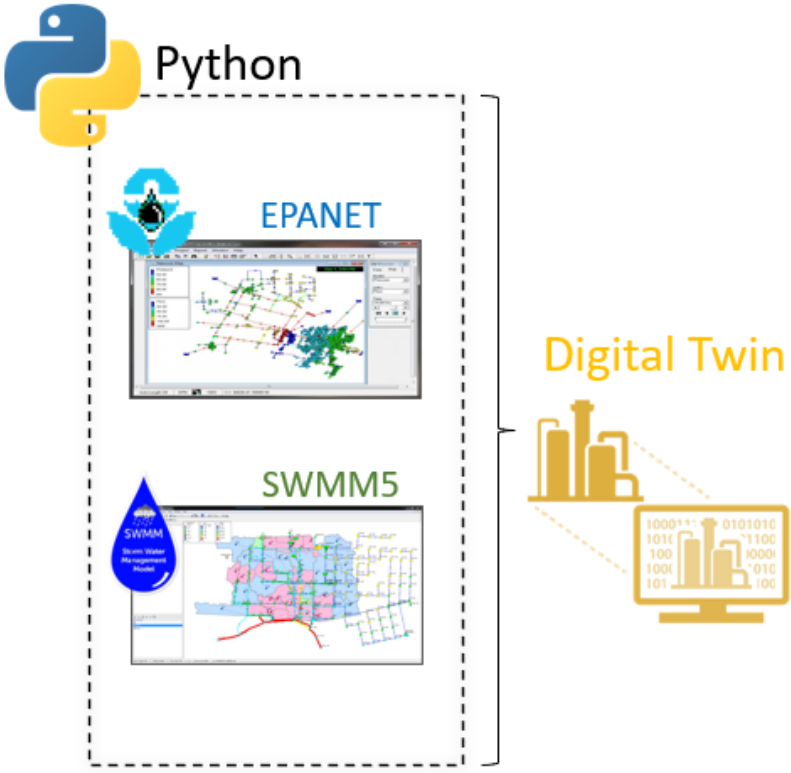
CHALMERS UNIVERSITY OF TECHNOLOGY | NTNU Norwegian University of Science and Technology



On the interconnectivity of urban water system models

Possibilities, limitations, and feasibilities

Nordic 5 tech master's thesis in Environmental Engineering track Urban Water

MAGDALENA JAURENA BELTRAMI

DEPARTMENT OF INFRASTRUCTURE AND ENVIRONMENTAL ENGINEERING

Division of Water Environment Technology

CHALMERS UNIVERSITY OF TECHNOLOGY

Gothenburg, Sweden 2023

www.chalmers.se

Abstract

The development of models in the urban water systems together with the increasing availability of real time data is allowing decision makers to evaluate different scenarios and take informed decisions. However, real time data is not always available and easy to get. The increasing use of smart water meters has made it possible to get data for the water distribution systems but for sewer systems it is not possible to have a good representation of the whole system due to the high cost and intensive maintenance of the sensors needed.

This thesis defines the interactions between the water distribution and the sewer system, to further suggest methods in which they can be modelled. The idea behind the thesis is to propose methods to reproduce the sewer flows in dry periods based on the demands and leaks from the water distribution network. The aim was to set the basis on how the interactions could be modelled and define the parameters that would need to be adjusted if more accuracy is required.

Acknowledgments

This project is the final documentation of my Master Thesis in order to achieve the Nordic´s master's degree in environmental engineering at both Norwegian University of Science and Technology (NTNU) and Chalmers University. The topic was suggested by Marius Møller Rokstad and Franz Tscheikner-Gratl from NTNU who together with Thomas Pettersson from Chalmers supervised my work.

I would like to express my gratitude to Marius and Franz, who despite the physical distance, they provided unwavering support and guided me every step of the way. Their invaluable assistance helped me untangle complex issues and navigate through challenges that arose during my project. Additionally, I extend my sincere appreciation to Thomas, who consistently championed my best interests throughout this time.

This thesis is dedicated to my parents, my sister and my aunt who always encouraged me to follow my own path and who I terribly miss. Also, to Bru who had to endure all my ups and downs during this last two years of master and still stands by my side.

Finally, I would like to thank all the people I met during my master and that will remain forever in my heart. Specially to my Nordic family -Lu, Sofi and Summy- which always made me feel at home.

Magdalena Jaurena Beltrami

01.06.2023

Gothenburg, Sweden

Contents

1 Introduction

Urban infrastructure has a dynamic nature, mainly due to changes in population and ageing systems. Together with this, high investments costs, and an increasing trend towards sustainability has required us to find new ways to address several problems of management and operation of urban water systems (Shi et al., 2019). The constant development of models in urban water systems and the availability of real-time data allow decision makers to evaluate different scenarios to take informed decisions (Li et al., 2020). With real-time data, water utilities can visualize the performance of their networks. Regarding planning, models can be used for asset management, development of master plans, contingency plans, and resiliency management (Romano & Kapelan, 2014). With a representative model and real data, it is possible to reproduce the behaviour of the network in long term demand forecast, design replacement plans and simulate emergency scenarios (Romano & Kapelan, 2014). Furthermore, concerning operation and maintenance, models provide a tool for decision support for applications such as getting a deep knowledge of the network, predicting the behaviour of the system, planning works in the network, optimize the network performance by preserving the quality and quantity of the water, minimizing the energy consumption, minimizing leaks, among others (Li et al., 2020).

When it comes to collecting data to feed the models, the vast placement of sensors and the increasing use of smart water meters has made it possible to get data for the water distribution systems (Zhang et al., 2021). Conversely, for sewer systems it is not possible to have a good representation of the whole system due to the high cost and intensive maintenance of the sensors needed. Therefore, generally it is possible to monitor the sewer network only in sparse locations. This implies that it is not currently realistic to get real-time data using only in-sewer observations (Lund et al., 2021). However, it is possible to establish a link between the water distribution and urban drainage system, as most of the water consumed will eventually end up in the sewer system (Lund et al., 2021). Several studies such as the one proposed by Zhang et al. (2021) and Lund et al. (2021) have managed to estimate the household demands in sewer based on data obtained from the water distribution system. Having a real-time sewer hydraulic model that simulates in real time the water depths and sewer flow across the whole network, allows to address several issues associated with sewer system such as overflows from CSO (Zhang et al., 2021). Together with in-sewer observations, it makes possible to detect and localize in-leaks and exfiltration, illicit discharges, deposits, and illicit connections to the sewer network [Lund et al. (2021) and Zhang et al. (2021)].

As cities continue to grow, new infrastructure must be built in order to provide service to the new areas. Regarding sewer systems, separate systems are the preferred solution as they disconnect the wastewater from stormwater. They come as a solution for reducing the combined sewer overflows which directly discharge unpolluted wastewater to the receiving water bodies causing environmental problems (Langeveld et al., 2012).

However, it is well known that large volumes of undesired water end up in the separate sewer system affecting its performance (Beheshti et al., 2015). The infiltration and inflow (I/I-water) of unwanted water can be caused due to faulty connections of stormwater outlets to the foul sewer system, infiltration from groundwater, drinking water leakages and other

extraneous sources (Beheshti & Sægrov, 2018a). I/I-water can represent high percentages of the volume that reaches the wastewater treatment plant. In Norway, Sola et al. (2018) estimated that for the studied plants the I/I-water in 2016 represented 66%, while for Sweden, Svenskt Vatten calculated an average of 49% in 2012.

The presence of I/I-water can cause several environmental, social, and economic impacts, and is a critical threat to sustainable urban water management and water infrastructural asset management (Beheshti et al., 2015). I/I-water represent a general reduction in the efficiency of sewer systems and treatment facilities (Beheshti & Sægrov, 2018a). They can cause overloading of the sewer network resulting in sanitary sewer overflows that can pose a threat to public health, as well as rising both the chemical an energy consumption in treatment plants (Beheshti & Sægrov, 2018a). The detection and location of I/I-water and illicit connections plays a major role for the efficient reduction and removal of I/I-water. Several qualitative methods can be used for the detection of I/I-water such as smoke testing, dye testing, DTS method, among others (Beheshti et al., 2015). However, they imply some physical method such as the addition of smoke, dyes or measuring temperature.

This master thesis aims to give a step forward into the development of models in the urban water systems by studying the interactions between the water distribution system and the sewer system. Furthermore, due to the relevance of in-leaks to the sewer system in countries such as Norway and Sweden, this thesis will study the possibility of relating out leaks from WDS to sewer systems, in order to address the I/I-water issue.

2 Aim and research questions

The aim of this Master Thesis is to develop a model-based proof of concept of the interactions between consumers water demand, out-leaks from drinking water network and in-leaks to the sewer network. This is done by setting a partly artificial hydraulic model for drinking water, to simulate leakages, and a model for sewer systems, to model the in-leaks. Through the analysis of different scenarios, the aim is to determine if it is technically possible and plausible a relation between the water and sewage network model. The water distribution model will be done using EPANET and the sewer network will be modeled in SWMM.

The following research questions are formulated:

> How is it plausible to build an interaction between the drinking water distribution model and the sewer model? In particular, how can both the household discharges and in leaks into the sewer system be modeled from the water demands? Which is the information needed to get a nearly realistic model?

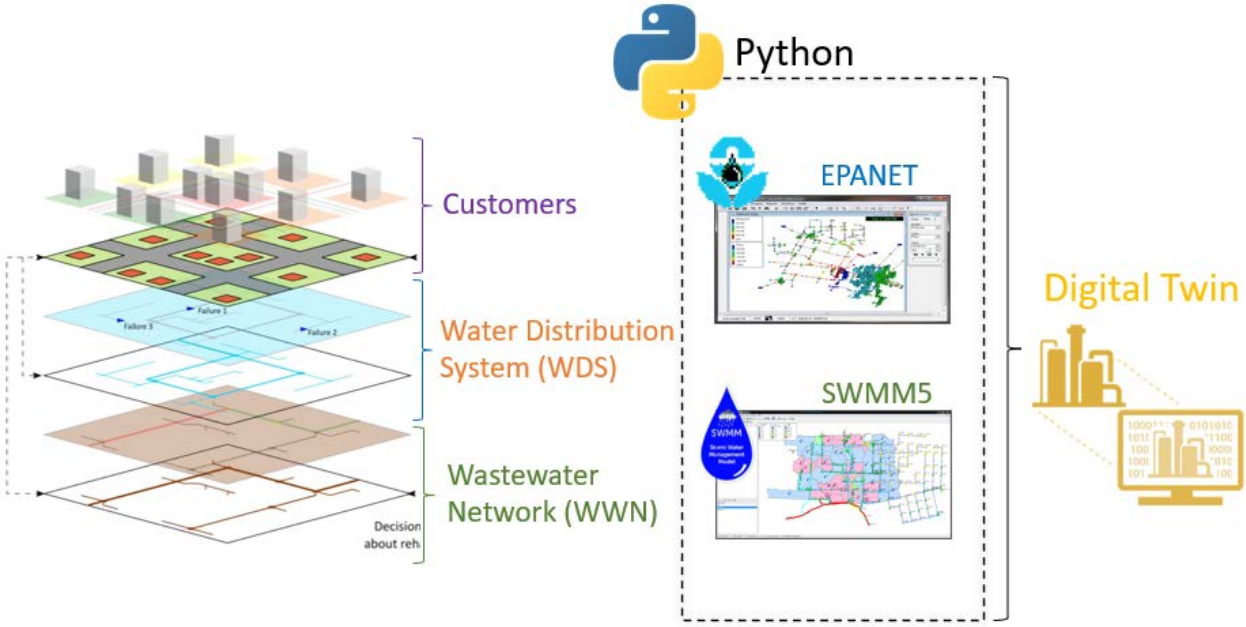


Figure 2-1: Visual aid for understanding master thesis aim. Provided by Marius Møller Rokstad and Franz Tscheikner-Gratl (NTNU) and modified.

3 Conceptual model

The conceptual model developed to identify and model the relations between water distribution system (WDS) and sewer network (SN) is presented in Figure 3-1.

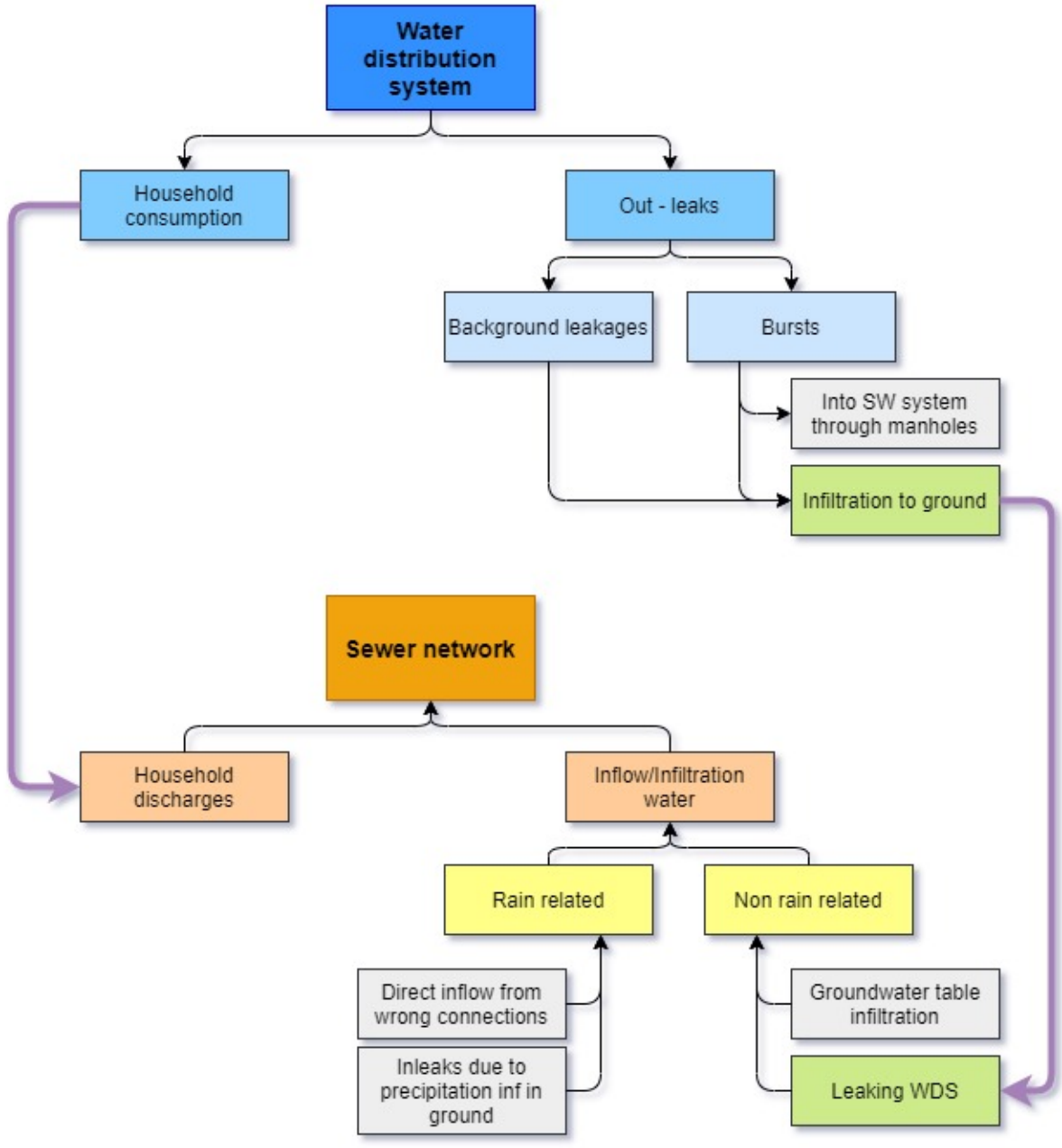


Figure 3-1: Conceptual model of interconnection between WDS and SN.

The demand required in the water distribution system is comprised by mainly the demands from users and possible leaks occurring in the network (Covelli et al., 2015). Leaks are a recurring problem for water utilities. In Norway it is estimated that around 30% of the water produced is leaked (Norwegian Water, 2018). Leaks from the water distribution system can be categorized in two ways according to their occurrence in the system (Berardi et al., 2015). Background leakages are almost constant leaks, through valves or joints (Covelli et al., 2015).

They are difficult to identify as the leaked flow is low, slowly emerging, and generally distributed in the pipe's length. They do not generate a clearly evident pressure drop in the network distribution; therefore, they remain unreported for long times, causing large volumes of water loss (Berardi et al., 2015). On the other side, pipe bursts refer to those leaks that occur because of a sudden failure in the pipe and they are associated with localized water losses (Covelli et al., 2015). In this case, the leak flow is high as well as the volume lost, due to the pipes being pressurized. However, they are usually repaired quickly due to them being reported by users as they are prone to causing major service disruptions (Berardi et al., 2015).

The water leaked will infiltrate into the ground nearby. In case of the pipe bursts, it can also be led to the drainage system if the leaked volume is large, however this scenario is not analyzed as the focus is on in leaks into sewer. Once infiltrated, it can become part of the groundwater or enter the sewer system through defects in the pipes.

With regard to the drinking water household demands, they represent the consumers demand over time. The consumption usually follows a daily pattern, being low in the night and having a peak in the morning before consumers go to work/school and in the afternoon. They also show seasonal variations, according to the characteristics of the area. Drinking water consumption can be intrinsically linked with sewer as a great percentage of the water consumed will end up in the sewer system (Lund et al., 2021). The household consumptions from drinking water can be used therefore to estimate the household discharges into the sewer system.

As seen in Figure 3-1, sewer flows can be characterized as household demands an infiltration/inflow water. The infiltration/inflow water precedence will depend on whether it is a rainy period or not (Lundblad & Backö, 2014). In rainy periods, wrong connections of drainage into sewer pipes in a separate system contribute to the I/I water. Furthermore, precipitation will infiltrate onto the ground and can enter the sewer through cracks and holes (Lundblad & Backö, 2014). On the other hand, in dry periods, infiltration can be caused by elevated groundwater table surrounding the pipe, independent of the hydrological cycle (Kragset, 2019). Furthermore, leaks from the water distribution system also contribute to infiltration into sewer systems (Beheshti & Sægrov, 2018a).

Looking at the diagram in Figure 3-1, two links can be established between the water distribution network and the sewer system. The household demands are related to the sewer household discharges, and the leaks from the water distribution system can end up in the sewer system.

When it comes to modelling, both the water and sewer networks can be well represented using EPANET and SWMM, respectively. Both software's were developed by the United States Environmental Protection Agency's (EPA) and are used worldwide to model the networks and are considered state of the art. Some studies (Lund et al., 2021; Zhang et al., 2021) have attempted to describe the link between household demands in both networks, but additional investigations need to be done in order to describe the leaks association. In the next chapters, the results of research and further modeling of the connections are presented. The idea was to build up the sewer flows from the water distribution model by representing its components in dry weather: household demands and infiltrations from sewer.

4 Case study

This study will be developed based on the water distribution network and sewer system from part of the Risvollan area, located in the south of Trondheim, Norway (see Figure 4-1). According to Beheshti & Sægrov, (2018a), Trondheim has both combined and separate systems, where the latter accounts for about 52% of the total network length. From a water balance in the city´s wastewater system in the period 2009-2011, it was detected that 46% of the total water delivered in dry period to the wastewater treatment plant came from I/I-water (Beheshti & Sægrov, 2018b). Trondheim municipality has declared solving this problem as high priority, as the percentage can increase further during wet weather (Beheshti & Sægrov, 2019).

Figure 4-1: Case study location. Taken from Google Earth.

The Risvollan catchment has an area of about 20 ha, and a population of 1112 inhabitants (Kragset, 2019). The buildings in the area are mainly residential but there is also a kindergarten. The altitude changes along the area, with presence of small hills and varying topography. According to the Geological Survey of Norway (2018) the soil consists of thick oceanic deposits, which is considered unsuitable for infiltration purposes. The catchment has a good database and there is an urban-hydrological measuring station at the outlet of the catchment, that allows to study hydrological parameters throughout the year (Beheshti et al., 2015).

Several studies (such as Kragset, (2019) and Beheshti et al. (2015)) have been conducted in this area, which have managed to construct models for both water distribution and sewer network. These models are used in this study and further developed to accomplish the aims of this thesis. In the following sections the given models are presented, providing an inside on the information they contain.

4.1 Water distribution network model

The water distribution network is modeled in EPANET and was provided by NTNU supervisors. It must be stated that this is not the real model for Risvollan water distribution network, as it could not be distributed for information security reasons. However, the model provided still represents the actual demands in the area and are placed within the area where the demand is generated.

The model representation is shown in Figure 4-2 together with the diameter and elevation distribution. It consists of 7.0 km of pipelines, represented by a total of 108 links. The diameters range from 51 to 250 mm (see Figure 4-3). The material of the pipes consists of Cast Iron, polyvinylchloride (PVC) and polyethylene (PE).

