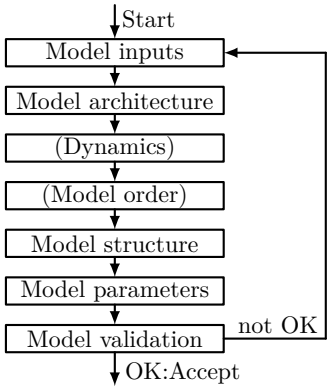


Figure D.2: System identification loop [30].

D.1.1 Choice of the model inputs

The first step in the identification is typically realized by a trial-and-error approach to find out the influence of the different inputs on the process with the help of prior knowledge. In complex systems, where the number of inputs are high and their influence is not so well defined, some data-driven input selection might be very helpful. In this case, using all inputs can lead to extremely high-dimensional approximation problems, which implies

the need for a huge number of parameters and increases the training time. Techniques for input data selection such as Principal Component Analysis(PCA) can be utilized in order to decide the relevance of certain inputs on the system. The main drawback of such techniques is that the relevance of an input is dependant only on the input data distribution, therefore sometimes highly relevant inputs are removed [30].

D.1.2 Choice of the model structure

The choice of the model architecture depends, among many factors on the type of the problem, the intended use, the insight into the real system behaviour, the complexity and the available data. The type of the problem can be for instance the approximation of a static system, or identification of the dynamics. In case of the WSS problem, both of them are considered. The intended use for the model architecture can differ whether the model is to be used for simulation, control or fault detection. Considering the insight, the complexity and the available data, three different modelling approaches can be distinguished[30]. These approaches are compared in *Table: D.1.*

	White box	**Gray box**	**Black box**
Information sources	First principles and insights.	Some insights and some data.	Data and experiments.
Features	Good understanding, high reliability and scalability.	⟷	Short development time and insight is not required.
Drawbacks	Well described process is required.	⟷	Not scalable and the accuracy is restricted by the available data.
Application	Planning, simulation and design for simple processes.	⟷	Data-driven model design.

Table D.1: Comparison of system identification modelling approaches[30].

As shown in *Table: D.1*, White box models are completely derived by first principles, i.e., by physical laws. The parameters and equations, describing the whole network can be determined by theoretical modelling, as it was done for our system in Chapter *4: System Modelling*. Black box models however, are based on measurement data, which means that the model describing the system is developed based on the characteristics of the data. In order to carry out a successful Black box modelling, very little insight or prior knowledge is required, however it is important to have a well-describing input-output data set. The combination between Black and White box models is called Gray box modelling. In this case, the knowledge from first principles and the information contained in the measurement data are both utilized. The blank fields in *Table: D.1* for Grey box models are left blank because the properties of such models lie between White and Black box modelling. Typically, when Grey box modelling is considered, the main goal is to overcome some of the most restrictive factors of the White and Black box approaches for the specific application. For example, some prior knowledge might be incorporated into a Black box model in order to ensure reasonable behaviour[30].

D.1.3 Model validation

The easiest type of validation is to check the quality of the model on the input-output data set. If this does not give satisfactory results, the model is not accepted. In this case, it can be either concluded that information is missing from the input or the model is not flexible enough to describe the corresponding input-output relations. In case if the performance achieved on the data set is acceptable, it is desirable to test the model on a new data set, especially if noise is present in the system. It should be noted however, that this new testing data should excite the system in the same operating regions, as the model was trained on. Otherwise, the model might fail the validation.

E. Overview of neural networks

In this part, a brief introduction is given to NNs, and more specifically, to RBFNNs.

E.1 Neural networks in general

From all possible realizations of functions with non-linear behaviour, almost all alternatives can be written in the following basis function formulation, if we assume that the number of basis functions is infinite [31]

$$\tilde{y} = \sum_{i=1}^{M} w_i^{(l)} \phi_i(u, w_i^{(nl)}). \tag{E.1}$$

The output $\tilde{y}$ is modelled as a weighted sum of M basis functions $\phi_i(\cdot)$.The basis functions are weighted with the linear parameters w_i, and they depend on the inputs and set of parameters gathered in $w_i^{(nl)}$. In order to realize a non-linear model, the basis functions need to be some kind of non-linear functions. The parameters of these non-linear functions can take part in the optimization or can be determined priorly. If the latter is the case, linear optimization can be used to find the optimal parameters $w_i^{(l)}$. Therefore, in the further discussion, *Equation: (E.1)* is considered in the form of

$$\tilde{y} = \sum_{i=1}^{M} w_i \phi_i(u). \tag{E.2}$$

The models in *Equation: (E.1)* and *Equation: (E.2)* are describing NNs. The illustration of an NNs is shown in *Figure E.1*.