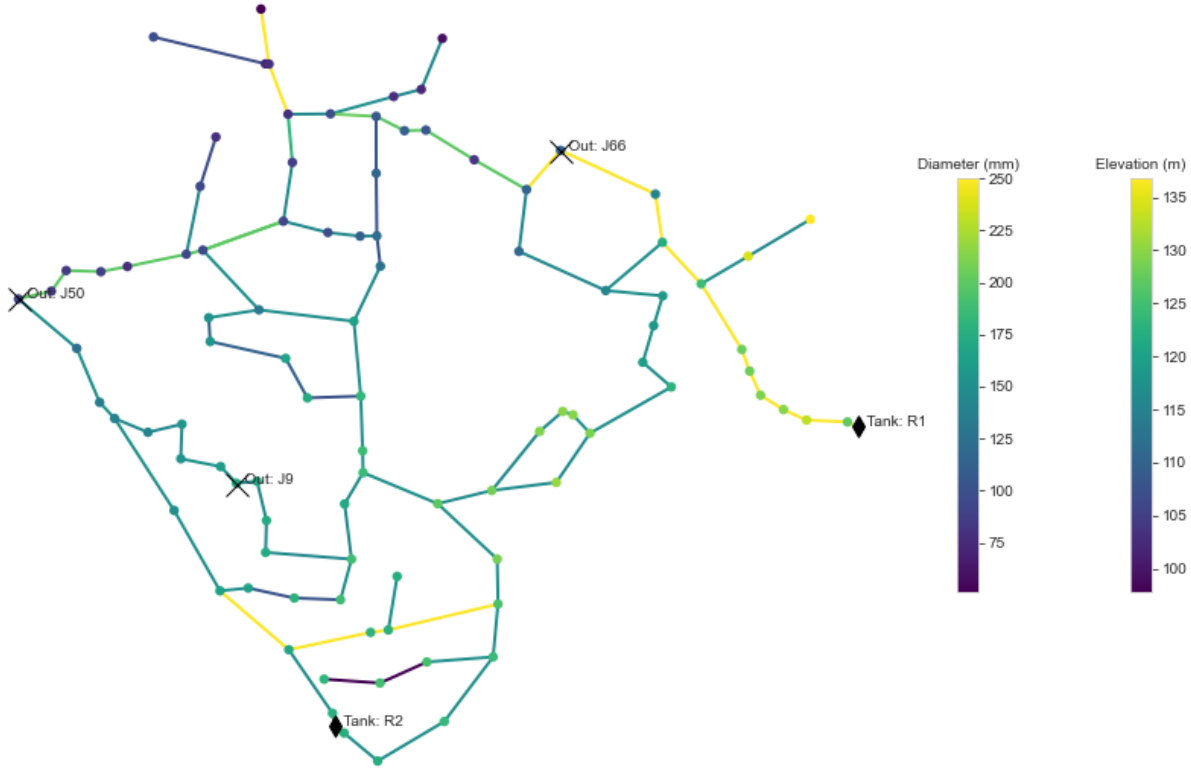


Figure 4-2: Water distribution network model.

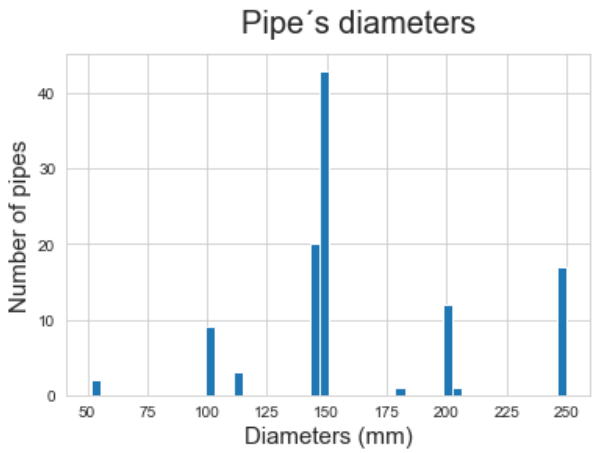


Figure 4-3: Distribution of pipe´s diameters, water distribution network model.

The household demands are given as two different daily patterns which are assigned to the nodes depending on the type of demand they are assumed to follow ("Household_1" and "Household_2". Moreover, leaks are represented as a constant demand pattern ("Background_leak"). Finally, as the study intends to be limited to Risvollan area, three other patterns were added which represent the water supply from the modeled pressure zone into other areas of the drinking water system ("External_demand"). These patterns are placed in the nodes represented as "External_demand" in Figure 4-2 and is where the boundary for the model is. The patterns are illustrated in Figure 4-4 where the pattern multipliers are displayed for each hour. Each node is assigned with a corresponding "base demand" that will multiply the patterns defined for the node in order to get the final demand.

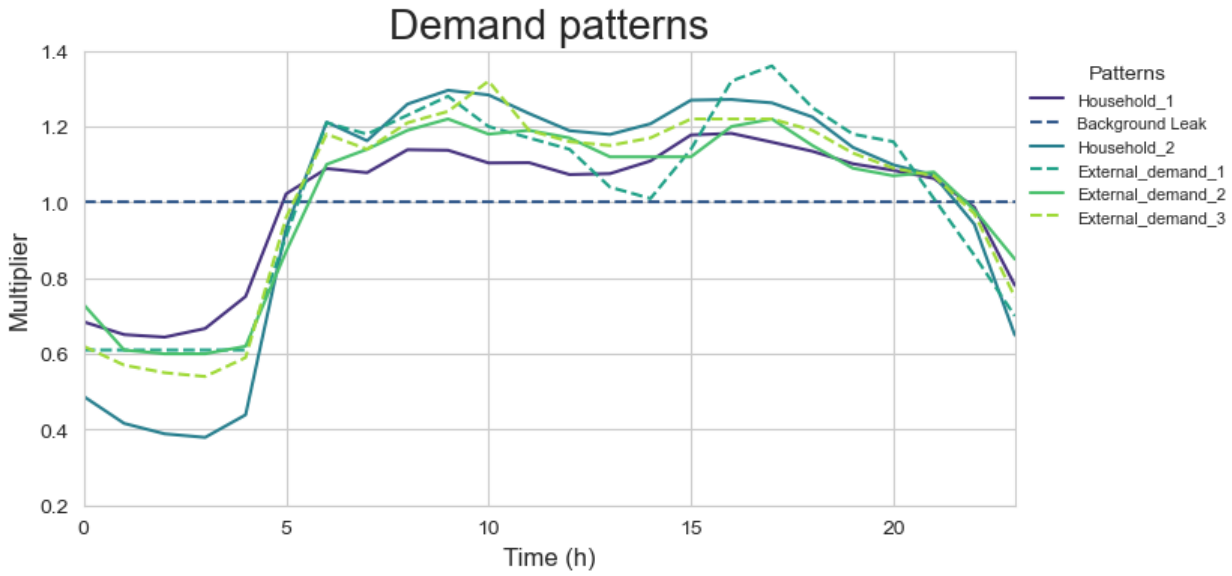


Figure 4-4: Demand patterns in the water distribution model.

4.2 Sewer model

The sewer model provided was first developed by Birgitte Taugbøl Kragset (Kragset, 2019). It is based on physical data of the pipe system provided by Trondheim municipality (see Figure 4-5). This model contains both sewer and stormwater pipe system. Even though the area has mainly separate sewer system, there are some zones where combined manholes were built. This means that both the sewer and stormwater pipes go through them. This is represented in the model as link connections between the two systems so that if the capacity of one of them is exceeded, it will overflow to the other. The model outline is presented in Figure 4-6.

The model consists of around 3.0 km of sewer pipes and 5.0 km of stormwater pipes. The diameters range from 125 mm to 800 mm (see Figure 4-7), and the pipe material is mainly concrete, but PVC is also used. Several sub-catchments were defined in order to represent surface runoff. They cover the whole study zone and can be differentiated between sub-catchments representing roof areas and other areas (e.g., parking lots, roads, lawns etc.).

Figure 4-5: Representation of sewer model in Risvollan area. Taken from Kragset (2019).

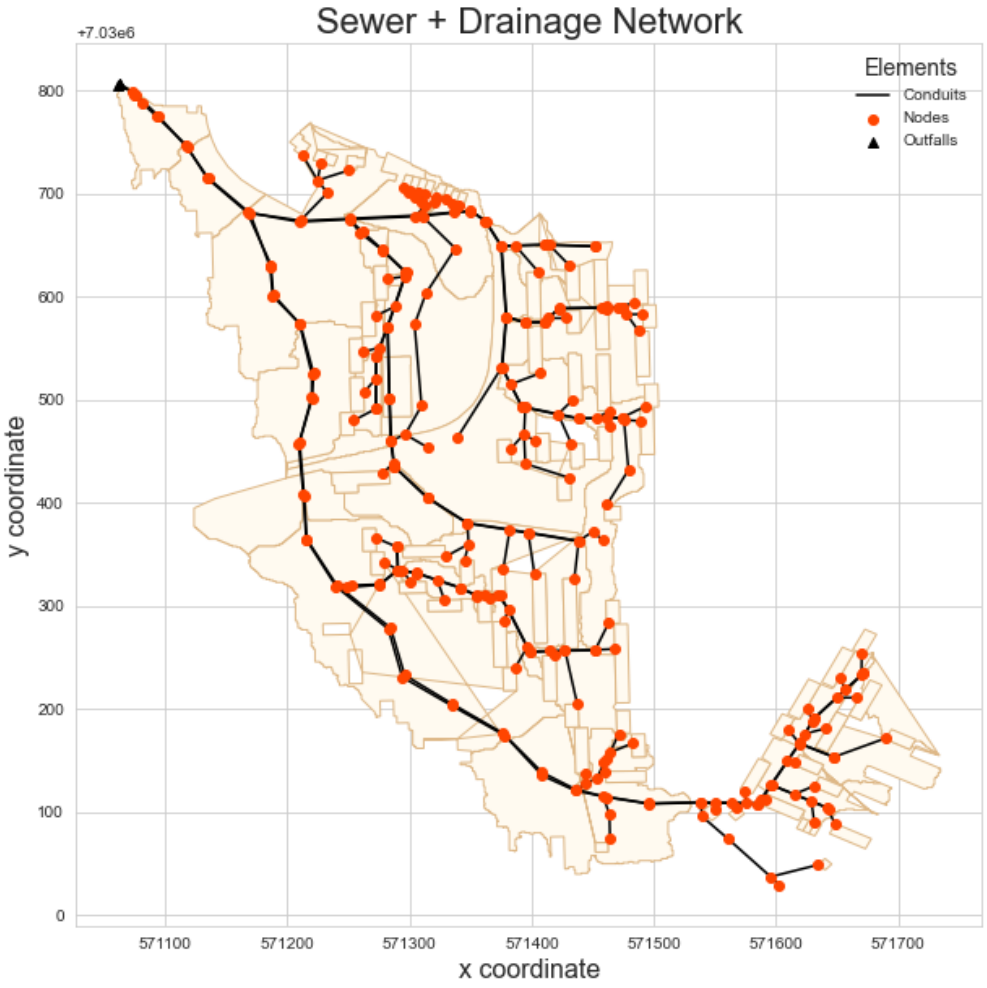
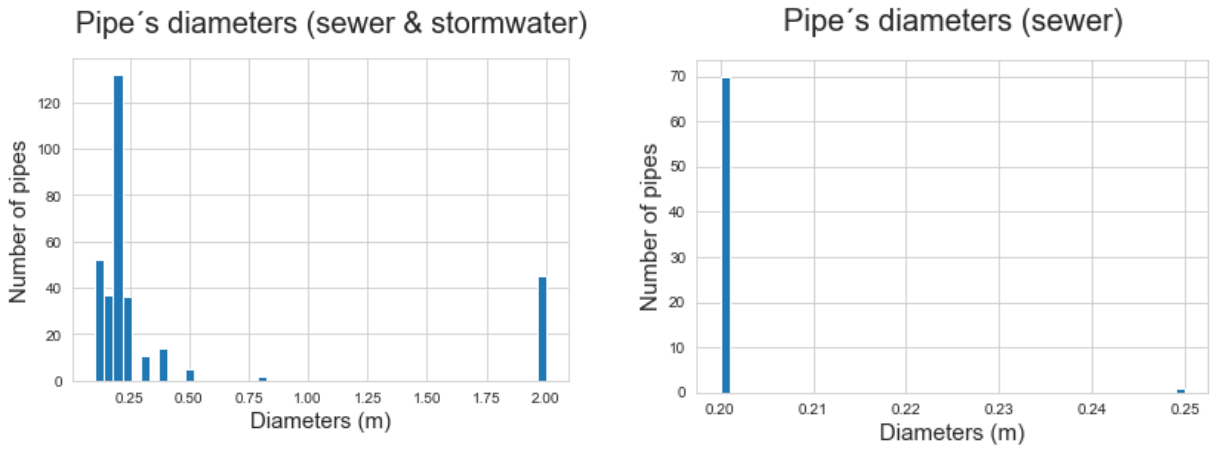


Figure 4-6: Model representation in SWMM.



(a) (b)

Figure 4-7: Pipe´s diameter distribution. a) sewer and stormwater diameter distribution, b) only sewer diameter distribution.

The model contains information to represent both dry and wet periods. Information for precipitation was obtained from a rain gauge present at a measurement station at the outlet of the catchment. Inflows from households were first estimated as a daily consumption pattern multiplied by the assumed water consumption (140 L/PE/day, PE refers to person equivalent). The model is calibrated for both dry and wet periods, as well as for the sewer and stormwater system. Calibration was based on data from a pressure sensor located in the sewer system (installed by NTNU) and data for both discharge and water level of stormwater obtained from Norwegian Water Resources and Energy Directorate-NVE. For the dry period, the calibration consisted of fitting the time of maximum and minimum flow from the model with a 7-days measured data. It must be stated that the assumption that the quantity input is 140L/PE/day was not modified during the calibration (for more information about calibration process refer to Kragset (2019)).

As a result of the calibration, inflows to the nodes in the model are defined as three types. First, the sewer baseflow from households was defined as a direct inflow, represented by a base demand for each area and a consumption pattern. Further, a constant inleak into the sewer was introduced in the "Dry weather" options in SWMM (see Appendix A: Inflows in SWMM, for more information regarding how to configure inflows in SWMM). This value represents the infiltration/inflow occurring during dry weather. It was determined by looking at the results from dry period modelling, and assuming that during nighttime the household discharges are almost zero. Lastly, the infiltration/inflow flow in wet conditions are represented by the RDII section in SWMM. RDII refers to rainfall dependent infiltration and inflow and is estimated using unit hydrographs defined by the rainfall volume that enters the sewer, the time to the peak of the unit hydrograph and the ratio between the time for recession to the time to peak (see Appendix A: Inflows in SWMM). In addition, a sewer shed area that contributes to RDII flow was given. In the model developed by Kragset (2019), RDII was calibrated only for the stormwater system and not for the sewer as data was not good enough to perform the calibration.

Figure 4-8 presents the flow in both the sewer and stormwater outfalls during a dry period simulation. The time period shown in the figure is the data availability provided in the sewer model for the study area.

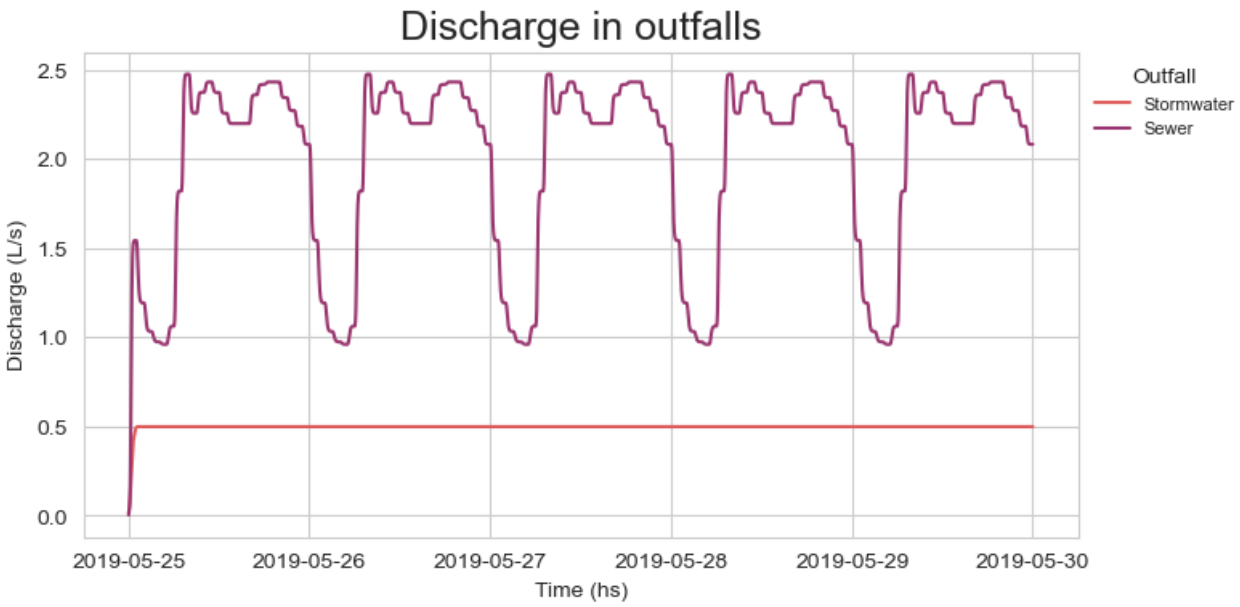


Figure 4-8: Discharge from stormwater and sewer outfalls in model.

5 Methodology

The next chapters describe the methods used in this thesis to understand and further model the relations between the water distribution network and the sewer network. The underlying idea was to set the basis on how the sewer flows can be obtained from the water distribution flows. To begin with, the systems are physically linked, where two approaches are suggested: one linking water node to sewer node, and another one linking the demands to a particular household that then discharges into a node in the sewer system. Then, the household demands are linked to the household discharges in the sewer model (see Figure 5-1). Last, a method to model the out leaks from the water system as inleaks into the sewer system is described (see Figure 5-1).

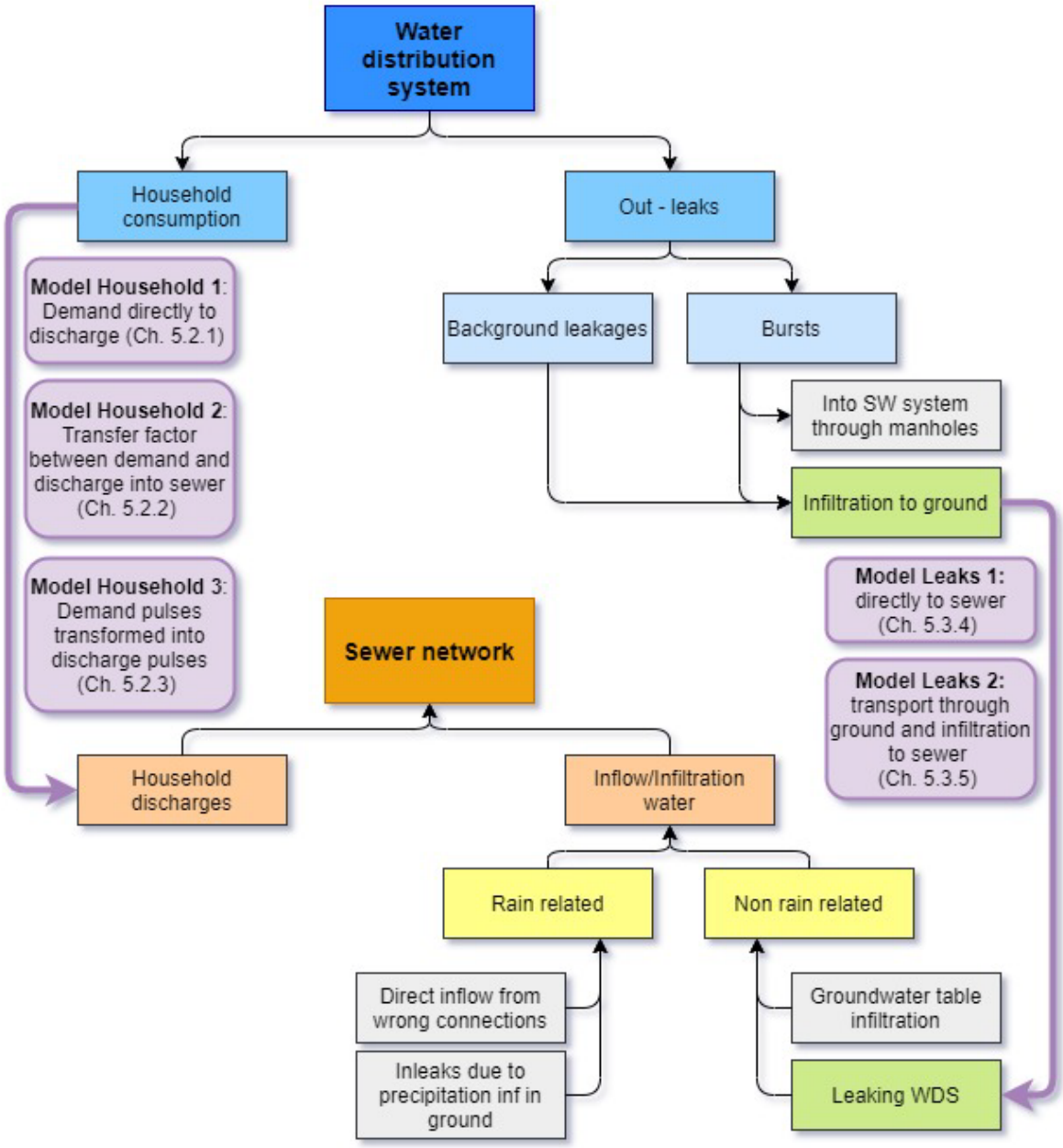


Figure 5-1: Conceptual model together with the models suggested in this study for each of the links.

This thesis was fully developed in Python, and the scripts can be found in Jaurena (2023). For the water distribution model, EPANET was used through the Water Network Tool for Resilience (WNTR) package in Python. WNTR is based in EPANET and features a versatile application programming interface (API) that enables adjustments to the network's operations and structure as well as the simulation of disruptive incidents and response actions (*Overview — WNTR 0.6.0dev Documentation*, n.d.). On the other hand, for the sewer system, SWMM was used through two software toolkits written in Python: PySWMM (*Overview — Pyswmm 1.2.0 Documentation*, n.d.) and Swmmio (*Swmmio — Swmmio 0.6.2 Documentation*, n.d.). Both tools allow to open and modify the SWMM model as well as to view the results from a simulation. However, for this thesis it was found more practical to interact with the flows setting with SWMMIO. PySWMM was then used to run and visualize the results.

5.1 Spatial linkage

To begin with, the physical interconnection between the water distribution system and the sewer system was defined. It is necessary to acknowledge that the spatial link aims to clearly define where the inflows to the sewer system will occur and to which demand/drinking water node they are associated. The idea behind the linkage of the two models resides on the understanding that the water consumed by a certain user, will be discharged (either fully or partially) into the near sewer system. Two strategies were developed for the spatial link for household demands. The first one directly connects the water nodes to its nearest sewer node, while the second one relates each household to a certain water demand from its closest water node and then assumed to discharge in the nearest sewer node. Both models are a simplification of reality, where connections of household to both the sewer and drinking water networks occur along the whole system and not only to the “nodes”.

5.1.1 Link Node to node

Generated water demands will probably be discharged in the sewer system nearby. This allows to define a first assumption where a water node will be related to its nearest sewer node (see Figure 5-2). As mentioned in (Zhang et al., 2021), this could result in unrealistic hydraulic status compared to the original system, but on the overall results the impact would be negligible. In practice, this was done by creating a buffer area around the sewer nodes and selecting the water nodes that fell in the defined area. This allowed to determine which are the nodes that are contributing to the discharge flows in the sewer system. Further, the relationship between systems was settled by an algorithm that identifies which is the closest sewer node for each selected water node according to the coordinates.

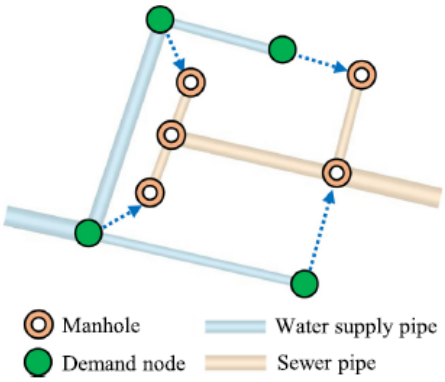


Figure 5-2: Node to node interconnectivity of water and sewer system. Taken from Zhang et al. (2021).

This method can result in two water nodes being assigned to the same sewer node. The script developed for this thesis does not take this into account, but a further improvement should be to consider the sum of both water demands being assigned to the sewer node.

5.1.2 Link through households

Based on the results obtained for the first linkage model (see Section 6.1.1), a second model for spatial link was proposed. Results for the first linkage model showed that for some sewer nodes the discharge was not accurately distributed. Therefore, this second model suggests to first associate the different households to a water demand. Then they were assumed to discharge that demand either fully or partially to the sewer system near the household (see Figure 5-3).

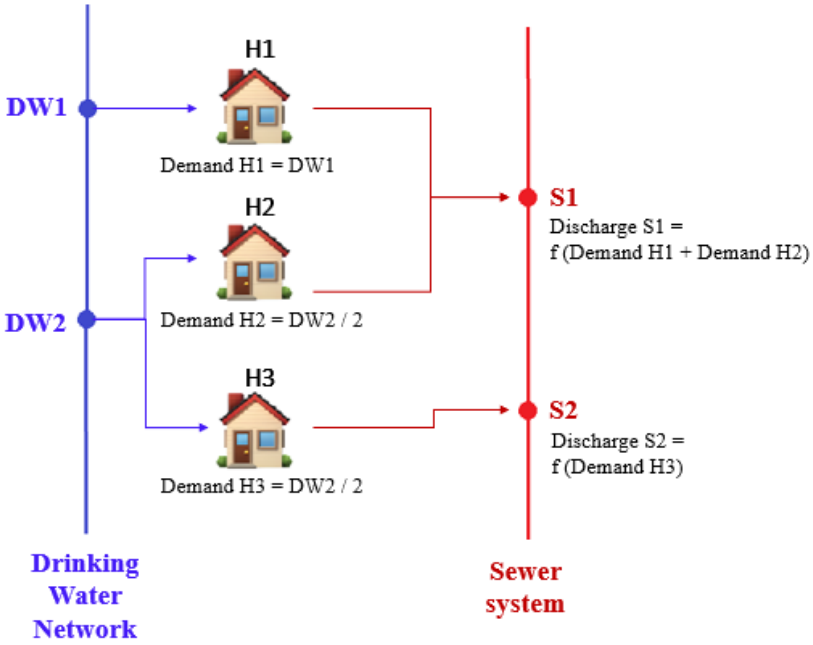


Figure 5-3: Proposed interconnectivity of water and sewer system.

In practice this was done by first identifying the location of the households in the area (using QGIS) for determining which was the nearest drinking water node to their centroids. This

resulted in a list of households with its corresponding node, where more than one could be linked to a single node. The demand for each household was calculated by dividing the demand of the water node by the number of households linked to that node. The same was done for the sewer system, where the nearest sewer node to each household´s centroid was identified and linked to the discharge of the different households. Moreover, one sewer node can receive the discharge of various households resulting in the inflow of that node being the sum of them. The link method is described in Figure 5-3.

5.2 Household demands

As mentioned previously, most of the water consumed by households will be discharged into the sewer system. In this study, three different ways on how to model this connection have been analyzed. In the first model suggested, the water demand is transformed directly into sewer discharge, similar to what is done in (Lund et al., 2021). The second model takes a step forward establishing that the discharges into the sewer system will be a function of the water consumption, a transfer factor, and a time delay (Zhang et al., 2021). Last, the third model transforms the water demand pulses into sewer discharges, considering both the residence time of drinking water in households and the attenuation as the liquid moves downstream from the house to the sewer (Elías-Maxil et al., 2014). It must be stated that the third model is only presented in this study and was not further developed.

5.2.1 Model H1: Demand directly to discharge

The first suggested model is based on the study made in Lund et al. (2021). They studied the possibility of smart meter consumption data being used to estimate the magnitude, timing, and spatial distribution of wastewater flow directly. In Lund et al. (2021), they assume measured data from both the water distribution and sewer systems is provided. As there are fewer sewer sensors, the smart water meters data is aggregated in order to be able to compare the data with the sensors in the sewer. Then, the aggregated water demand is placed in the sewer network as inflow. According to the result of the difference between the observed and the simulated sewer discharges, Lund et al. (2021) suggests different potential reasons for the deviations which are summarized in Appendix B: Study by Lund et al. (2021). As a result of their study, they suggest that it is conceivable and promising to directly estimate the wastewater flow using data from smart meters as input to hydrodynamic urban drainage models.

This concept was adapted to this master thesis, where smart meter consumption was not provided, and instead a water distribution model with the daily demands was given. The water demands were extracted from the water distribution model after having run an extended period simulation. Then the demands were placed into the sewer model as external inflows according to the link proposed in Chapter 5.1.

5.2.2 Model H2: Transfer factor between water demand and discharge into sewer

The second model studied for transforming household demand into sewer discharge is taken from Zhang et al. (2021) and adapted to the information available for this thesis. Zhang et

al. (2021) describes the relationship between water consumption and sewer discharge for certain user i as:

$$d_i = F(q_i, k_i, t_i) \quad (5\text{-}1)$$

Which indicates that the discharge is a function of the water consumption for user i (q_i), a certain transfer factor (k_i) and the delay time (t_i). This connection is based on the fact that losses will unavoidably occur while being transported inside the users' facilities.

The transfer factor represents the amount of water consumed by node i that has been collected by its associated manhole. It is suggested in Zhang et al. (2021) that the underlying relationship between nodal water use and manhole inputs can be modeled as a linear transfer function. Furthermore, according to Bailey et al. (2019) the transfer factor can be assumed to be approximately constant over time as the user properties remain almost unchanged over a short period of time. However, the characteristics of the water users, such as the user categories (commercial users or typical resident users), and the patterns of water use, can affect this component (Bailey et al., 2019). Due to this, the transfer factor must be calibrated for each demand node using the nodal water consumption data and sewer observations from the foul sewer system (such as sewer flow rates or water depth in manholes).

With regards to t_i, it represents the time it will take for the water consumed in node i to be discharged in the corresponding sewer node. The value of t_i can be affected by the user properties such as: the characteristics of the water supply area connected to the demand node, and the physical characteristics (such as length and slopes) of the connecting sewer pipes between users and the related manholes. Typically, t_i varies from a few minutes up to 15 minutes (Zhang et al., 2021). As a simplification, this time was dismissed since the output from the water distribution model was given with a time resolution of one hour (hourly multipliers in the nodes), and this is longer that the maximum delay time.

As mentioned previously the ki factors needed to be calibrated for the entire sewer system with n manholes that receive external inflows (water consumption). The transfer factor can be written as $K = [k_1, k_2, \dots k_n]^T$, and the objective function to minimize to calibrate it was defined by Zhang et al. (2021)

as:

$$Min: F(K) = \sum_{t=T_w}^{T} \left(\sum_{i=1}^{M} \left[g\left(h_i^O(t)\right) - g\left(h_i^S(t)\right) \right]^2 + \sum_{j=1}^{N} \left[g\left(f_j^O(t)\right) - g\left(f_j^S(t)\right) \right]^2 \right) \quad (5\text{-}2)$$

Where:

- T is the time period with observations used for the calibration.
- T_w is the warming-up time period for model setting up.
- M, N are respectively the numbers of observed water depths at manholes and flow rates in the sewer pipes.
- $h_i^O(t), f_j^O(t)$ are the observed water depth at manhole i and the observed flow rate at sewer pipe j for time t

- $h_i^S(t), f_j^S(t)$ are the simulated water depth at manhole i and the simulated flow rate at sewer pipe j for time t
- g ( ) is a linear function used to convert water depths and pipe flow rates into the same scale. The function used for this study is the one presented in Zhang et al. (2021) where $g(x) = \frac{x - x_{min}}{x_{max} - x_{min}}$, where x is the observed or simulated values in the monitoring points (either depth or flow), x_{min} and x_{max} are the minimum and maximum values determined by observations throughout the data on the calibration period.

The values for the simulated manhole depth and pipe flow, were obtained from running simulations in SWMM, where the nodes inflow for each timestep was determined by the following expressions:

$$D(T) = \begin{bmatrix} d_1(t_1), d_2(t_1), \dots d_n(t_1) \\ d_1(t_2), d_2(t_2), \dots d_n(t_2) \\ \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \\ d_1(T), d_2(T), \dots d_n(T) \end{bmatrix} \quad (5\text{-}3)$$

$$d_i(t) = k_i \times q_i(t) \quad (5\text{-}4)$$

$$k^{min} \leq k_i \leq k^{max} \; i = 1,2, \dots n \quad (5\text{-}5)$$

Where:

- $D(T)$ is a matrix with dimensions $T \times n$, which represents the inflows of the manholes across the total time period T.
- $d_i(t)$ is the inflow rate at manhole i at time t.
- $q_i(t)$ is the water consumption for node i at time t , extracted from the water distribution model.
- k^{min}, k^{max} are the minimum and maximum value that the transfer factor k can take for the optimization, in this case 0.7 and 1.0 are the minimum and maximum values respectively and were determined by engineering experience as suggested in Zhang et al. (2021).

In this study the optimization was solved by combining the use of an Evolutionary Algorithm and the software SWMM (run by both '*pySWMM*' and '*swmmio*' in Python). The EA used was a differential algorithm which is run by the use of the package SciPy in Python ('*scipy.optimize.differential_evolution*'). More information on the EA used to find the optimal transfer factors can be found in Appendix C: Differential evolution algorithm.

It is worth noting that the calibration of the k factors was done once and can be done based on historical data. After calibration, the model can predict the sewer discharge in real time if the water demands are provided in real time, and by the use of k factors.

5.2.3 Model H3: Demand pulses transformed into discharge pulses.

This studied model is based on Elías-Maxil et al. (2014) and E. J. Pieterse-Quirijns et al. (2012). The basis of the model is to transform water demand patterns obtained from statistical and demographic data, into sewer discharges taking into consideration the time delay and the attenuation when the liquid is transported from one system to the other.

First, they assumed that the water demand patterns (Q) in each household can be represented as the addition of rectangular water pulses (B), as presented in Figure 5-4. The pulses represent the demand required by the appliances present in each household, whose usage will depend on the habits and age of the inhabitants of each household. Monte Carlo method is used to propagate these rectangular pulses per appliance with an intensity (I), duration (D) and time of occurrence in the day (τ)(Blokker et al., 2010).

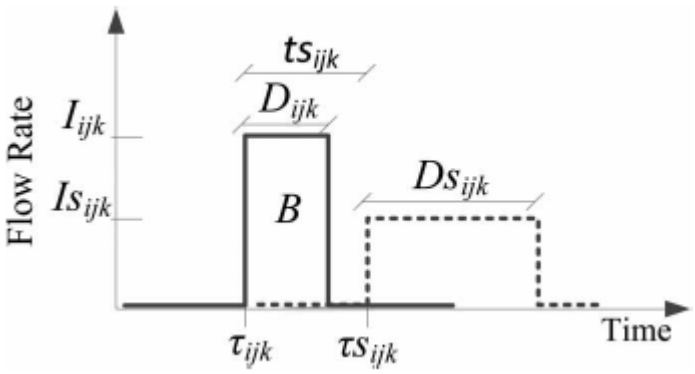


Figure 5-4: Block function of a rectangular water pulse (filled line), and sewer discharge pulse (dotted line). Taken from Elías-Maxil et al. (2014).

The household demand can therefore be described mathematically by Equation (5-6) where B is a flow value and i, j, and k represent the busy times, users and equivalent appliances of a house, described by an intensity I within the time interval $[\tau, \tau + D]$ and 0 in the remaining time period T (24 hours) (Elías-Maxil et al., 2014).

$$Q = \sum_{k=1}^{M} \sum_{j=1}^{N} \sum_{i=1}^{F_{jk}} B\left(I_{ijk} D_{ijk} \tau_{ijk}\right) \qquad (5\text{-}6)$$

It is worth noting that the duration and intensity of the pulse depends on the appliance used, while the time of occurrence will be related to the age, occupation, and inhabitants of the household (Elías-Maxil et al., 2014). The information required for the Monte Carlo method is detailed below, and it can be obtained from surveys:

- For the equivalent appliances (i): Penetration rate of the appliances (the percentage of households that have the appliance), which is linked to the size of the house.
- For the users (j): information about the number of users in a household, divided by age, gender, and occupation.

- For the frequency of use of the appliances (k): this is obtained from probabilistic relations of the users and the type of appliances they use.

This method for generating water demand patterns was further developed in a stochastic model called SIMDEUM which stands for "SIMulation of water Demand, an End-Use Model". It is based on statistical information regarding water user's household composition (number and age), home installed water appliances (typical flow and volume) and consumers behavior related to water (number of toilet flushes, duration of showers, preference of time to use appliances) (Blokker et al., 2010). Figure 5-5 shows a brief description on how SIMDEUM works.

SIMDEUM determines for each appliance when, who, and for what purposes it is utilized resulting in a demand pattern of cold and hot water for each appliance. The demand pattern of a home, office, hotel, or nursing home is generated by adding the demand patterns of all appliances (I. Pieterse-Quirijns, 2014). SIMDEUM has certain distinctive features that set it apart from conventional demand models. These include its physical foundation, stochastic character, small spatial scale (customer tap) and temporal scales (1 s) (Blokker et al., 2017). At different aggregation levels from 1 s to 1 h, from a single family to a small town, SIMDEUM produces genuine needs and realistic variations in those demands, during the day as well as across days and between demand nodes (Blokker et al., 2017).

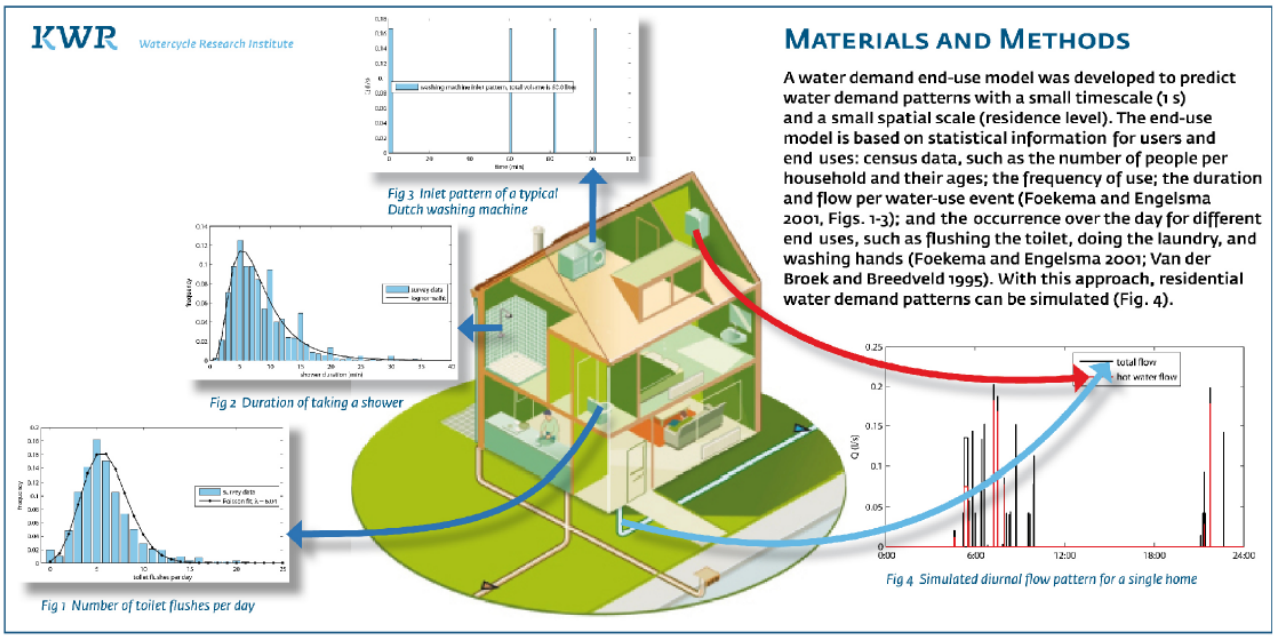


Figure 5-5: Explanation on how SIMDEUM DW works, taken from Blokker et al. (2017)

Therefore, if the statistics of users and consumption is available for the study area, the demand pulses can be obtained. Then, the water pulse is transformed into wastewater generation pulse inside the households as shown in Figure 5-4. After the water is used by an appliance, it is discharged into the sewer after a residence time (t_s), so the expected time of discharge can be described as $\tau_s = \tau + ts$ (Elías − Maxil et al., 2014). The duration of the discharge into the sewer is represented by Ds and depends on the type of appliance. Moreover, the discharge´s intensity depends both on the volume discharge and the duration of every discharge and can be determined as: $I_s = D \times I/D_s$. Table 5-1 presents values for D, I, t_s and D_s extracted from Elías-Maxil et al. (2014). Once the sewer pules are obtained, the total discharges can be calculated and introduced into the sewer models as external inflows.

Table 5-1: Duration, intensity, and time delay. Taken from Elías-Maxil et al. (2014)

Equivalent Appliance	D (s)	I (L/s)	t_s (s)	D_s (s)
Shower	600	0.123	45	Same as D
Kitchen tap	16|48|15|37	0.083|0.125	30	Same as D
Toilet	45–106	0.042|0.884	0	9
Bathroom tap	40|15	0.042|0.042	2700	Same as D
Washing machine	120[a]	0.167|0.083	3840|1260	300[a]
Dish water	21[a]	0.19[a]	1800[a]	120[a]

|: Separation of sub-activities or cycles.
[a]The same parameter was included in the remaining three cycles.

In conclusion, this method requires first to get the water demands as rectangular pulses in order to transform them using the time lag and discharge time into sewer discharges. Detailed information from the study area needs to be provided to get pulses that represent accurate enough the demand of the area. As mentioned before, the data needed includes information about the composition of the household (number, age occupation), together with the type and number of appliances present and statistics about the duration and intensity of each appliance. Then, further data is required to transform the water pulses into sewer pulses such as the time delay between the consumption and the discharge and the duration of the discharge.

This method is not developed in this master thesis due to lack of data together with time limitations. If information from previous studies such as Elías-Maxil et al. (2014) and Blokker et al. (2010) where to be used, verification need to be done to see if they match with the behaviors of the area under study, as mentioned studies are done using Netherland´s statistics and population behaviors. The water pulses could be generated by SIMDEUM whose Python version is freely available, but still under development. However, the package is built on statistics from the Netherlands and was not trivial to change. All things considered, this method is seen as a good opportunity to link SIMDEUM with sewer discharges and needs further work in order to implement it.

5.3 Leaks modelling

Ageing infrastructure together with lack of maintenance and high repair costs leads to having leaks in the water distribution system. Additionally, the sewer pipes can have cracks that allow external water to enter the system. This section aims to describe a method to represent the in leaks into the sewer system due to background leaks and burst leaks occurring in the drinking water system.

Leak to in-leak modelling consists of 1) defining the spatial link between the leaking water pipe and the receiving sewer system; 2) quantifying the flow of water that is being leaked from the water distribution system; 3) determining how the flow is being transported through the ground and infiltrating into the sewer system.