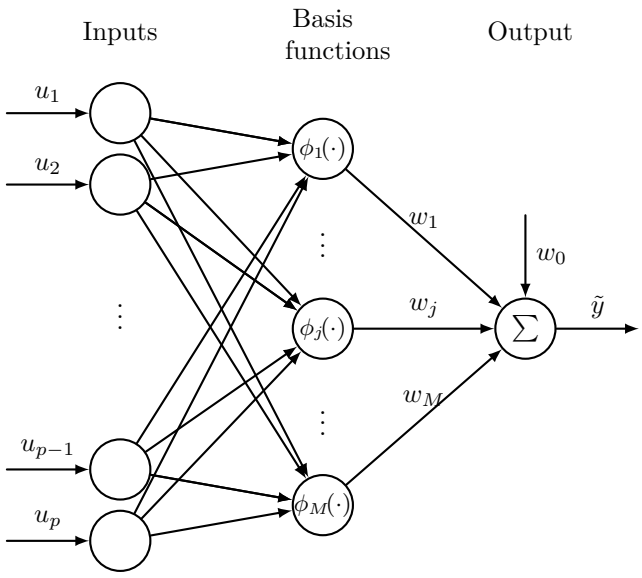


Figure E.1: Block diagram of a neural network with multiple inputs.

Typically, in literature such as [30, 31], a NN is distinguished from a non-NN network, when its basis functions are of the same type. In the terminology of NNs, the network in *Figure E.1* is described as follows. The node at the output is called the output neuron,

and all output neurons together form the output layer. In the example network, in *Figure E.1*, only one output is considered, therefore the output layer consists of one neuron. The set of M nodes in the center, each of which realizing a basis function, is called the hidden-layer. The inputs are associated with neurons and together they form the input layer. Furthermore, the linear parameters of the network associated with the output neuron(s) are called output weights. The output neuron is usually the linear combination of the basis functions in the hidden layer, with an additional possible offset w_0, called the bias. This offset parameter adjusts the operating point.

The basis functions in the NN formulation are generally multidimensional, defined by the number of inputs. For all NN approaches, however, the multi-variate basis functions are constructed by simple one dimensional functions[30]. This function is called the activation function. Such construction mechanism in the context of NNs is shown in *Figure E.2*

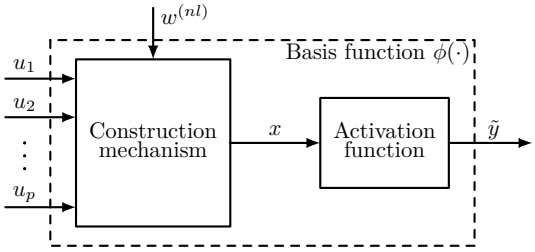


Figure E.2: Operation of construction mechanism [30].

The basis function of the neurons is based on the construction mechanism, that maps the inputs to a scalar x with the help of some non-linear parameters. The activation function then non-linearly transforms the scalar x to the neuron output $\tilde{y}$.

E.1.1 Radial construction

Among several construction mechanisms, the radial construction is further discussed. In this approach, the scalar x is calculated as the distance between the inputs and the center of the basis functions such that

$$x = ||u - \mu|| = \sqrt{(u-\mu)^T(u-\mu)}, \tag{E.3}$$

where $\mu = (\mu_1 \ \mu_2 \ ... \ \mu_M)^T$ is the center vector of the basis functions. The radial construction is utilized for RBF networks, which is discussed in the following sections.

E.2 RBF networks

In RBF networks, the first task is to calculate the Euclidean norm, i.e. the distance of the input and center vectors. This is the radial construction mechanism, which is shown in *Figure E.2*. In the second part, this distance x is transformed by the activation function. Therefore, an RBF network is a class of single hidden layer feedforward networks, expressed as the linear combination of radially symmetric non-linear basis functions [32]. Typically the choice for the basis functions is the Gaussian function, which is formulated in *Equation: (E.4)*

$$\phi(u, \mu_k, \psi_k) = exp\Big(-\frac{||u-\mu_k||^2}{2\psi_k^2}\Big), \tag{E.4}$$

where $\mu_k \in \mathbb{R}^M$ determines the center of the RBFs, $\psi_k \in \mathbb{R}^M$ is the standard deviation of the Gaussian, i.e the width parameter and $||\cdot||$ is the Euclidean norm. Depending on the choice of μ parameters, the RBFs can overlap each other to capture the information from the input data, while the width parameters ψ_k control the amount of these overlapping

basis functions. An example is shown in *Figure E.3*, where the influence of these parameters is illustrated on one RBF neuron, with a single input u.

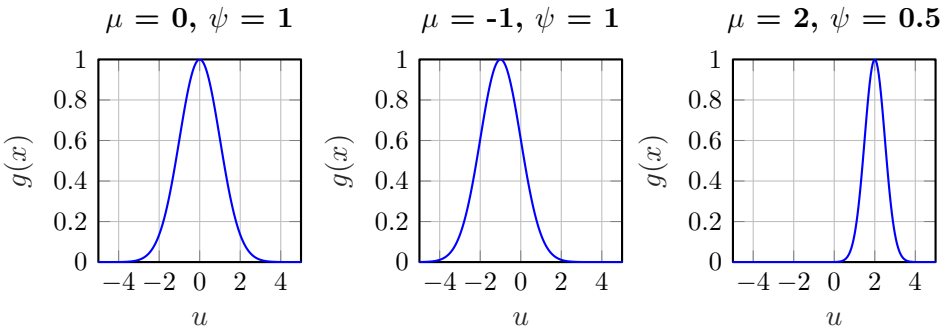


Figure E.3: The width and the position of the Gaussian basis function in terms of ψ and μ.

Therefore, when an RBF network is chosen for identification, three types of parameters need to be considered. The output weights w_i are linear parameters. They define the heights of the basis functions and the bias, i.e the offset value. The centers μ are non-linear parameters of the hidden layer neurons. They determine the positions of the basis functions. Furthermore, the standard deviations ψ are also non-linear parameters of the hidden-layer neurons. During the optimization, these parameters have to be somehow determined and optimized. *Figure E.4* illustrates the interpolation capability of RBF networks.

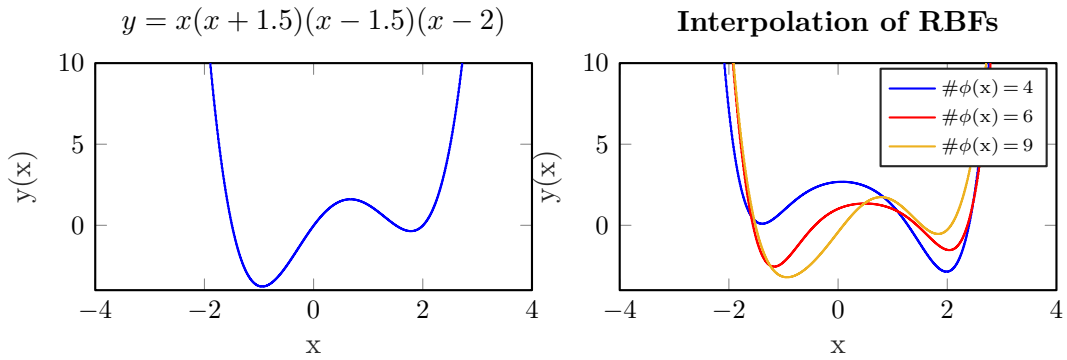


Figure E.4: Interpolation of an RBF network with different number of basis functions.

In *Figure E.4*, a 4^{th}-order polynomial is approximated with different number of Gaussian basis functions. As can be seen, with increased number of basis functions the estimation error decreases, and eventually a perfect interpolation can be achieved.

E.2.1 RBF training

The process of estimating the parameters of the basis functions ϕ_i consists of two problems. One task is to determine a set of M RBFs and then the other task is to estimate the optimal output weights w_i and bias parameters w_0. Therefore, the RBF network parameters are generally determined in a two-stage procedure, while in other type of networks the parameters are determined simultaneously by non-linear optimization[32]. As the output layer of RBF networks linearly combines the Gaussian basis functions, most strategies for RBF networks determine the parameters μ and ψ of the hidden layer, and then the output layer weighs w by using some linear optimization methods, such as LS methods [30]. This two-step procedure allows to vary only the weights between the hidden and output layer of the network during the training.