To quantify the flow coming out of the water pipes, it was differentiated between pipe bursts and background leakages. For pipe bursts, the flow was determined by using the method proposed by Fuchs-Hanusch et al. (2016). For background leakages, they were extracted directly from the water distribution model, where they were provided as constant demand patterns. This was due to lack of information about the user's consumption in order to be

able to determine the flow by for example the Minimum Night Flow method. Both methods are thoroughly described in the following sections.

As for the link between leaks and in-leaks, first a model connecting the leaks directly to the sewer systems was suggested, to then develop a more advance method where the transport on the ground was considered as well as the characteristics of the sewer defects.

5.3.1 Spatial link

In Norway water and sewer networks are placed in the same trench. Water leaking from the water distribution network will become an inleak into the sewer system if they are located close enough that the water leaked can be transported from one system to the other. Taking this into consideration, first it was determined which of the water pipes could have an influence on the sewer pipes if leaks happen. This was done by creating a buffer area around the sewer pipes, and a fictional node in the middle of the water pipes that will act as a representative of the latter (see Figure 5-6). If the fictional node was contained by the buffer area of the sewer, the leaks from the water distribution system will have the potential to be transformed into in-leaks to the sewer. Then, in-leaks to the sewer were placed as inflows in the nearest sewer node to the fictional node.

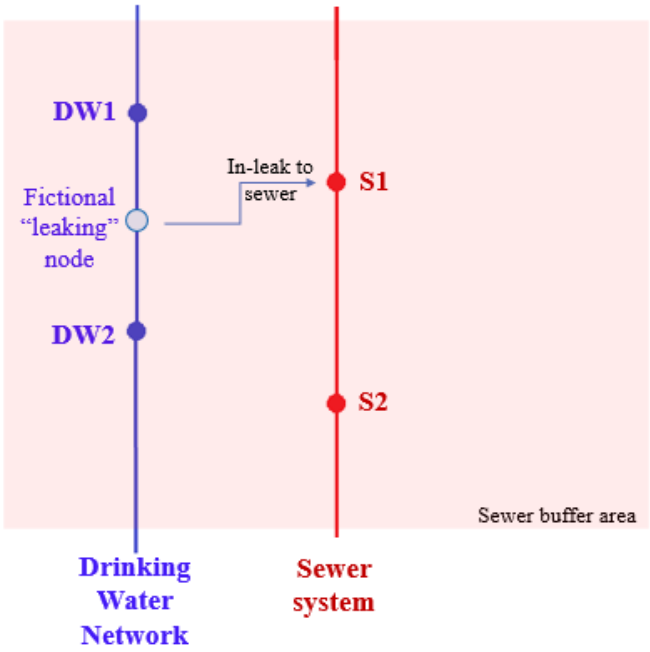


Figure 5-6: Spatial link for modelling leaks.

5.3.2 Determine burst outflow

The method proposed by Fuchs-Hanusch et al. (2016) was used to determine the flow coming out of a pipe burst. Outflow from leaks is usually represented by Torricelli´s law (where the flow through an orifice is proportional to the square root of the pressure at the orifice). However, it has been shown that the pressure dependency with the leakage is larger than the Torricelli´s law. As a result, the FAVAD (Fixed and variable area discharge Concept) concept was developed which suggests that cracks in elastic materials enlarge depending on

the pressure. In EPANET a pipe burst outflow can be modelled using an emitter which follows Equation (5-7).

$$Q = c_e \times h^{e_e} \quad (5\text{-}7)$$

Where:

- Q = discharge rate when the pipe breaks
- h = nodal pressure head
- e_e = emitter exponent
- c_e = emitter coefficient

The emitter exponent can have values of 0.5 for Torricelli´s law, while for FAVAD it can have values up to 2.0. Additionally, c_e can be calculated for an emitter as:

$$c_e = c_d \times A \times \sqrt{2g} \quad (5\text{-}8)$$

Where:

- c_d = Discharge coefficient. It can vary between 0.5 and 0.7 according to the Reynolds number and shape of the orifice (Lambert, 2001). In this case 0.6 was used.
- A = Area of the crack
- g = Acceleration due to gravity

Furthermore, Cassa et al. (2010) developed the equation to represent the leakage outflow in the FAVAD context. This expression comes as a result of several studies investigating the behavior of elastically deforming leaks so as to be able to estimate emitter exponent and coefficient with various boundary conditions.

$$Q = c_d \times \sqrt{2g} \times (A_0 \times h^{0.5} + m \times h^{1.5}) \quad (5\text{-}9)$$

Where:

- A_0 = Initial crack size under pressurized conditions
- m = head-area slope (potential of a pipe to increase the leak size under pressure). It depends on specific parameters from the pipe such as Young´s modulus, E, and the type of failure mode.

Fuchs-Hanusch et al. (2016) propose a systematic model of leakage outflow from pipe bursts using EPANET and including the FAVAD concept described in (5-9). They suggest an iterative approach to determine the emitter exponent and coefficient based on the research by Van Zyl & Cassa (2014). The steps for the iteration are described below:

1. Suppose a pipe burst in a pipe:
    a. Create a dummy node in EPANET in the middle of the pipe under study.
    b. Extract the material parameters of the pipe under study and determine the Young´s module (E), the pipe diameter (d) and the wall thickness as it will be needed to calculate the head-area slope. The values used for this thesis are presented in Table 5-2.

Table 5-2: Young Modulus, and thickness depending on material and diameter used for this thesis. SJK refers to Cast Iron, PVC to polyvinylchloride and (PE) polyethylene.

Material	E(GPa)
SJK	81.00
PVC	3.30
PE	3.45

Material	Diameter (mm)	Thickness (mm)
SJK	100	5.74
SJK	150	7.11
SJK	200	8.18
SJK	250	9.27
PVC	99	5.30
PVC	144	7.70
PVC	180	7.70
PVC	203	10.80
PE	51	3.00
PE	102	6.70
PE	114	8.10

c. Run a simulation with the emitter coefficient set to zero (indicating no pipe burst) and determine the pressure head (h) at the dummy node.
d. Compute the leakage outflow using equation (5-9) and the calculation approach developed by (Cassa & Van Zyl, 2013) for the head-area slope which depends on the type of crack:
    i. For longitudinal cracks:

$$m_{long} = \frac{2.93157 \cdot d^{0.3379} \cdot L_c \cdot 10^{0.5997(\log L_c)^2} \cdot \rho \cdot g}{E \cdot t^{1.746}} \quad (5\text{-}10)$$

$$A_{crack} = L_c \cdot W_c \quad (5\text{-}11)$$

    ii. For circumferential cracks:

$$m_{circ} = \frac{1.648 \cdot 10^{-5} \cdot L_c^{4.880} \cdot \sigma_l^{1.092} \cdot 10^{0.828(\log L_c)^2} \cdot \rho \cdot g}{E \cdot t^{0.338} \cdot d^{0.186}} \quad (5\text{-}12)$$

$$\sigma_l = \frac{h \cdot d}{4 \cdot t} \quad (5\text{-}13)$$

$$A_{crack} = P_c \cdot d \cdot \pi \cdot W_c \quad (5\text{-}14)$$

Where:

- m = Head area slope
- A_{crack} = Area of the crack
- L_c = Length of the crack
- W_c = Width of the crack

- d = Pipe diameter
- P_c = Impacted perimeter of the pipe
- E = Young´s modulus
- t = Wall thickness of the pipe
- σ_l = Longitudinal stress of the pipe
- ρ = Density of the fluid transported by the pipe
- g = Gravity´s acceleration

The values used for W_c, L_c and P_c for this thesis were the ones used by Fuchs-Hanusch et al. (2016) and are presented in Table 5-3 and Table 5-4. Several studies such as WSAA (2003), Sorge (2007) and Friedl (2012), have shown that for cast iron (CI) the failure modes that can occur are: longitudinal crack, circumferential crack, corrosion cluster, shard fracture and leaking joint. However, for plastic pipes such as PVC and PE the most probable cracks are longitudinal.

Table 5-3 Parameters for longitudinal cracks,taken from Fuchs-Hanusch et al. (2016). *Where AC refers to asbestos cement, CI to cast iron, and PE to polyethylene.*

Material	Wc_{min} (mm)	Wc_{max} (mm)	Lc_{min} (mm)	Lc_{max} (mm)
PE	0.1	2	500	2000
AC	0.1	2	500	2000
CI	0.1	2	500	2000

Table 5-4: Parameters for Circular crack, taken from Fuchs-Hanusch et al. (2016). *Where AC refers to asbestos cement, CI to cast iron, and PE to polyethylene.*

Material	Wc_{min} (mm)	Wc_{max} (mm)	Pc_{min} (−)	Pc_{max} (−)
PE	0.1	2	0.1	0.4
CI	0.1	2	0.1	0.4

e. Calculate the emitter exponent and coefficient with Equations (5-15) to (5-17) from Van Zyl & Cassa (2014). These equations are derived from the ones presented by Cassa & Van Zyl (2013) as EPANET can calculate pressure dependent demands with e_e and c_e.

$$L_N = leakage\ number = \frac{m \cdot h}{A_0} \quad (5\text{-}15)$$

$$e_e = \frac{1.5 \cdot L_N + 0.5}{L_N + 1} \quad (5\text{-}16)$$

$$c_e = Q \cdot h^{-e_e} \quad (5\text{-}17)$$

f. Insert the emitter coefficient and exponent into the model and run a simulation to extract the pressure head in the dummy node which is now leaking. Further repeat steps d) to e).

2. Steps d) to f) are repeated until the leakage outflow converges.
3. Once it has converged, a dummy node is assigned to the following pipe based on the original model. Steps 1) to 2) are repeated until all the pipes of interest are covered.

This procedure requires knowledge about the type of crack occurring in the burst, together with its dimensions. In order to get an approximate outflow from a burst occurring in any of the pipes in the water distribution system provided, Monte Carlo simulations were done varying the type of crack together with the length and width of the crack or the impacted perimeter of the pipe using a uniform distribution.

5.3.3 Determine background leak´s flow

Background leaks are usually more difficult to identify as they do not generate an evident pressure drop in the distribution network, the leaked flow is low and are generally distributed along the pipes (Berardi et al., 2015). One of the generally used methods to detect background leakages is the Minimum night flow method (García & Cabrera, 2006). This method requires the measurement of night flow to assess the flow rate supplied to a certain area during the hours of less flow in the night. To determine the losses, two types of calculations are required. First, by subtracting the user's night consumption from the metered night flow, the water losses at the minimum night flow hour are calculated. Then, they need to be extrapolated for the full period of 24 hours. This is done by a simplified method that relates the flow and pressure together with the pressure along the day (García & Cabrera, 2006).

However, for this thesis is not possible to apply this method as the information provided does not include measurements from the consumer´s demand. Nevertheless, the water distribution model provided includes an estimation of the background leakages. They are inserted in the model as a constant pattern in the nodes and a certain base demand associated with the specific node. Therefore, in this case the background leakages are extracted from the water distribution model.

5.3.4 Model L1: From water leak directly to sewer in-leak

In order to get a first estimate of the behavior of the leak into the sewer, the outflows coming out of the leaking water pipe were connected directly to the sewer. This model is fictional as there will probably be a transport through the soil before infiltrating into the sewer. It could happen however that if the pipe burst is big, the outflow can end up running in the surface and entering the sewer system directly through manholes. Nevertheless, this situation would probably cause a collapse of the soil package or even in the sewer pipe, not representing the aimed effect that is the infiltration into the sewer.

5.3.5 Model L2: Water leak transported through soil and transformed into sewer in-leak

Once the water has leaked from the water distribution network, it will be transported in the soil surrounding the pipe and will eventually end up in the sewer system if the pipes have cracks. When investigating how to model in-leaks caused by the water distribution network into the sewer system few supporting literature was found. However, some studies such as

DeSilva et al. (2005) and Karpf & Krebs (2011) suggest a method to represent groundwater infiltration into sewer. This method together with some engineering assumptions was adopted in this master thesis to develop an infiltration model into sewer caused by leaks of the water distribution system.

DeSilva et al. (2005) and Karpf & Krebs (2011) approach to describing infiltration into the sewer due to groundwater is based on Darcy´s Law, and calculate infiltration as:

$$Q_{in} = k_f \cdot A_{leak} \cdot \frac{\Delta h}{\Delta l} \qquad (5\text{-}18)$$

Where:

- Q_{in} = Flow into the sewer
- k_f = Hydraulic conductivity of the surrounding soil
- A_{leak} = Area of the defect in the sewer
- Δh = Difference between groundwater level and the pipe defect (see Figure 5-7)
- Δl = Colmation layer, that is assumed to be 10cm according to DeSilva et al. (2005) (see Figure 5-7).

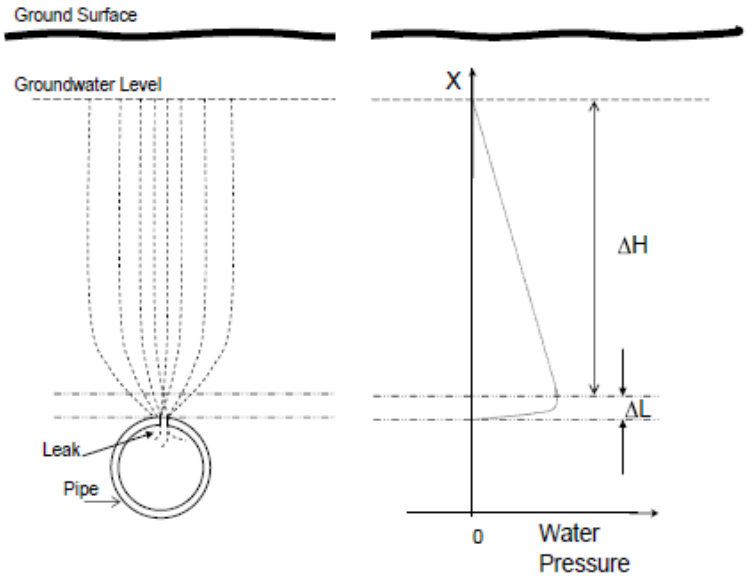


Figure 5-7: Groundwater infiltration into sewer. Taken from DeSilva et al. (2005).

It is relevant to mention that for this master thesis no data about previous infiltration to sewer due to background leakages or burst was available. Therefore, the assumptions made to model the process are the ones that were considered to best represent and simplify the problem. The idea was to set the basis for the analysis of the infiltration to sewer of water leaks, which it will need further study to get a fully approved model. The assumptions made to represent the in-leaks to the sewer system due to out-leaks in the water distribution network were the following:

- The water distribution pipe is located above the sewer pipe in the trench, with a fixed distance of 0.5m.

- The leaks from the water distribution pipes occur in the "bottom" part of the pipe and therefore the water will spread downwards (see Figure 5-8).
- The trench is filled with gravel and it is well drained, thus it will conduct the water in the horizontal and vertical direction once it is out of the water pipe. This means that the water will spread along the length and width of the trench in which the pipe is located and accumulate until it covers the distance between the water pipe and sewer pipe. Then, it will start to infiltrate into the sewer.
- It is assumed that the water that comes out and does not either accumulate in the area between the water pipe (see Figure 5-8) and the sewer or infiltrate into the sewer, it is getting out of the analyzed system. This means that it can be further infiltrating into the ground below the sewer pipe, expanding in the trench further than the pipe´s length or infiltrating beyond the trench walls. Nevertheless, the excess water does not play a role in the infiltrated water to the sewer pipe that is being analyzed.
- The groundwater table is below the sewer pipes, and hence will not influence the infiltration into the sewer.
- It is assumed that the wastewater in the sewer pipes is not flowing under pressure and the in leak into the sewer will therefore be discharging in free flow.
- Once the water is out of the water pipe, the pressure is dissipated by the surrounding soil package. Therefore, the relevant variable from the out-leak will be its flow and not the pressure.

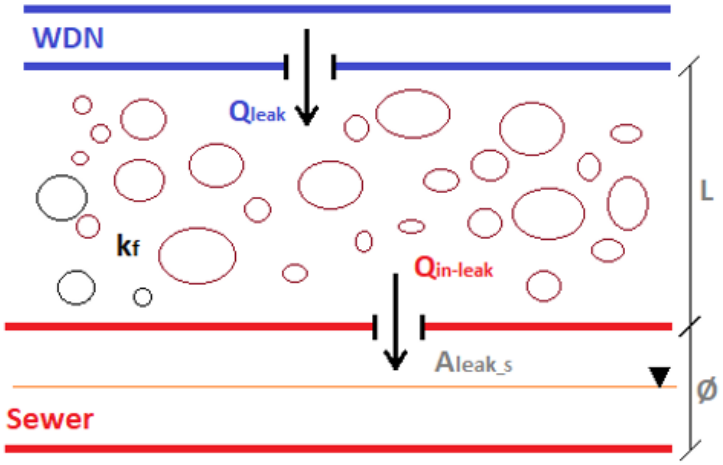


Figure 5-8: Schematic representation of the variables needed for the transport from water leak into sewer in-leak.

In the following sections the process is further described for both pipe burst and background leakages due to their intermittent and constant characteristics respectively.

a) Pipe burst

As described previously, a pipe burst occurs spontaneously, and it is usually rapidly detected and repaired. In order to model the in-leaks, different situations must be identified according to the time since the burst was initiated. Once the pipe burst happens (t_0), the water will infiltrate into the sewer when: a) the accumulated out-flow reaches the sewer; b) when the time exceeds the time required for water to travel through the soil before reaching the sewer (t_s). This time will depend on the characteristics of the soil matrix and the distance between the water and sewer pipe, as it is assumed that the pressure of the water coming out of the

leak is dissipated. After this, the leaking volume will have the potential to infiltrate into the sewer.

The infiltration flow into the sewer will be given by Equation (5-18). For this model, Δh will be described by $h(t)$, which aims to represent the water movement through the soil from the burst leak to the sewer defect. This process is assumed to be like a tank being filled up and discharging through an orifice. The orifice in this case is the sewer crack, which is assumed to discharge freely into the sewer pipe. The 'tank' is the soil matrix, delimited by the length and width of the trench and the distance between the water pipe and the sewer. The volume delimited by these dimensions will be referred to as maximum volume. Once the water accumulated reaches this value, it will be assumed to be leaving the studied system and therefore not contribute to the infiltration into the sewer. When the pipe burst is fixed (t_f), the leak will cease, but water can still infiltrate into the sewer pipe if the soil matrix contains water from previous steps. The process is described thoroughly below:

1. When $t < t_0$:
   The pipe burst has not yet occurred, therefore:
   - $Q_{leak}(t) = 0$
   - $h(t) = 0$
   - $Q_{in-leak}(t) = 0$

2. When $t_0 \leq t < t_0 + t_s$:

   Where t_s can be calculated as $t_s = k_f/L$, being k_f is the soil´s permeability and L the distance between the water distribution pipe and the sewer pipe. During this time period, water starts to flow out of the water distribution pipe and infiltrating in the soil. There will be no infiltration into the sewer until $t = t_0 + t_s$ or if the maximum volume is reached. The maximum volume is given by:

$$V_{max} = L \cdot A_t = L \cdot (W_t \cdot L_{pipe})$$

(5-19)

   Where:
   - L is the distance between water pipe and sewer defect (assumed 50cm for this thesis)
   - A_t is the area of the trench, which can be calculated as $A_t = W_t \cdot L_{pipe}$ being W_t the width of the trench and L_{pipe} the length of the pipe under study.

   As A_t is constant for a given pipe, the infiltration will occur when $h(t) \geq L$. Therefore, the infiltration process for these timesteps can be described as follows:

   - $\boldsymbol{Q_{leak}(t) > 0}$, it´s value will be the flow calculated previously as leak from the water distribution network.
   - $Q_{in-leak}(t)$ will depend on the water level in the soil ("tank") in time t:

$$h^* = h(t-1) + [Q_{leak}(t) - Q_{in-leak}(t-1)\,] \cdot \frac{\Delta t}{A_t}$$

(5-20)

Where $h(t-1)$ and $Q_{in-leak}(t-1)$ is the value of the water level in the soil and in-leak flow for the previous timestep respectively, and Δt is the time between timesteps. Then,

a) If $h^* < L \rightarrow \begin{cases} h(t) = h^* \\ \boldsymbol{Q_{in-leak}(t)} = 0 \end{cases}$

(5-21)

b) If $h^* \geq L \rightarrow \begin{cases} h(t) = L \\ \boldsymbol{Q_{in-leak}(t)} = \boldsymbol{A_{leak}} \cdot \frac{\boldsymbol{k_f}}{\boldsymbol{\Delta l}} \cdot \boldsymbol{h(t)} \end{cases}$

(5-22)

3. When $t_0 + t_s \leq t < t_f$:

   Between these timesteps the water from the leak will have already reached the sewer defect (if it had not done it before) and will infiltrate into the sewer. As in the previous case, $h(t)$ will be limited by the maximum storage assumed for the soil package, therefore:

   - $\boldsymbol{Q_{leak}(t) > 0}$
   - $Q_{in-leak}(t)$ will depend on the storage height in time t:

$$\boldsymbol{Q_{in-leak}(t)} = \boldsymbol{A_{leak}} \cdot \frac{\boldsymbol{k_f}}{\boldsymbol{\Delta l}} \cdot \boldsymbol{h(t)} \; with \begin{cases} h(t) = h^* & if\ h^* < L \\ h(t) = L & if\ h^* \geq L \end{cases}$$

(5-23)

   With h^* being calculated as in Equation ((5-20).

4. When $t \geq t_f$:

   In these timesteps, the leak has been repaired but there might still be water in the soil column with potential to infiltrate into the sewer.