E.2.2 Determination of the hidden layer parameters

The parameters μ and ψ of the neurons are not determined through optimization. Instead, one approach is to determine these parameters only by using the input data set u. The most commonly applied approach for such tasks is the application of clustering techniques [30]. The clustering of the input data allows to determine the center μ of the basis functions, according to the distribution of the input data set. Thus, many RBFs are placed in regions where data is dense, and few RBFs are placed in regions where data is less dense.

The most commonly applied clustering technique in RBF networks is the k-means algorithm [30]. A cluster in the input data set can be defined as a group of data that are more similar to each other than data belonging to other clusters. Initially, the values for the M number of center points is chosen[1]. This can be done by choosing randomly M different data samples. Then the algorithm assigns all data samples to their nearest cluster center. When the M clusters are separated, a new mean is calculated for each cluster. Each cluster center is set to the mean of its cluster such that

$$c_k = \frac{1}{N_k} \sum_{i \in \mathcal{A}_k} u_i, \tag{E.5}$$

where i runs over the N_k data samples that belong to cluster k, which are in the set $\mathcal{A}_k$ and N_k is the number of elements in the set $\mathcal{A}_k$. Eventually, the algorithm stops when the center values c_k do not change. The c_k center parameters of the clusters then define the centers μ_k of the basis functions. Therefore, the k-means algorithm determines the centers μ, in an unsupervised manner, meaning that the RBFs are not automatically moving to the regions where they are required for a good approximation of the system. However, if the number of RBF neurons is sufficient, the NN can give a good approximation of the process. Compared to the supervised learning techniques, where the basis function parameters are to be optimized, the increased number of neurons does not cause increased training time, since both the k-means and LS algorithms are very fast [30].

After the clustering is completed, and the c_k cluster centers are assigned to the μ basis function centers, the width parameters ψ can be determined by the k-nearest rule. This method assigns each RBF a width parameter proportional to the variance of the values in the corresponding cluster such that

$$\psi^2 = \frac{1}{N_k} \sum_{i \in \mathcal{A}_k} ||u_i - \mu_k||^2. \tag{E.6}$$

With *Equation: (E.5)* and *Equation: (E.6)*, the two parameters of the basis functions are defined.

[1] The M number of center points of the clusters correspond to the number of basis functions, thus the number of neurons.

F. RBFNN-based identification on an example network

In Appendix F. a numerical example for the system identification is given.

F.1 The network in EPANET

The reason for testing the identification method on data from a simple EPANET model is that, in simulation, several different operations of the network can be studied. On a real world network, typically the operation or control cannot be taken to its extremes, as this would mean undesired service to the customers . Therefore, the network shown in *Figure F.1* has been created to mimic the behaviour of a multi-inlet, single-WT system, such that the identification method, presented in Chapter *5: System identification* is being tested.

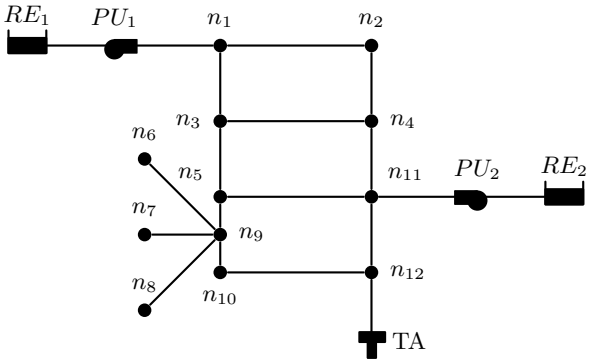


Figure F.1: Multi-inlet, Single-WT example network.

The network shown in *Figure F.1* consists of two pumping stations PU_1 and PU_2 and a Water Tank TA. Furthermore, reservoirs are present in the network, as in EPANET reservoirs are required for simulating pumps. These reservoirs RE_1 and RE_2 are different from WTs in the sense that they cannot be emptied in the simulation. Additionally, the physical and operational properties of the network can be found in *Appendix: G.*

The controls implemented on the network are based on simple time scheduling of the two pumping stations. Both pumps turn off when the pressure head in the WT is above $19.85[m]$. PU_2 turns on again if the pressure head in the WT decreases to $19.55[m]$. Furthermore, PU_2 turns on again only if the pressure head in node n_1 exceeds $43[m]$. The initial pressure head in the WT is $19.5[m]$.

F.2 Identification and validation

The identification is carried out on the simulation data, exported from EPANET. In order to validate the identified model, scenarios with increased and decreased consumption patterns are taken intro account. It is worth noting, that the control properties of the network are kept the same, as any modification in control would mean a new operating point in the WSS. Thus, the model is tested by modifying the total demand and therefore creating different inlet pressure and flow settings for the two pumping stations. The identification is carried out on the network when the total consumption is σ_1. The validation is done when the total consumption is increased(σ_2) and when decreased(σ_3), respectively. Furthermore, all three consumption curves are periodic with a time period

of 24 hours, however the validation has been done on a 72 hours long test data. *Figure F.2* shows the three different consumption patterns.

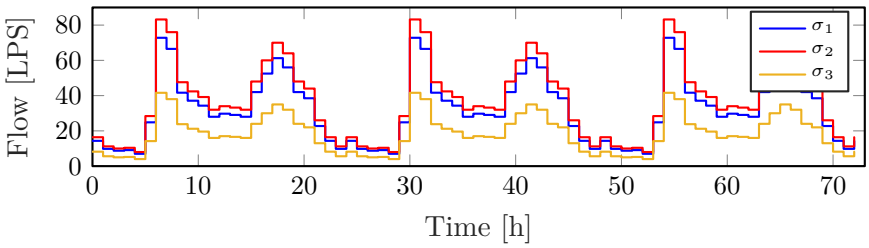


Figure F.2: Consumption patterns used for identification and validation.

The inlet flows of the two pumping stations PU_1 and PU_2, and the WT pressure $\hat{p}$ are shown in *Figure F.3*.

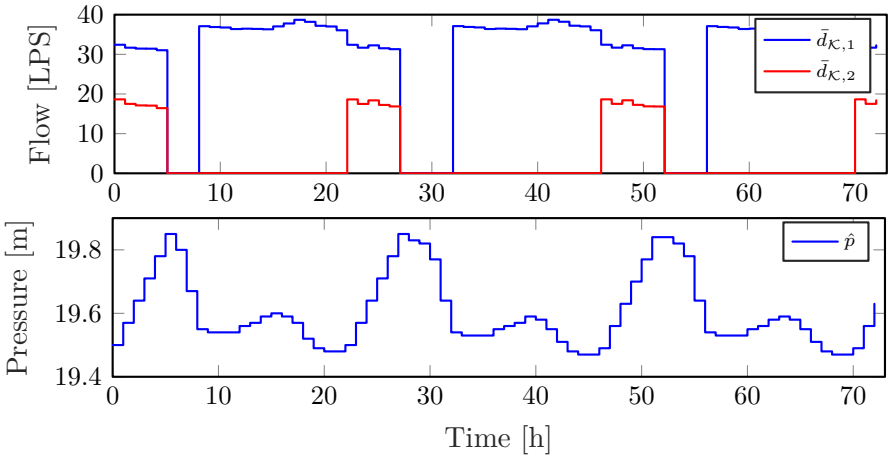


Figure F.3: Inlet flows of PU_1 and PU_2(above) and the pressure in the WT(below) under σ_1 total demands.

F.2.1 Identification on σ_1 total demands

For the two inlet pressures $\bar{p}_{\mathcal{K}1}$ and $\bar{p}_{\mathcal{K}2}$, the identification has been carried out with ten RBF neurons. The performance goal has been chosen to the order of 10^{-4}. *Figure F.4* shows the mean of squared errors in terms of RBF neurons.

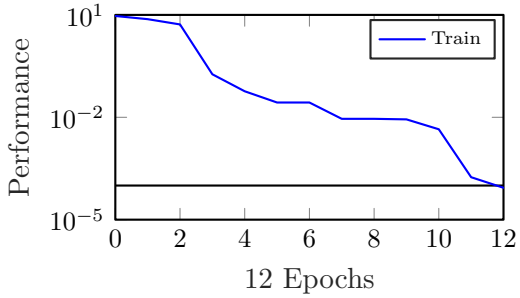


Figure F.4: The network's performance according to the mean of squared errors.