   - $Q_{leak}(t) = 0$
   - $Q_{in-leak}(t)$ will depend on the storage height in time t:

$$h^{**} = h(t-1) + [-Q_{in-leak}(t-1)\,] \cdot \frac{\Delta t}{A_t}$$

(5-24)

a) If $h^{**} > 0 \rightarrow \begin{cases} h(t) = h^{**} \\ \boldsymbol{Q_{in-leak}(t)} = A_{leak} \cdot \frac{k_f}{\boldsymbol{\Delta l}} \cdot h(t) \end{cases}$

(5-25)

b) If $h^{**} \leq 0 \rightarrow \begin{cases} h(t) = 0 \\ \boldsymbol{Q_{in-leak}(t)} = 0 \end{cases}$

(5-26)

The leak area for this master thesis was estimated from the curves presented by DeSilva et al. (2005), see Figure 5-9. These curves are presented for vitreous-clay pipe in sandy soil, but it is stated in DeSilva et al. (2005) that more curves for concrete and PVC are available. The later curves were not easily found and therefore for this master thesis the curves in Figure 5-9 were used as an estimation of the leak area, that will need further adjustments.

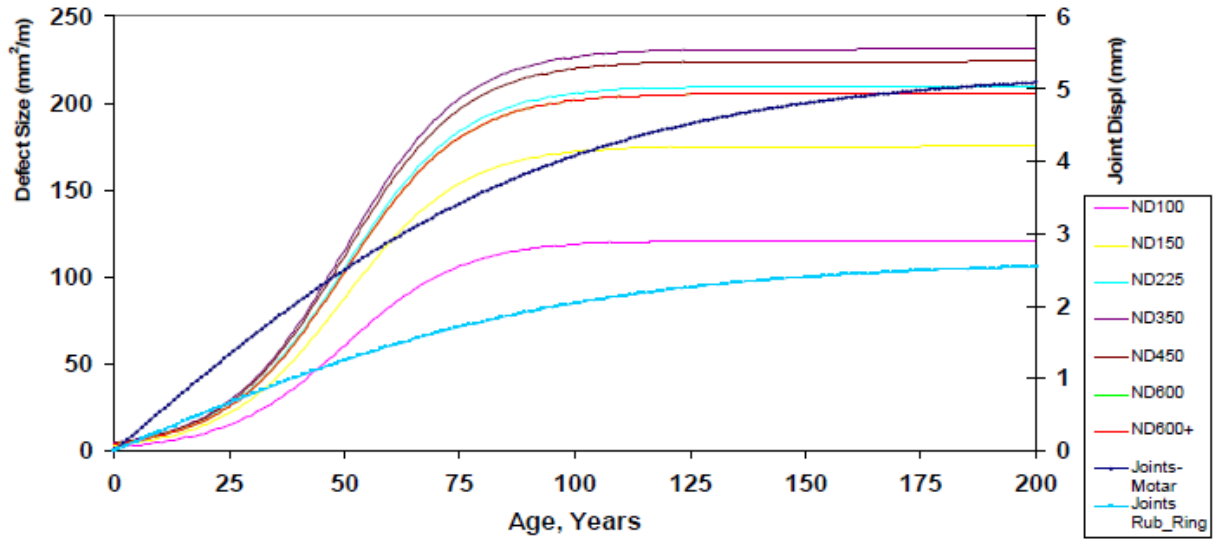


Figure 5-9: Leak area in sewer, depending on age and nominal diameter (ND). Taken from DeSilva et al. (2005).

b) Background leakage

Background leaks can be assumed to be constant, as they remain undetected for long periods of time. This characteristic allows to define the transport through the soil for all times as follows:

- $\boldsymbol{Q_{leak}(t) > 0}$**,** for every t in the simulation
- $h(t) = L$, as it is assumed that the water has already reached the sewer in previous timesteps as the leak has been going on indefinitely and will therefore cover the distance between the water pipe and the sewer defect.
- $\boldsymbol{Q_{in-leak}(t)}$ will be limited by the flow coming out of the water pipe, the potential flow to infiltrate can be calculated as

$$q = A_{leak} \cdot \frac{k_f}{\Delta l} \cdot h(t) \qquad (5\text{-}27)$$

$$\boldsymbol{Q_{in-leak}(t)} = \begin{cases} \boldsymbol{q} & if\ q \leq Q_{leak} \\ \boldsymbol{Q_{leak}} & if\ q > Q_{leak} \end{cases} \qquad (5\text{-}28)$$

5.4 Model performance evaluators

To assess the model performance, several statistical metrics are used in this study. The metrics are absolute percentage error (APE), mean absolute percentage error (MAPE), the coefficient of determination (R^2), the Nash-Sutcliffe model efficiency (NSE), and the Kling-Gupta Efficiency (KGE). The metrics were selected because of their wide use in assessing model performance within the water resources (Zhang et al., 2021). The mathematical description of each of the methods is presented below:

APE and MAPE:

$$APE = \left| \frac{Q_{i,O} - Q_{i,S}}{Q_{i,O}} \right| \times 100\% \quad (5\text{-}29)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{Q_{i,O} - Q_{i,S}}{Q_{i,O}} \right| \times 100\% \quad (5\text{-}30)$$

Where:

- $n =$ total number of data points
- $Q_{i,O} =$ is the observed value.
- $Q_{i,S} =$ is the simulated value.

A lower value of either APE or MAPE indicates a general good performance of the model.

Coefficient of determination R^2:

$$R^2 = \frac{\left[\sum_i (Q_{i,O} - \overline{Q_O})(Q_{i,S} - \overline{Q_S})\right]^2}{\sum_i (Q_{i,O} - \overline{Q_O})^2 \sum_i (Q_{i,S} - \overline{Q_S})^2} \quad (5\text{-}31)$$

Where:

- $Q_{i,O}$ and $Q_{i,S}$ are the observed and simulated i^{th} value respectively.
- $\overline{Q_O}$ and $\overline{Q_S}$ are the mean of the observed and simulated values respectively.

A high R^2 value indicates greater model performance.

Nash-Sutcliffe model efficiency (NSE)(Nash & Sutcliffe, 1970)**:**

$$NSE = 1 - \frac{\sum_{i=1}^{n} (Q_{i,O} - Q_{i,S})^2}{\sum_{i=1}^{n} (Q_{i,O} - \overline{Q_O})^2} \quad (5\text{-}32)$$

In this case, a bigger NSE value indicates a better model performance.

Kling-Gupta Efficiency (KGE)(Gupta et al., 2009)**:**

$$KGE = 1 - \sqrt{(R-1)^2 + (\alpha-1)^2 + (\beta-1)^2} \quad (5\text{-}33)$$

Where:

- $R =$ is the correlation coefficient between the observed and simulated data
- $\alpha = \frac{\sigma_s}{\sigma_m}$ with σ_s and σ_m being the standard deviation of the simulated and observed data respectively
- $\beta = \frac{\mu_s}{\mu_m}$ with μ_s and μ_m being the mean of the simulated and observed data respectively

With KGE = 1 reflecting the best model performance, a large value of KGE suggests that simulations can better match data.

6 Results and discussions

6.1 Household demands

This chapter presents the results obtained for Model H1 and Model H2 to link household demands. The results are divided according to the physical link used to relate the water demand to the discharges in the sewer system, either node to node (LNN) or through the households (LTH). In order to have an estimation of the model performance, the values obtained for both models are compared to the values obtained when running the original sewer model developed by Kragset (2019). These values are used as no measured data is provided. It must be considered that they are the result of a model, and therefore will have its own uncertainties and errors. From now one the outputs from the initial sewer model will be identified as "initial sewer model"

6.1.1 Physical Link: Node to node (LNN)

To begin with, the physical link was determined. A buffer area around the sewer nodes of 5m was created in order to select the drinking water nodes that contributed to the discharge in the sewer system. For the model under study only 14 nodes from the drinking water model fulfilled the criteria. Then, the nearest sewer node to each of these 14 drinking water nodes was selected. Initially, no sewer node had two or more drinking water nodes contributing to the discharge. Both drinking water and sewer network are presented in Figure 6-1. The drinking water nodes contributing to the sewer network are shown as blue diamonds in Figure 6-1.

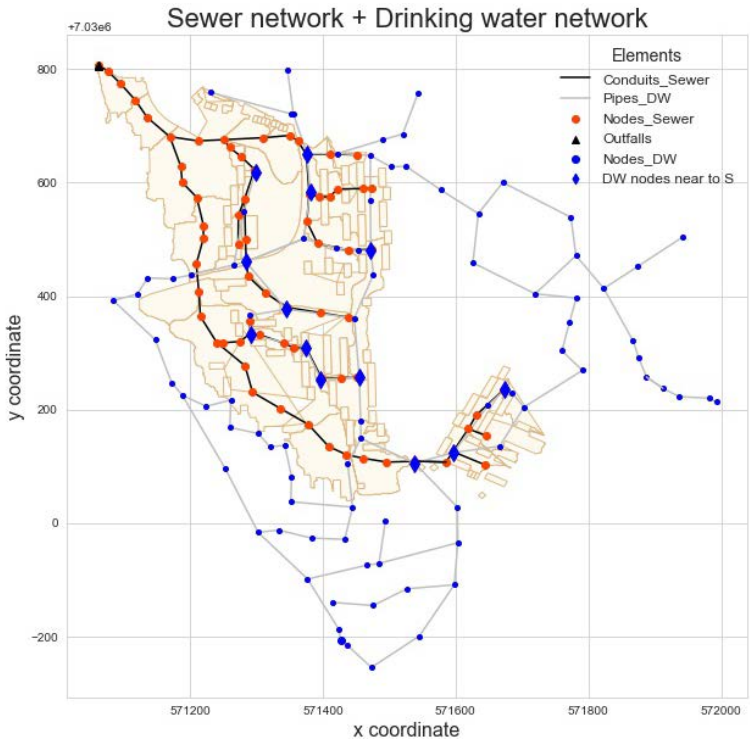


Figure 6-1: Drinking water system and sewer system

6.1.1.1 Results Model 1: Direct to sewer, using LNN

For this model the household demands were inserted directly into the sewer, following the procedure stated in Section 5.2.1. The sewer nodes that receive inflow from the water distribution network are illustrated in yellow in Figure 6-2. The demands added into the sewer are the demands generated by the households, without the background leakages. The background leakages were extracted from the original water distribution network according to the procedure described in Section 5.3.3.

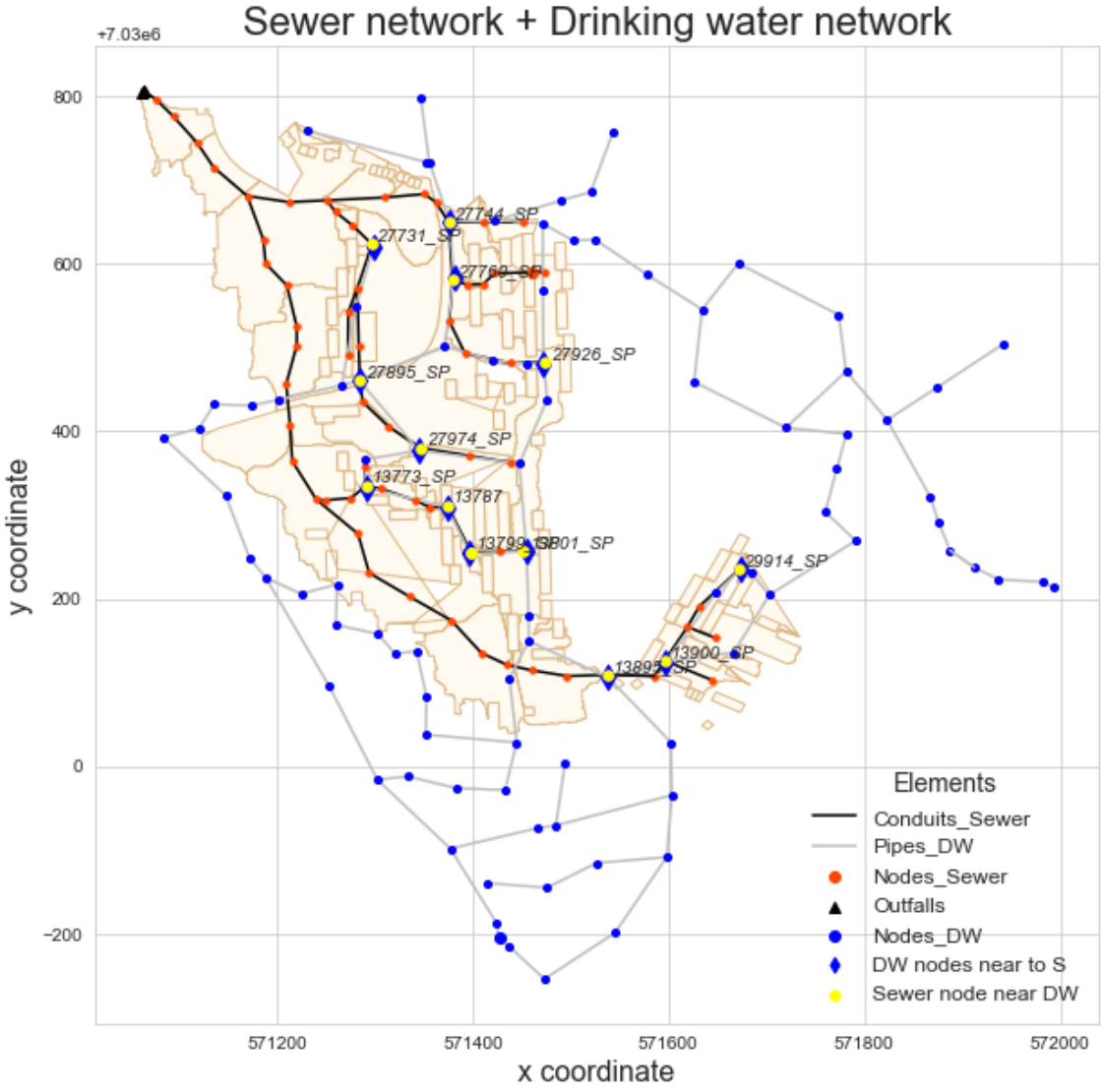


Figure 6-2: Sewer nodes with inflows from drinking water system

The flows obtained using Model 1 are compared to the values obtained when running the original sewer model developed by Kragset (2019). It can be observed that the patterns for both simulated and original series follow the same pattern but in most of the cases the values differ, being the simulated generally higher. Additionally, the water peaks and valleys occur before the sewer peaks.

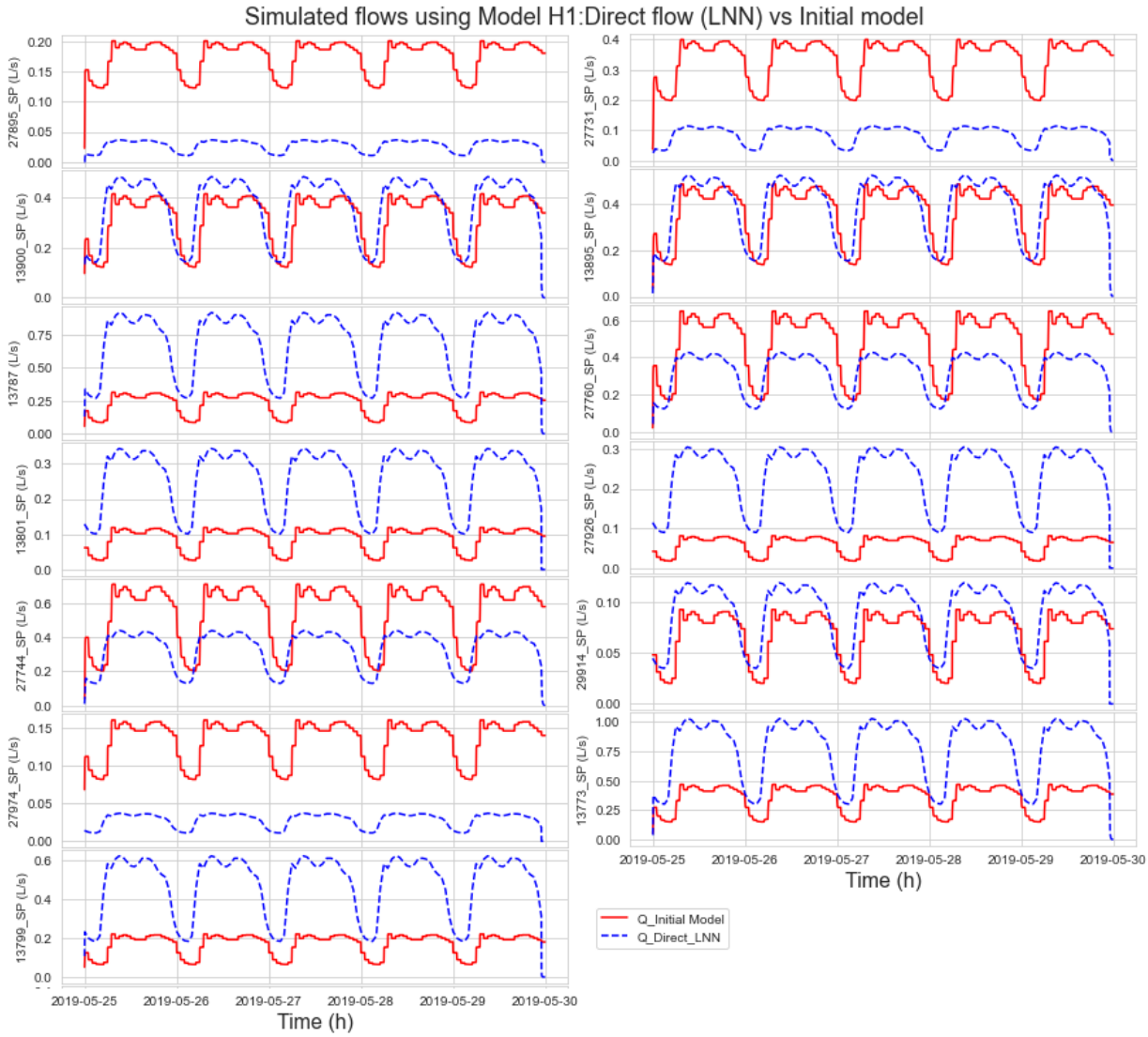


Figure 6-3: Simulated flows in nodes for household demand using Model 1 (LNN) compared to the initial sewer model flows.

Furthermore, several model performance evaluators were calculated. The value obtained for each sewer node with drinking water inflow is presented in Table 6-1, together with the mean value for the whole system considering said nodes. The values obtained for the different performance evaluators show that the simulated values are far from the ones in the original model. One source of error could be that this thesis is working with models and not actual measurements from the drinking water system and the sewer system. Therefore, the error is being propagated throughout the modelling increasing the uncertainties. The differences between the models are further discussed in Section 6.1.3.

However, the simulated values follow the same form as the ones in the initial model, which can be related to relatively good results of R^2 (indicator for the shape). They have the peaks and minimum values almost at the same time, having a delay on the initial model values. Furthermore, the simulated values are generally higher than the initial model ones, evidencing the need for a dampening factor between the demand and the discharge.

The main takeaway is that the results suggest an existing relationship between the household demands and the discharges into the sewer system. Furthermore, they imply the need for a delay factor and a transfer factor between them. These two parameters are discussed in the next model.

Table 6-1: Result of model performance evaluators for the studied nodes, using Model 1 for Household demands with LNN spatial link.

SewerNode	MAPE	R^2	NSE	KGE
27895_SP	1.001211	0.550698	-28.780508	-0.095988
13900_SP	1.001067	0.538523	0.063931	0.620827
13787	1.00134	0.539411	-38.411933	-1.763036
13801_SP	1.00135	0.536609	-32.37707	-1.615963
27744_SP	1.001111	0.555544	-0.895138	0.412606
27974_SP	1.001202	0.530174	-14.944241	-0.06042
13799_SP	1.001328	0.537807	-36.299184	-1.660496
27731_SP	1.001192	0.547766	-11.941361	0.023841
13895_SP	1.001041	0.547625	0.356494	0.707071
27760_SP	1.0011	0.548351	-0.501706	0.456861
27926_SP	1.0014	0.536608	-69.208198	-2.834534
29914_SP	1.001142	0.536609	-0.436263	0.537412
13773_SP	1.001264	0.545998	-16.684333	-0.816734

6.1.1.2 Results Model 2: Transfer factor, using LNN

The second studied model implied the calibration of a transfer factor between the water demand and the sewer discharge. The function to optimize in order to calibrate the transfer factors depends on the observed and simulated flows in pipes and depth in manholes (see Equation (5-2)). According to what it is proposed in Zhang et al. (2021), their method suggest that the flows and depths observed are obtained from sensors installed in the sewer pipes. However, for this master thesis no real measured data was provided. Therefore, the calibration was done by assuming that the flows and depths observed were the ones obtained after running a simulation of the initial sewer model supplied, with the given external flows defined (as specified in Section 4.2).

As mentioned in Section 5.2.2 to optimize the target function, an evolutionary algorithm is to be used. However, for this thesis it took long time to run the simulations with few generations and number of iterations. Therefore, no reasonable values were obtained for the transfer factor. Not finding the optimal solution for the transfer factors could have been due to the calibration being done based on data from a model which has uncertainties on its own and is not accurate enough to perform a calibration.

For this reason, a sensibility analysis was done using Monte Carlo simulations. The distribution chosen for the transfer factor was a uniform distribution with the values limited between 0.7 and 1.0 based on Zhang et al. (2021). The Monte Carlo simulations were used to

derive the mean value, 95th and 5th percentile of the flows obtained in the sewer when the water demand was multiplied by the transfer factor.