As it is shown in *Figure F.4*, a very accurate fit is achieved for the model which is capable of generalizing the output equation. It is worth to mention that although more RBF neurons results in a more accurate fit, however it does not generalize the model as well as ten RBF neurons. In case of increasing the number of neurons, the identification resulted

in overfit and the validation failed. The identification of $\bar{p}_{\mathcal{K}1}$ and $\bar{p}_{\mathcal{K}2}$ inlet pressures in case of σ_1 total demand is shown in *Figure F.5*.

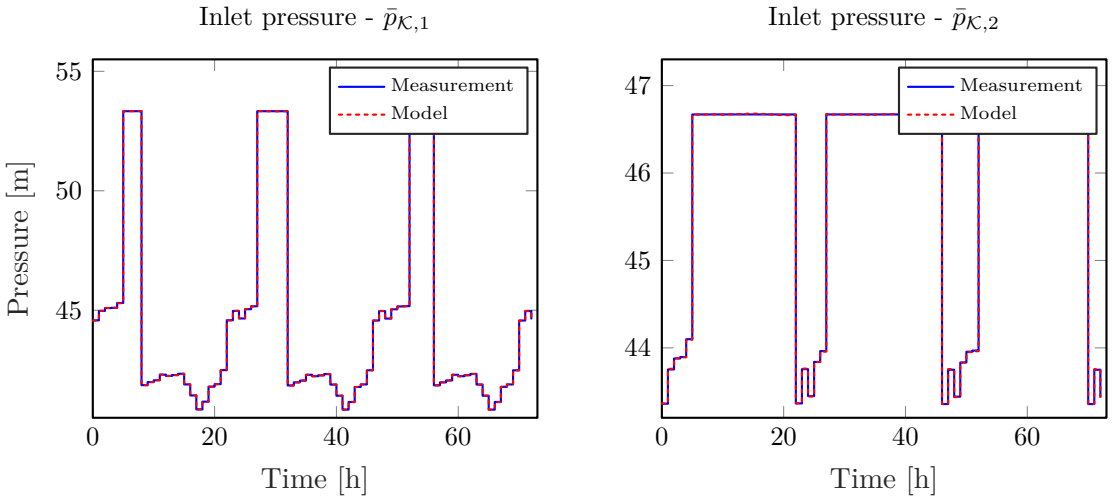


Figure F.5: Identification of the inlet pressures $\bar{p}_{\mathcal{K}1}$ and $\bar{p}_{\mathcal{K}2}$.

As can be seen in *Figure F.5*, the inlet pressures $\bar{p}_{\mathcal{K}1}$ and $\bar{p}_{\mathcal{K}2}$ of the identified model and the measurements are almost identical.

For the WT pressure $\hat{p}$, the identification has been carried out with six RBF neurons. The identification on the 72 hours-long data did not result in good fit and generalization, therefore the measurement data from the EPANET simulation has been extended to a 140 hours-long data set. *Figure F.6* shows the mean squared errors in terms of RBF neurons on the identified network model.

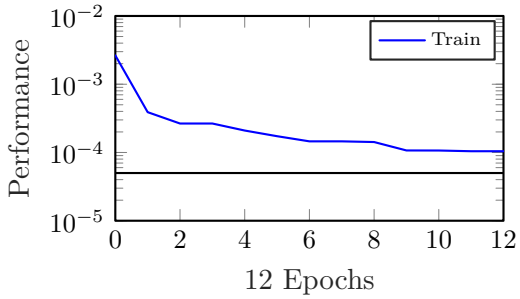


Figure F.6: The network's performance according to the mean of squared errors.

Figure F.6 shows, that adding more neurons to the network, the performance does not increase significantly. The network has been tested with more than six neurons, however the best validation result turned out to be the best at $N = 6$ RBF neurons. The WT pressure is shown in *Figure F.7*