Figure 6-4 shows the comparison between the initial model flows in nodes (meaning the flows obtained from the original sewer system), the flows from Model 1, and the results for the mean average flow derived from the Monte Carlo simulations incorporating the transfer factor. Results show that the flows obtained from Model 2 are closer to the initial model values.

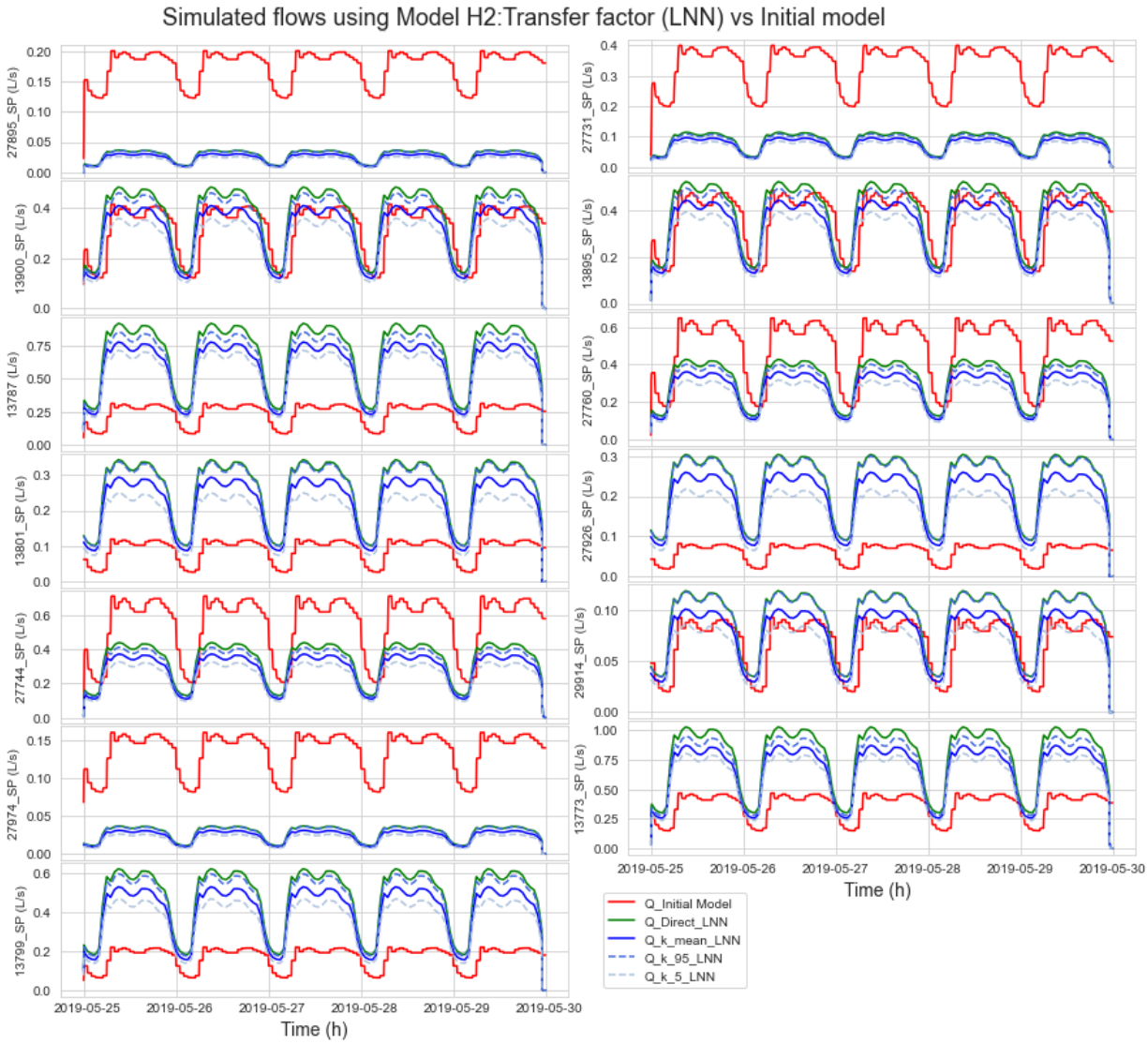


Figure 6-4: Monte Carlo simulations. "Q_Initial Model" indicates the series from the original sewer model. "Q_Direct_LNN" indicates the series of household discharge obtained by the first model. "Q_k_mean_LNN", "Q_k_95_LNN" and "Q_k_5_LNN" represent the mean, 95th percentile and 5th percentile from the flow values derived by Monte Carlo simulations, all using the link node to node spatial connection (LNN)

Additionally, model performance evaluators where calculated and are shown in Table 6-2. As well as in the previous model, the results indicate a low level of accuracy in the predictions. Reasons for the differences are further discussed in Section 6.1.3.

Table 6-2: Result of model performance evaluators for the studied nodes, using the Model 2 for Household demands with LNN spatial link.

SewerNode	MAPE	R^2	NSE	KGE
27895_SP	1.001215	0.552022	-30.589384	-0.147477
13900_SP	1.00102	0.539105	0.442911	0.730868
13787	1.001305	0.539983	-22.615373	-1.128638
13801_SP	1.001318	0.536609	-19.760407	-1.04351
27744_SP	1.001133	0.55747	-1.711742	0.288507
27974_SP	1.001208	0.530174	-16.283503	-0.118466
13799_SP	1.001293	0.538075	-21.498621	-1.061589
27731_SP	1.0012	0.549616	-13.373512	-0.045045
13895_SP	1.001027	0.54906	0.429847	0.720171
27760_SP	1.001124	0.549589	-1.244127	0.327247
27926_SP	1.001371	0.536608	-44.416388	-2.066619
29914_SP	1.001081	0.536609	0.348676	0.707583
13773_SP	1.00122	0.547289	-8.553194	-0.33787

Looking at the results obtained from the two models proposed to link the water household demands to the sewer household discharges, several conclusions can be drawn. First, for most of the nodes the water demands are higher than the sewer flows. This could be explained by an inaccurate distribution of the demands in the water distribution nodes or when assigning them to the sewer nodes. Additionally, in order to get a more representative model, a delay factor is needed as the simulated sewer flow peaks occur before the observed ones. Then, looking at the confidence intervals derived by the Monte Carlo simulations, two situations can be identified. On the one hand, if the observed values fall outside the confidence interval, there is a need to adjust the person equivalents that were assigned to the drinking water node which is generating the discharge into sewer. On the other hand, if the observed series falls inside the confidence interval, the water demand can be adjusted with the transfer factor in order to represent the sewer discharges accurately.

6.1.2 Physical Link: Through households (LTH)

Based on the results obtained for the node-to-node physical link, a second method was proposed in order to improve the outcome of the model. The method is described in Section 5.1.2, and the results of the two models for transforming water demand into sewer discharges using this physical link are presented in the following sections (see Section 0 and 6.1.2.2).

For this link, first each household centroids were determined using QGIS. Then, each unit was associated with its nearest water node. The demand for the household was calculated as the water node´s demand divided by the number of houses associated to it. Furthermore, the demand is transformed into discharge to the nearest sewer node for each unit, accumulating the demands if more than one unit discharged to the same sewer node.

This method aimed to have a better distribution of the demands into the sewer system, by associating the demands through the households. Additionally, it allowed to represent those demands that could have been left out because of the water nodes being far from the sewer nodes, but that still have an influence due to the households being located near both.

Figure 6-5 shows the centroids for each household and the relative location to both the water and sewer systems.

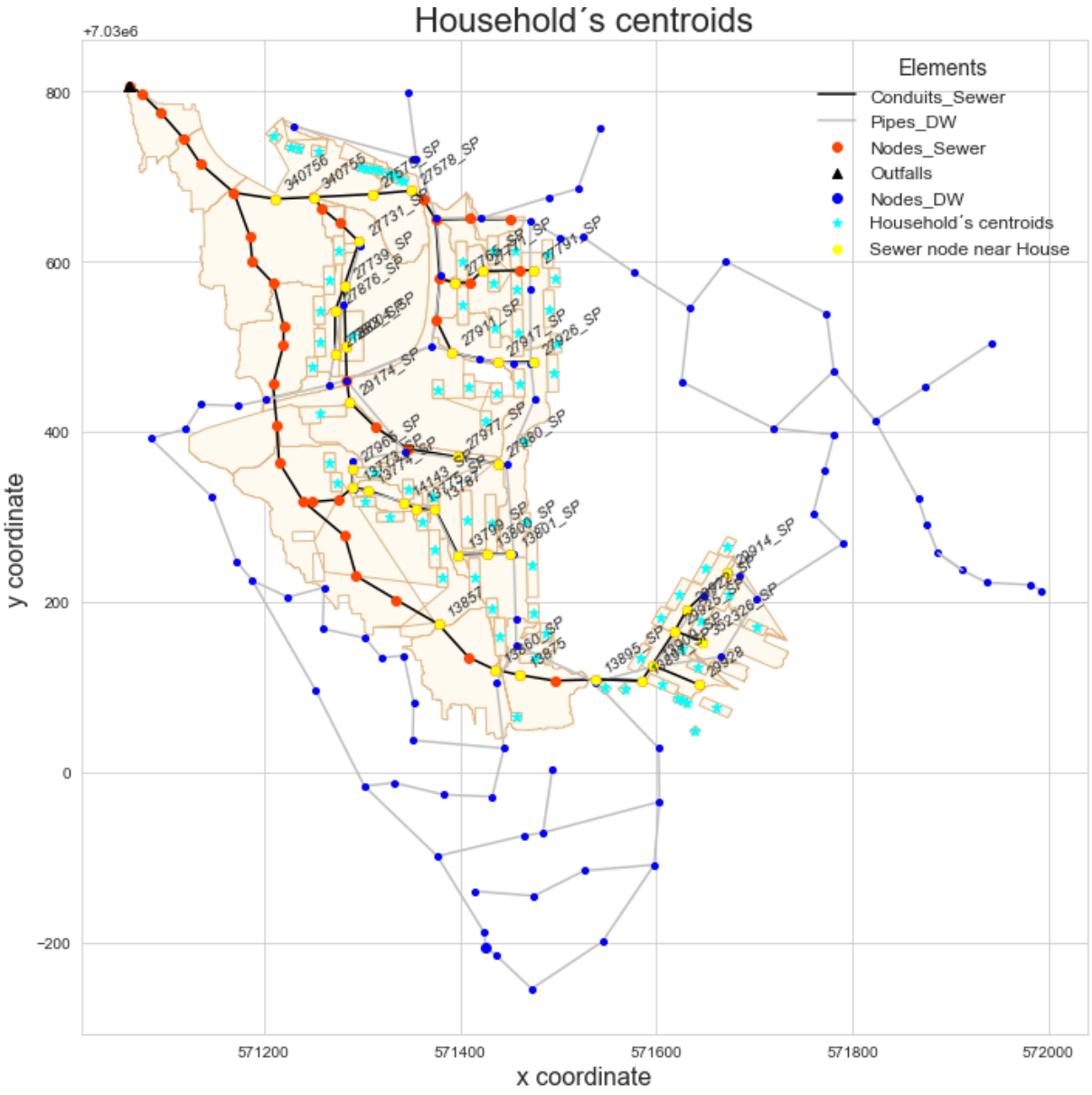


Figure 6-5: Households centroids and relative location to drinking water and sewer network.

6.1.2.1 Results Model 1: Direct to sewer, using LTH

This section presents the results of placing the water demands directly into the sewer system, when both systems are linked through the LTH (link through household) method. Figure 6-6 shows the initial model flow in the sewer system and the simulated flow. The results for all the nodes with inflow with this method can be found in Appendix D: Outflow results for LTH method for both model H1 and H2. It can be observed that for some nodes the simulated flow still has a higher value than the one in the original model. However, for some nodes there is a better approximation of the flows. It can also be seen that in some nodes an inflow was assigned where there was none in the original model. This could mean that probably the discharge associated to those households is allocated downstream the studied node.

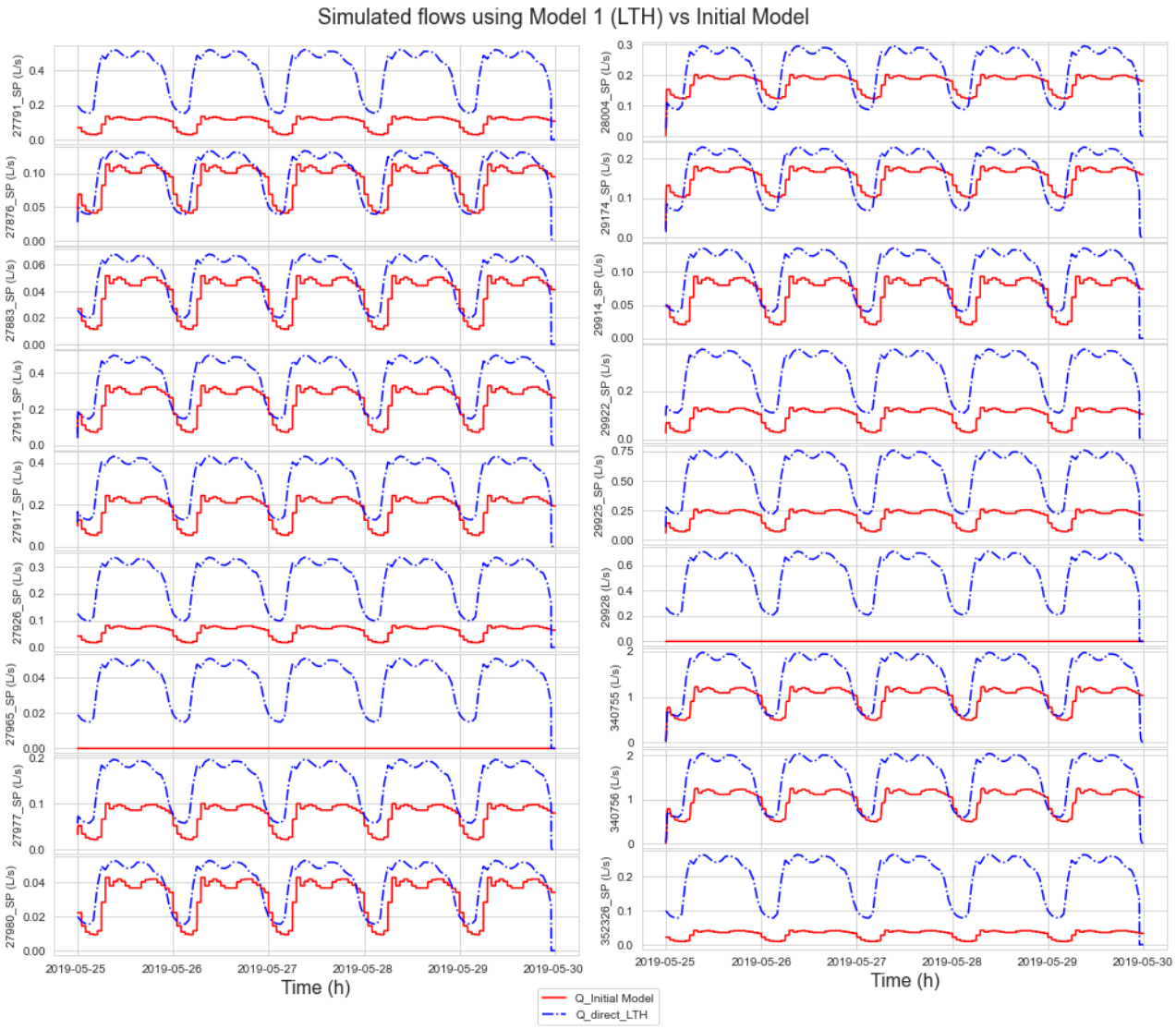


Figure 6-6: Simulated flows in nodes for household demand using Model 1 (LTH) compared to the flows in the Initial sewer model.

Additionally, the performance evaluators were calculated and can be found in of Figure 11-10 in Appendix E: Performance evaluators for LTH model. Similarly to what was obtained previously, they show poor results. However, it is possible to see a relationship between the water demands and the discharges into the sewer. As well as for the previous physical link model, it can be implied that the curve obtained in the original model and in the simulated model follow the same pattern, and that both a transfer factor (to adjust the value of the flows) and delay factor are needed to get a better approximation of the flows.

6.1.2.2 Results Model 2: Transfer factor, using LTH

This method required the calibration of the transfer factor. Due to the calibration taking a long time to run and not throwing good results (see Section 6.1.1.2), a Monte Carlo analysis was done to study the influence of the transfer factor on the discharges. The distribution chosen for the transfer factor was a uniform distribution with the values limited between 0.7 and 1.0 based on Zhang et al. (2021).

Figure 6-7 present the original model flow, the flow obtained from Model 1 (direct demands to sewer) together with the mean value for the flow with transfer factor and the 95^{th} and 5^{th} percentile obtained from the Monte Carlo simulations. The results for all the nodes with inflow with this method can be found in Appendix D: Outflow results for LTH method for both model H1 and H2. It can be seen that this model gives a better approximation of the flows as observed also in Section 6.1.1.2. However, there is still a difference in some of the nodes. As mentioned previously, it can be due to an inaccurate distribution of the Person Equivalents in the water model. It could also be due to an under estimation of the flows in the original sewer model, as it is also based on assumptions on the flow and location of the discharges (further discussion is presented in Section 6.2.4).

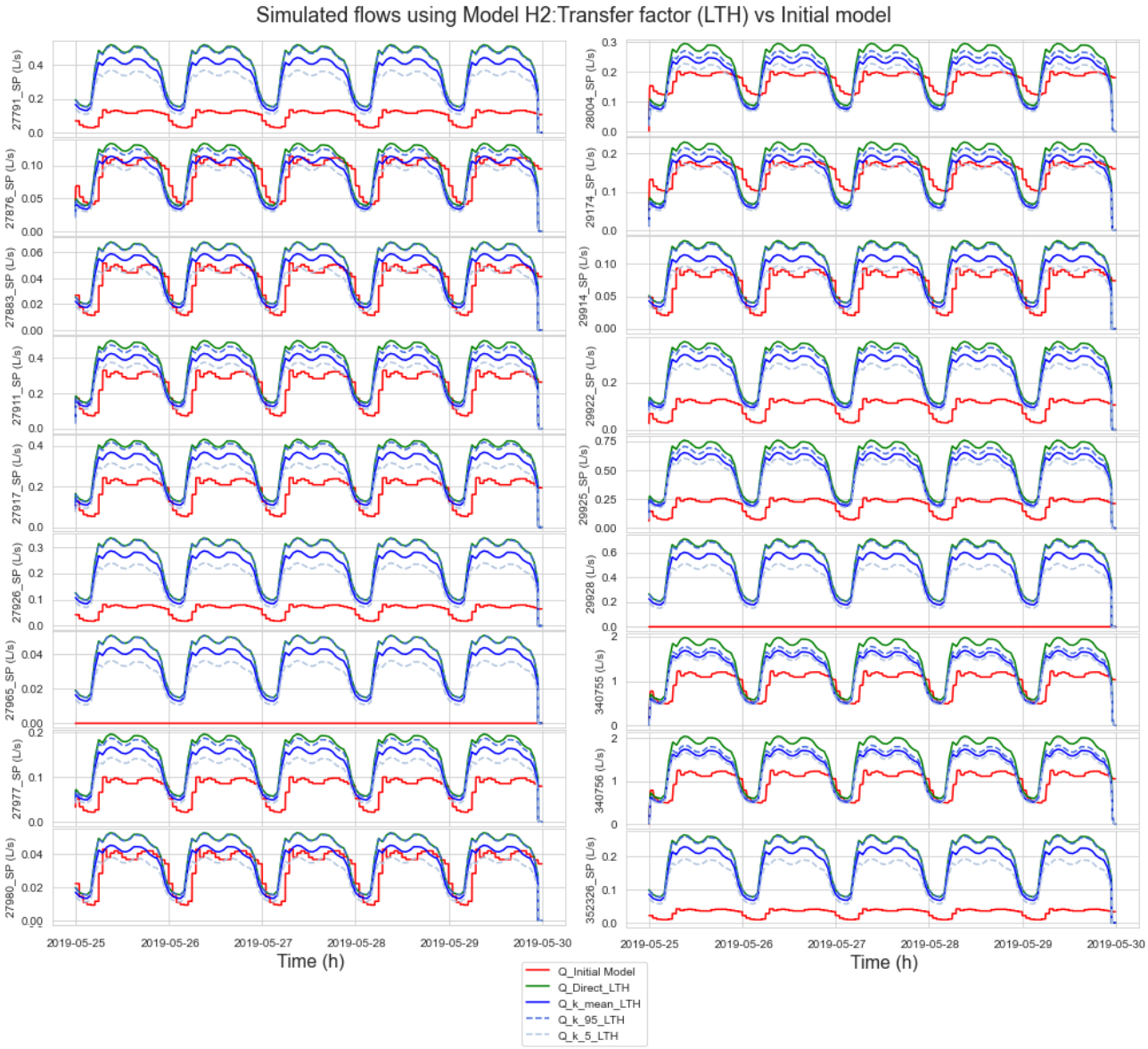


Figure 6-7: Monte Carlo simulations. "Q_Observed" indicates the series from the original sewer model. "Q_Direct_LTH" indicates the series of household discharge obtained by Model 1. "Q_k_mean_LTH", "Q_k_95_LTH" and "Q_k_5_LTH" represent the mean, 95th percentile and 5th percentile from the flow values derived by Monte Carlo simulations, all using the link through household spatial connection (LTH).

6.1.3 Analysis of differences between initial model and simulated models

This section seeks to understand the differences between the values observed for the initial model provided and the simulated models obtained from transforming the water demand into sewer discharges. It must be noted that obtaining a perfect matching model is not part of the aim of this thesis, being the goal to establish the interaction between the water and sewer system and suggest possible ways to model those interactions. However, for the sake of completeness the differences between the results are analyzed.

To evaluate the general performance of the models, the flows obtained at one of the nodes near the discharge are presented in Figure 6-8. The plotted timeseries correspond to the results obtained for both ways to physically link the networks, using Model H2 where a transfer factor is used to transform the water demands into sewer discharges. The initial thought was that connecting the demands through the households and then into the sewer system (see Section 5.1.2) would give better results than linking node to node (method LNN, Section 5.1.1). This was because it is believed to represent more accurately the real behavior, where houses take the water from the nearby drinking water pipe and then discharge it to the closest sewer pipe. Additionally, this method incorporated demands that could have left aside in the LNN method, as in the latter only the nodes near to the sewer contributed to the discharge. However, results show that for the LTH link the flows are much higher than the given in the initial model provided.

Reasons for the gap in flows could be because of inaccuracies in the allocation of the demands or inflows either on the water model or in the sewer model provided. It must be considered that the water demands are taken from a model and no real measurements, and that the results are being compared to a model from the sewer system. Therefore, they both include their own errors and uncertainties. To compare the models, the sum of the average daily consumption was calculated. For the initial model, it was calculated as the sum of the average daily consumption of all the nodes that had inflows in dry period (see Section 4.2). For the simulated model, it was calculated from the demands of all the nodes that form part of a household demand and therefore discharge into the sewer. For the simulated LTH model, the result was much higher than for the initial model. This could be due to the fact that the initial model, even though it was calibrated against measured data, it mainly consisted of adjusting the peaks and minimum flow. The quantity was maintained constant for the calibration being assumed to be 140L/PE/day based on previous studies (see Section 4.2 and Kragset (2019) for more information on the calibration process).

Therefore, it is relevant to compare the models developed with measure data to determine their accuracy. As no data was available, the initial models will be continued to be used during the thesis.

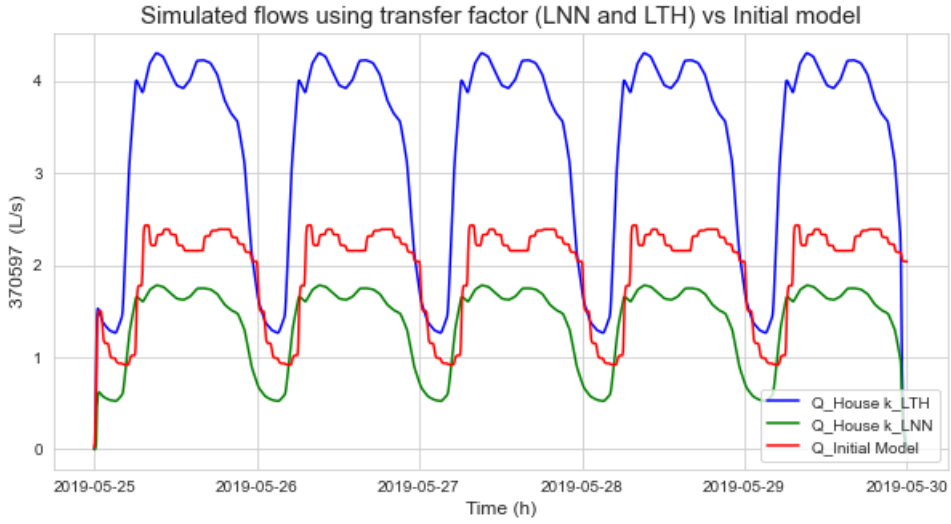


Figure 6-8: Comparison of simulated flows using transfer factor for both LNN and LTH physical link, with the flows obtained in the initial model for one of the nodes near the discharge.

Furthermore, in case both water and sewer model provided are accurate, which was the transfer factor needed in order to get more accurate results. Figure 6-9 shows the results assuming a constant transfer factor of 0.4 for all nodes. It can be observed that the initial model and the simulated one coincide more in flow values. However, literature suggests that the transfer factor is around 0.7 and 1.0, being 0.4 too low. Further studies will be needed to accurately define the transfer factor if better approximations of the sewer discharges from water demands are required.

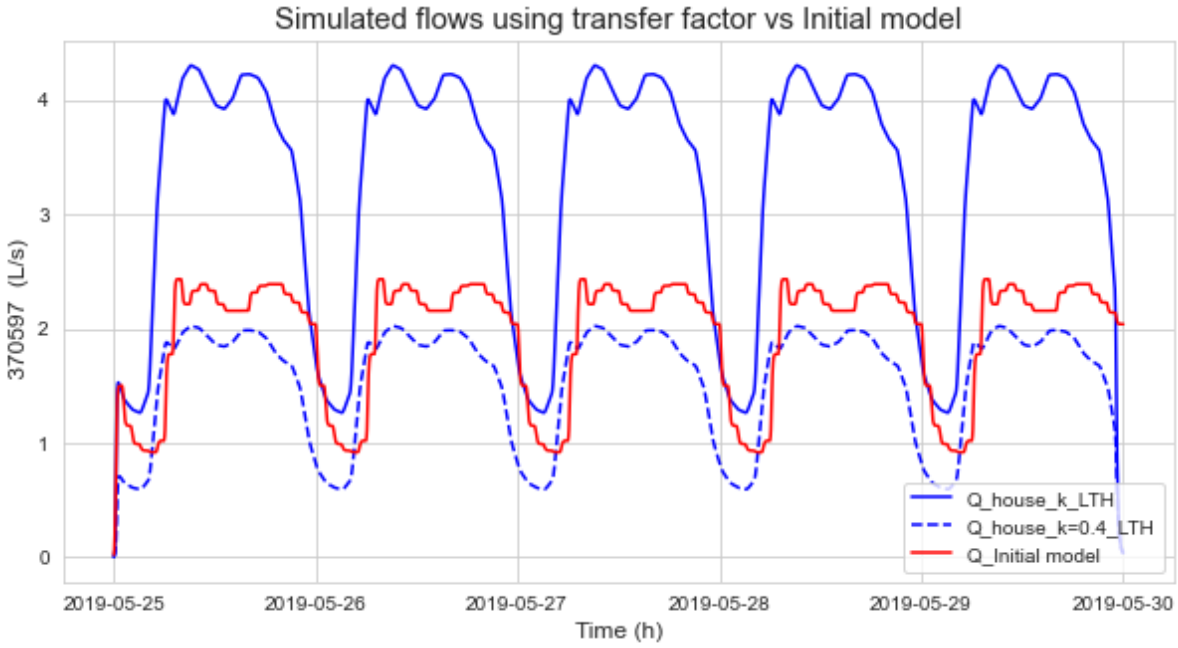


Figure 6-9: Comparison of simulated flows using model LTH, with the transfer factor obtained in Section 6.1.1.2 and changing the transfer factor to k=0.4, with the flows obtained in the initial model for one of the nodes near the discharge.

Additionally, it can be seen that when studying the flows node by node, the results get better when changing the transfer factor to 0.4 as it can be seen in Figure 6-10. The results for all the nodes with inflow with this method can be found in Appendix D: Outflow results for LTH method for both model H1 and H2.

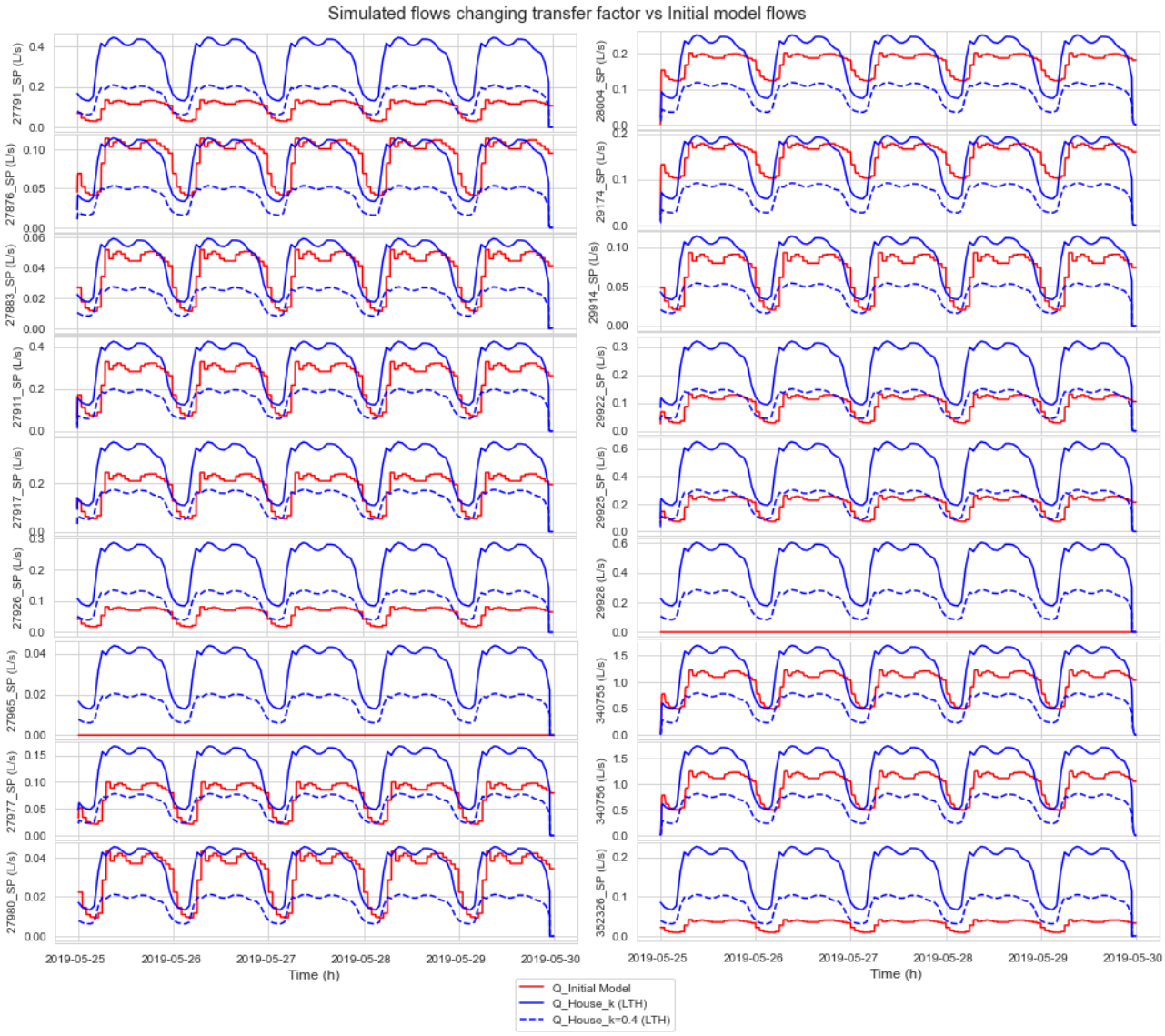


Figure 6-10: Comparison of flows using model H2 using transfer factor k=0.4 and the mean k factor obtained in section 6.1.2.2, with link method LTH and the flows in the initial model.

6.2 Out-Leaks to In-leaks

This chapter aims to present the results obtained when modelling the out-leaks from the water distribution network to in-leaks in the sewer network. Two models were used: L1, where the out-leak is taken directly to the sewer; and L2 where the water leak is transported through soil into the sewer system. The methodology used for this method was described in Sections 0 and 5.3.5.

6.2.1 Spatial link

To begin with, the spatial link was established in order to define where the outflows from the water pipe will be related to the inflow in the sewer system. As stated in Section 5.3.1 infiltration will occur if the water and sewer pipe are located close enough for the water to reach the sewer. The water pipes were represented by a fictional leaking node, which is located in the middle point of each pipe. Then, a buffer area around the sewer pipe is created and those fictional nodes that are included in it will have the potential to leak into the sewer. Figure 6-11 shows the fictional nodes that will have an influence on the sewer flow. It can be seen that they represent those water pipes that are almost in the same location than their adjacent sewer pipe.

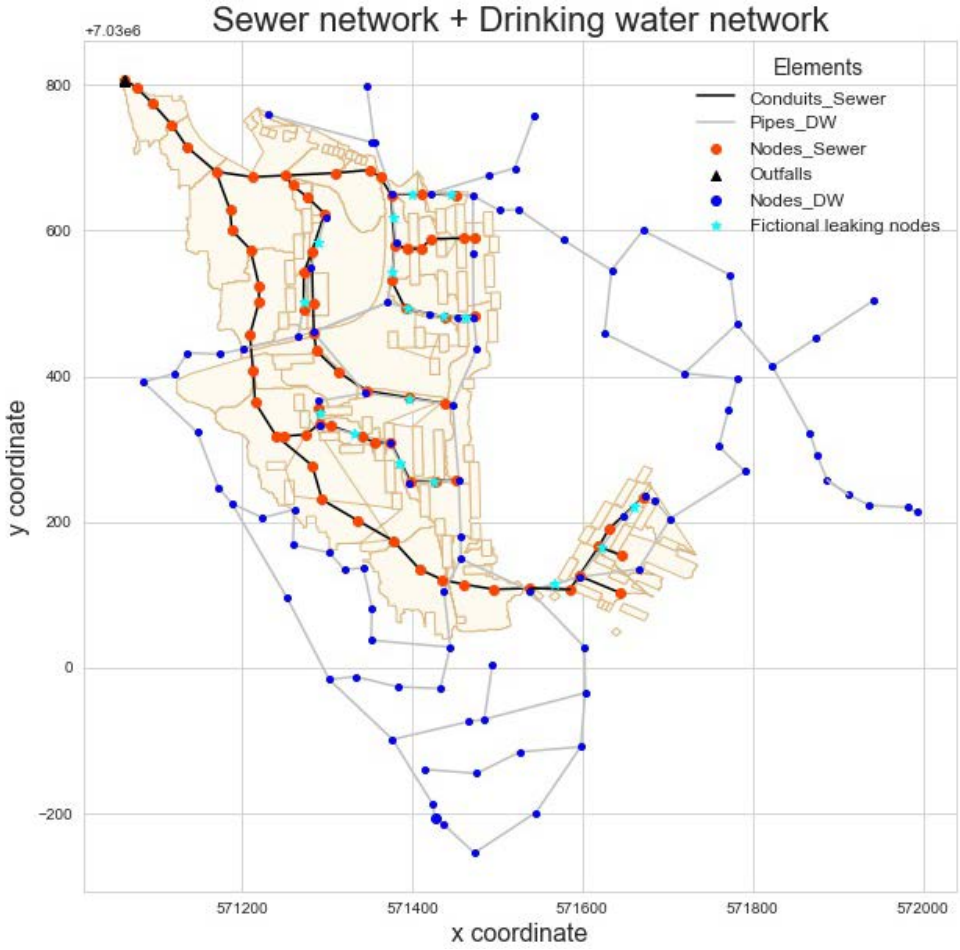


Figure 6-11: Fictional leaking nodes near sewer system.

6.2.2 Determine out-flows

This section presents the results of estimating the flow that will leak from the water distribution pipe for both pipe bursts and background leakages.

6.2.2.1 Pipe burst outflow

In order to get the outflow from a pipe burst according to the method described in Section 5.3.2, information regarding the pipe´s material is needed, such as the young modulus and the thickness of the pipe together with an estimation of the crack size is required. In this case, the values used are the ones presented by Fuchs-Hanusch et al. (2016) and are presented in Table 5-2, Table 5-3 and Table 5-4. As described in Section 5.3.2, a fictional node was created in the middle of each pipe where an emitter was placed in order to model the burst. By iterating over the emitter coefficient and emitter exponent the potential outflow due to a pipe burst for each pipe was determined.

The outflow from the water leak depends on the type of crack, which is linked to the pipe material. Several studies (WSAA (2003), Sorge (2007) and Friedl (2012)), have shown that Cast Iron can have any type or defect, but plastic pipes would most likely have a longitudinal crack. The outflow was calculated for every pipe for both types of crack and are presented in Figure 6-12 and Figure 6-13 (longitudinal and circular crack respectively). They are the result of the mean value obtained after doing a Monte Carlo analysis varying the size of the crack for the different types of cracks. It can be observed that for longitudinal cracks the values are higher, being the highest values in the pipes surrounding the tank. No data was available regarding previous pipe burst in order to determine the reliability of the results obtained. Therefore, they were compared to the results obtained by Fuchs-Hanusch et al. (2016) which are on the same order as the ones obtained for the studied model. By doing this, the aim was to have a general idea if the values were reasonable and according to what was expected.

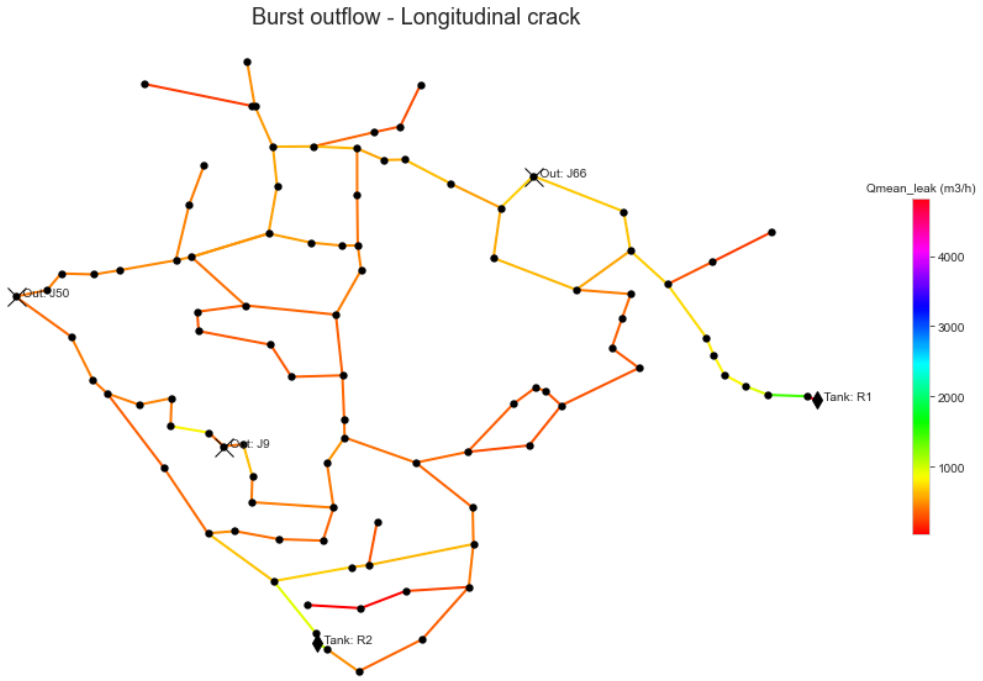


Figure 6-12: Potential burst leak outflow: Longitudinal crack.

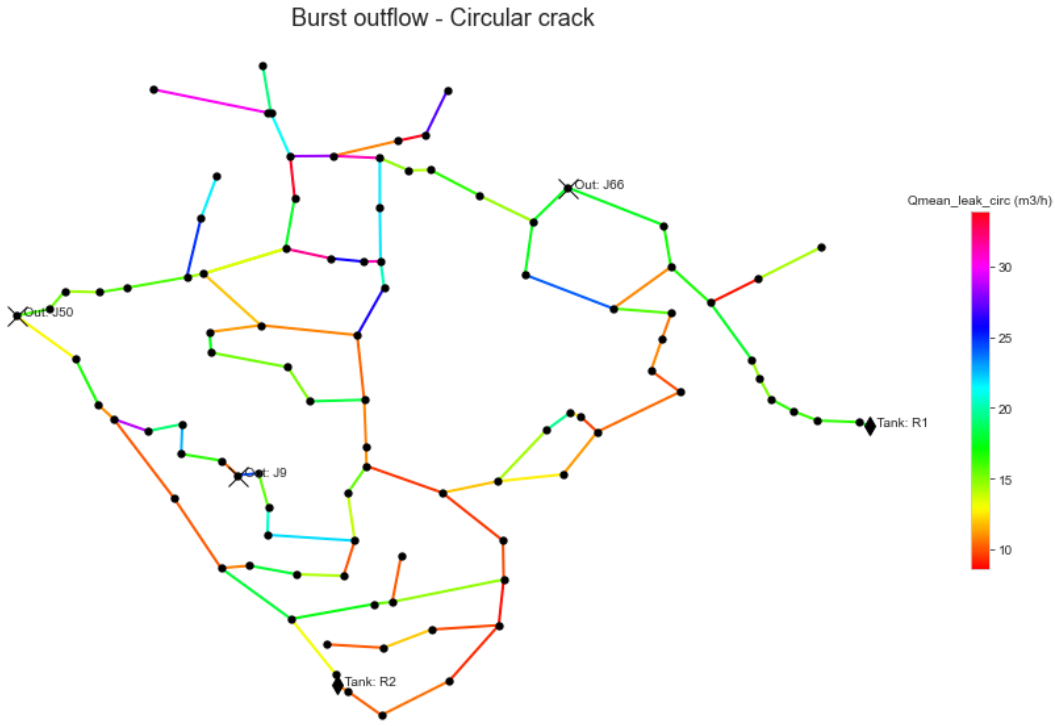


Figure 6-13: Potential burst leak outflow: Circular crack.