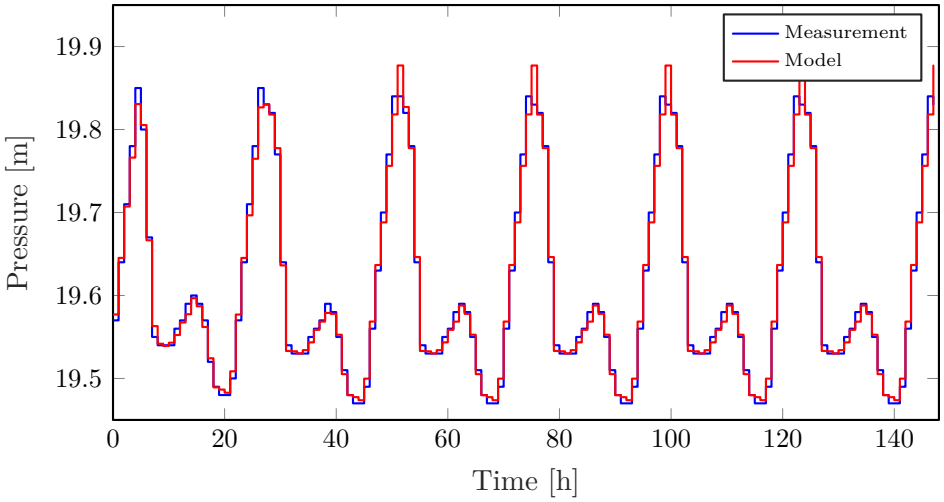


Figure F.7: Identification of the WT pressure $\hat{p}$.

The output parameter vectors $\theta_{\mathcal{K}1}$, $\theta_{\mathcal{K}2}$ and the state parameter vector $\theta_{\mathcal{S}1}$, including the output weights, linear parameters and biases are shown in *Equation: (F.1)*.

$$\theta_{\mathcal{K}1} = \begin{pmatrix} -321.27 \\ -101.81 \\ -391.47 \\ -382.66 \\ -4.48 \cdot 10^6 \\ 4.55 \cdot 10^6 \\ 9.23 \cdot 10^3 \\ 4.48 \cdot 10^6 \\ 394.62 \\ -4.57 \cdot 10^6 \\ -0.27 \\ 37.33 \end{pmatrix}, \quad \theta_{\mathcal{K}2} = \begin{pmatrix} -102.36 \\ 10.88 \\ -89.94 \\ 115.19 \\ -3.28 \cdot 10^5 \\ 1.42 \cdot 10^5 \\ -1.78 \cdot 10^3 \\ 3.27 \cdot 10^5 \\ 35.35 \\ -1.39 \cdot 10^5 \\ -0.33 \\ 60.14 \end{pmatrix} \quad \theta_{\mathcal{S}1} = \begin{pmatrix} 4.15 \cdot 10^3 \\ 2.64 \cdot 10^4 \\ -5.15 \cdot 10^4 \\ 2.51 \cdot 10^4 \\ 0.42 \\ -4.15 \cdot 10^3 \\ 3.4 \cdot 10^{-3} \\ -8.08 \cdot 10^{-4} \\ -5.8 \cdot 10^{-3} \\ 0.16 \end{pmatrix} \tag{F.1}$$

F.2.2 Validation on σ_2 total demands

With the increase in total water consumption, the controls change. The new inlet flows of the two pumping stations PU_1, PU_2 and the pressure in the WT is shown in *Figure F.8*.

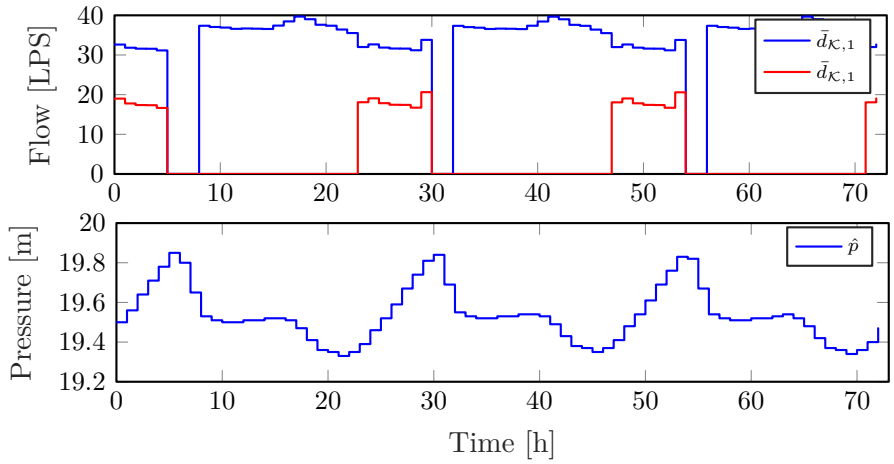


Figure F.8: Inlet flows of PU_1 and PU_2(above) and the pressure in the WT(below).

The simulation results for $\bar{p}_{\mathcal{K}1}$ and $\bar{p}_{\mathcal{K}2}$ inlet pressures on the new validation data when σ_2 is the total demand in the network are shown in *Figure F.9*

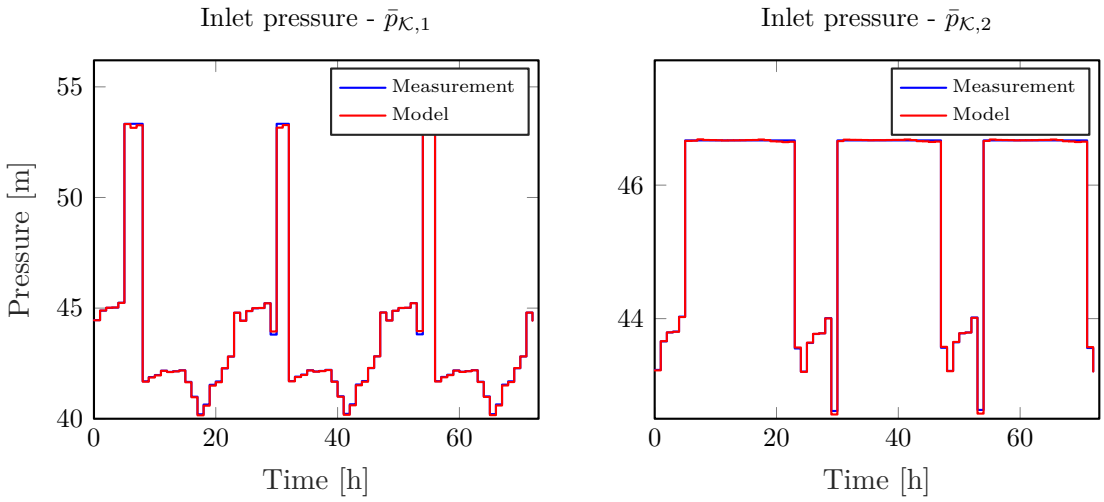


Figure F.9: Simulation of $\bar{p}_{\mathcal{K}1}$ and $\bar{p}_{\mathcal{K}2}$ pressures on validation data.

The simulation results for $\hat{p}$ WT pressure on the new validation data when σ_2 is the total demand in the network are shown in *Figure F.10*

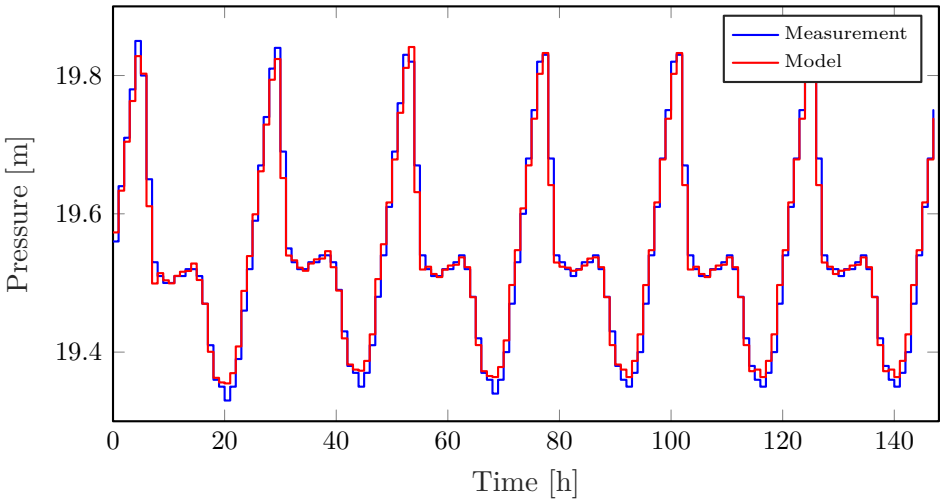


Figure F.10: Identification of the WT pressure $\hat{p}$.

F.2.3 Validation on σ_3 total demands

(in progress)

G. Example network description

In this part of the appendix, the corresponding physical parameters, control properties and consumption properties of the Multi-inlet, Single-WT example network are listed.

H. Measurements