6.2.2.2 Background leaks outflow

As it was described in Section 5.3.3 the background leaks were extracted from the water distribution model. The information was stored in the water nodes as constant demand

patterns with an associated base demand according to the node. Then, the leaks outflow was associated to a pipe by assigning the sum of half the background leak for the starting and end node:

$$Q_{background\ leak,\ \ j} = \frac{1}{2} \cdot \left(Q_{bl,\ \ start\ node_j} + Q_{bl,\ \ end\ node_j} \right)$$

Where:

- $Q_{background\ leak,\ \ j}$ = Background leak for pipe j
- $Q_{bl,\ \ start\ node_j}$ = Background leak for the starting node of pipe j
- $Q_{bl,\ \ end\ node_j}$ = Background leak for the end node of pipe j

The results for each pipe are presented in Figure 6-14.

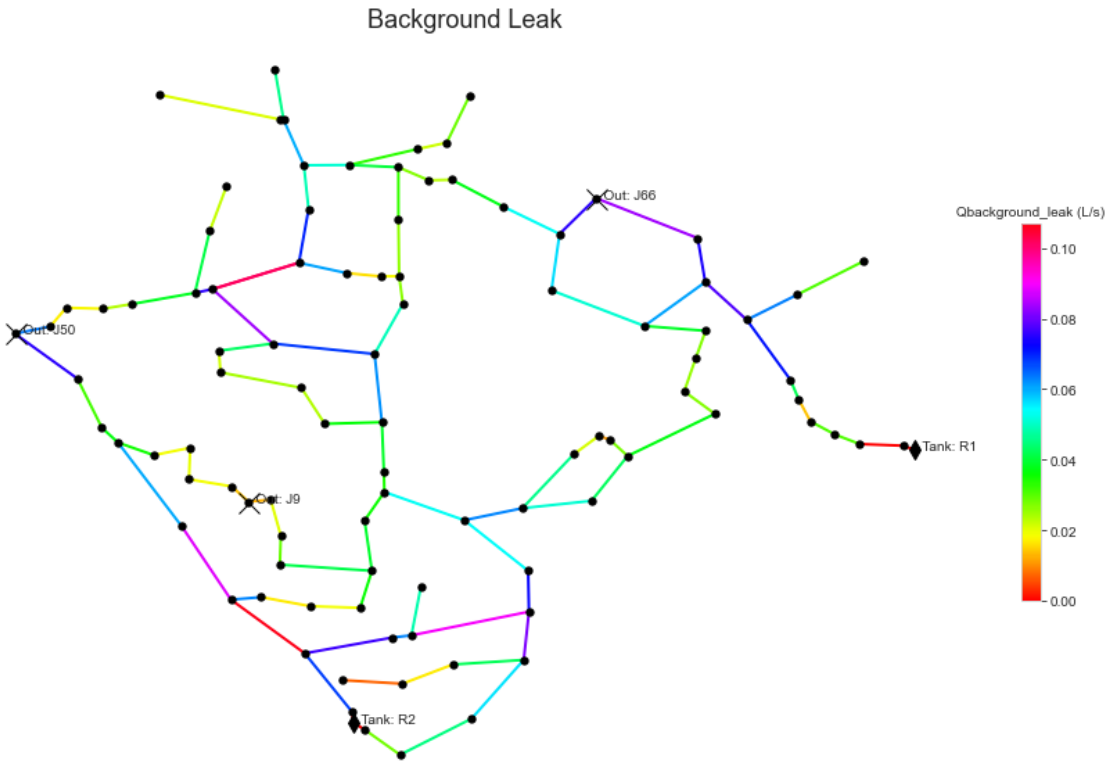


Figure 6-14: Background leaks flow.

6.2.3 Results Model L1: From water leak directly to sewer in-leak

The results obtained for both background leaks and pipe burst, considering that the out flow goes directly into the sewer, are presented in Figure 6-15 and Figure 6-16 respectively. Background leakages are assumed to happen in all pipes, and its flow will depend on the material, age, maintenance, among others. With regard to the pipe burst, it was assumed that it will happen in one pipe at a time. Therefore, a random pipe was selected together with a random starting time and duration before it is repaired. The results are shown for pipe "27911_SP" starting at 2:30 on the 25/5/2019 and ending at 8:30 the same day.

From Figure 6-15 it can be observed that for the background leaks as they are constant in time it will affect the whole timeseries, therefore obtaining an offset timeseries. However, for the pipe burst as they can be identified more easily, once they are detected they are

rapidly repaired. This makes the flow into the sewer have a peak when the burst happens, then maintained flow until it is repaired, where it goes back to the regular flow in the sewer pipe (see Figure 6-16). Due to the high flow coming out from the burst, the sewer will probably exceed its capacity if it were to be transported directly to the sewer. Therefore, it was necessary to implement a certain transport though the soil that will delay the infiltration, and a leak area that will limit the flow into the sewer.

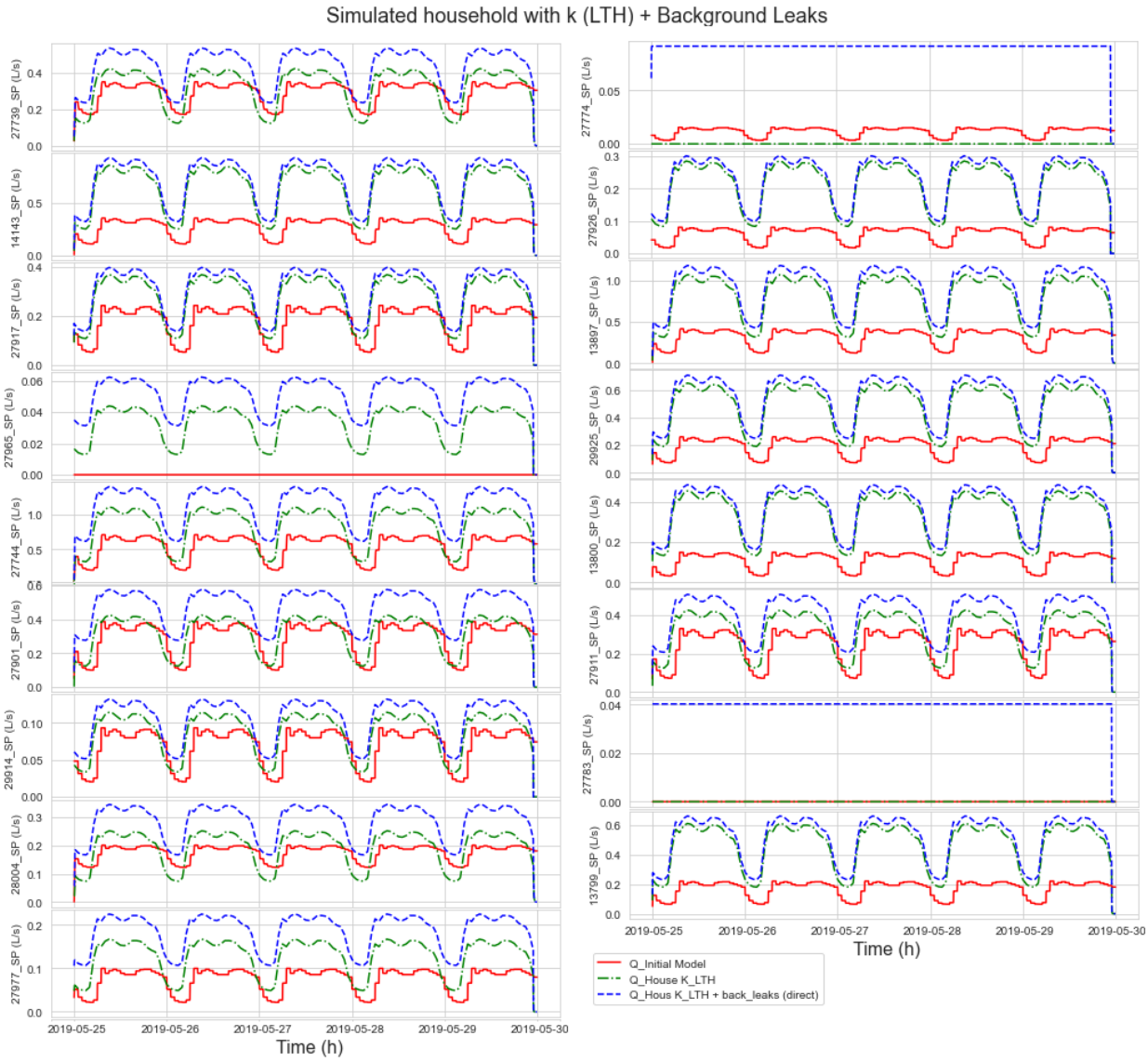


Figure 6-15: Background leakages directly to sewer

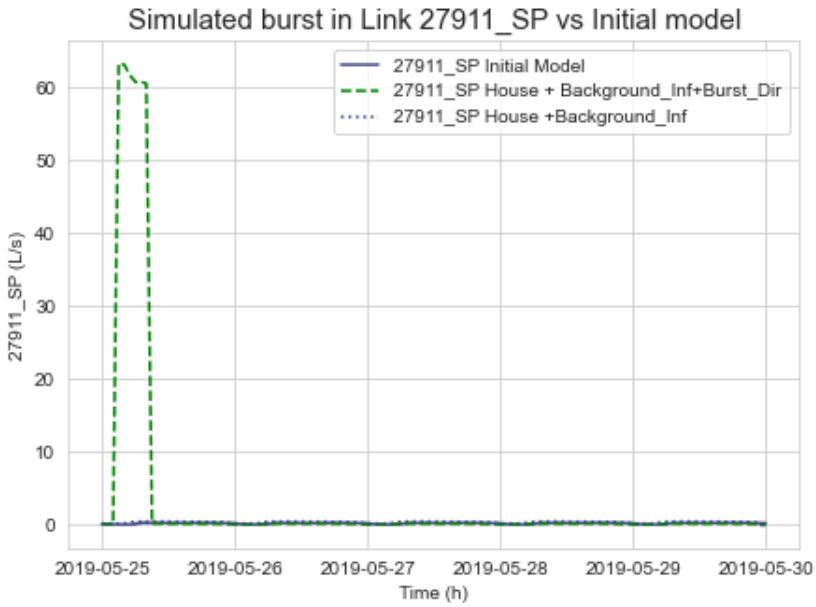


Figure 6-16: Outflow from burst directly to sewer. Simulated burst in Link 27911_SP (randomly chosen).

6.2.4 Results Model L2: Water leak transported through soil and transformed into sewer in-leak

This section presents the results obtained after modelling the out leaks into the sewer according to section 5.3.5. As stated before, this model takes as a basis the formula for infiltration into sewer suggested by Karpf & Krebs (2011) and varies the distance between the water level and the sewer defect according to several assumptions.

For background leakages, it was assumed that they are constant through time and that they have reached the sewer defect. Therefore, the water height above the defect remains constant throughout time. Additionally, the infiltration into sewer is limited by the amount of water leaked. Figure 6-17 shows the observed flows (those that are obtained from the original sewer model), the flows obtained for the household demands model, and the results for background leaks when modelling both directly into sewer or by infiltrating into the ground. It can be seen that the results for both ways of modelling are equal, meaning that the estimated flow considering infiltration is bigger than the actual flow coming out from the pipe. It must be stated that no data is available to verify the assumptions taken to model the infiltration of background leaks. However, the assumptions taken are considered reasonable for having an initial approximation of the process given the information availability.

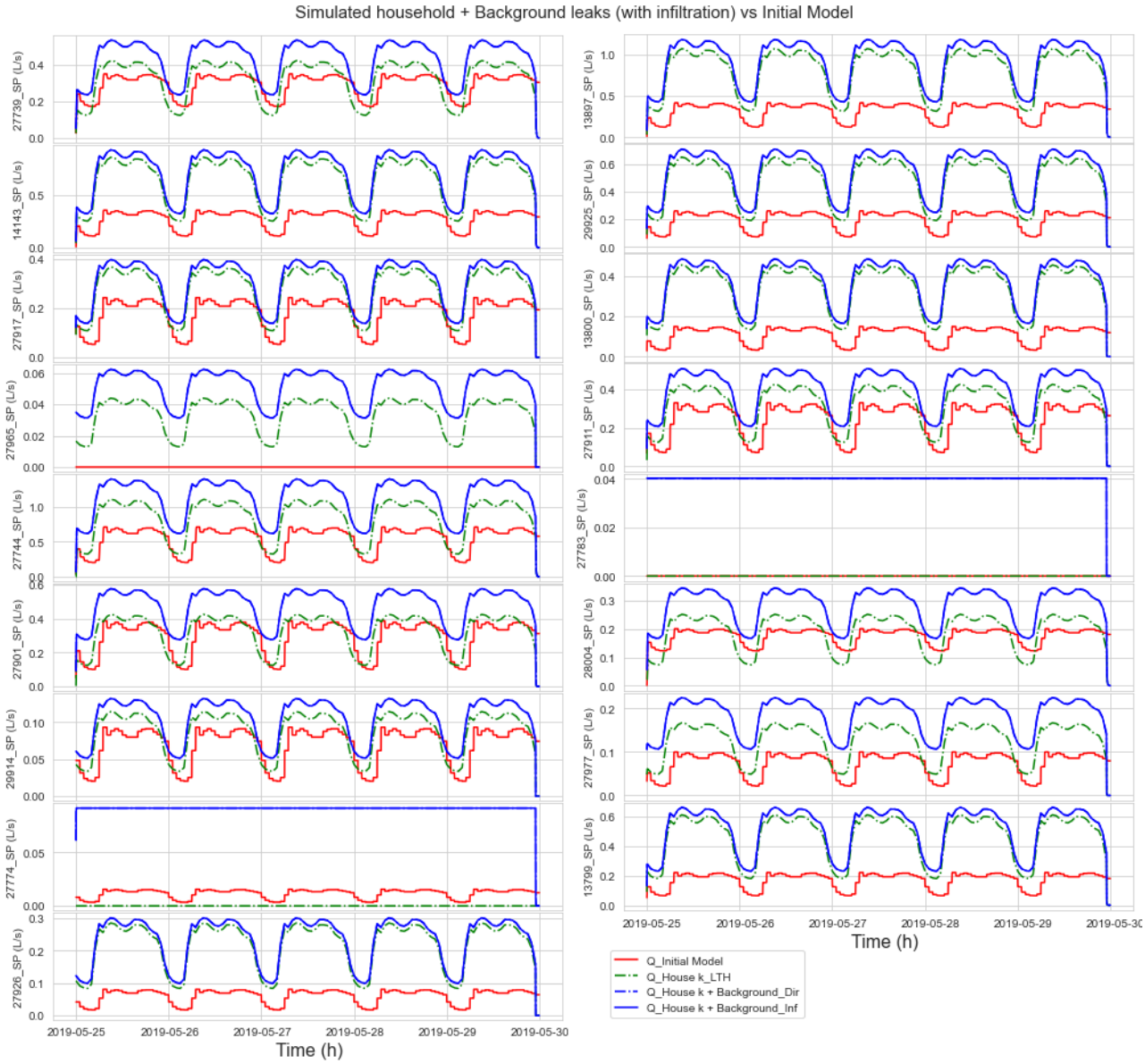


Figure 6-17: Background leakages considering infiltration to the ground.

With regard to infiltration into sewer due to pipe burst, the model also takes as a basis the infiltration formula proposed by Karpf & Krebs (2011). However, in this case the variation of the water level above the sewer defect will depend on the time that has passed since the burst, together with the maximum volume that is available between the water pipe and the sewer. It will also depend on the infiltration velocity of the soil, as well as on the difference between the flow going out of the burst and the in-flow to the sewer (see Section 5.3.5).

Figure 6-18 and Figure 6-19 present both the variation of water level above the sewer defect, and the infiltrated flow to the sewer. It can be seen that both are null until the pipe burst occurs, moment where the value increases until it reaches the maximum volume. This maximum is maintained until the pipe burst is repaired, to then slowly decrease as the water starts to infiltrate into the sewer. Figure 6-20 shows the total flow in the sewer for the whole time period. It can be observed that in this case the inflow from the pipe burst does not have a great influence on the flow.

As well as for the background leaks, no information was available to validate the suggested model. However, it is aimed to take a first step into the design of a model that can estimate in-leaks to the sewer due to leaks from the water distribution network. The assumptions were made based on engineering experience and will need to be further studied to confirm, improve, or deny them.

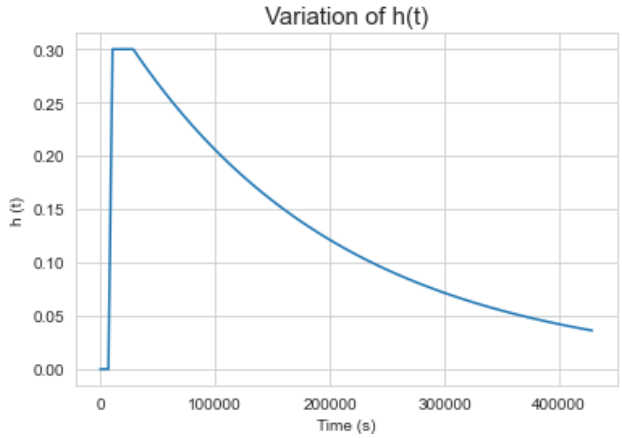


Figure 6-18: Variation of h(t) in the soil package when a pipe burst occurs.

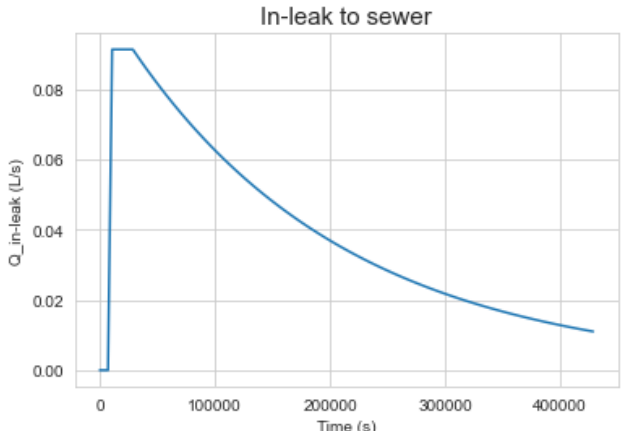


Figure 6-19: In leak to sewer due to pipe burst, for pipe 27911_SP (chosen randomly), considering infiltration into the soil.

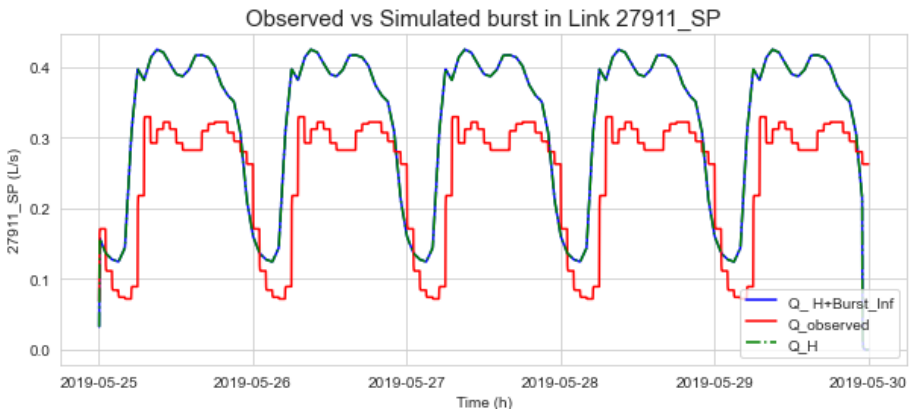


Figure 6-20: Total flow through pipe 27911_SP when pipe burst occurs, considering infiltration through the soil and into the sewer.

7 Limitations

One of the limitations of this thesis is that the data provided for both water distribution system and sewer network are based on models and not real measured data. This could lead to, for example, inaccurate distribution of the households and their demands that makes more difficult to get matching results when comparing the simulated and observed discharges. This could be solved if the information for the smart water meters was available. It would help improve the distribution of the households to get a better idea where they are taking their demand from and where they are discharging. Also, having some measurements in sewer would help with the calibration of the k factor to get a more accurate model.

Additionally, the Model H2 which incorporates a transfer factor between the water demand and the sewer discharge does not include a time lag between them, being the demands transformed instantly into discharges. This has proven to be not accurate as results show a clear delay of the discharge. With regard to model H3, where demand pulses are transformed to discharge pulses, a limitation is the data availability to reconstruct the pulses. This is due to the fact that are based on statistics about the area and consumption patterns.

Another limitation resides in the estimation of burst outflow being based on statistics on burst size as it can happen that in reality burst are bigger and will influence the infiltration into sewer. The background leakages in this case were extracted from the model, however this is not the usual situation when only measurements are available. In that case, a method to determine the background leaks would need to be chosen to then estimate the infiltration to the sewer.

As mentioned before limited data and literature was available that dealt with how to determine infiltration from water pipe to sewer. This implied to develop a model based on assumptions that seemed reasonable and that need to be further investigated to confirm the method. Additionally, no information was available regarding pipe´s burst to compare the results obtained with real data for either calibrate or rectify the proposed method.

Relevant parameters for the development of the leak model were roughly estimated based on literature found for areas with similar characteristics. However, they can highly impact the results and would need further investigation. These parameters are for example the leak area for the sewer inflow which is quite unpredictable and varied for sewer system. Furthermore, the soil surrounding the pipes was assumed to be gravel, under the assumption that it was correctly built. However, this is not always true as many times during construction soils are mixed. Therefore, infiltration test would need to be performed in the area to determine the soil infiltration coefficient.

Furthermore, the leak model does not take into consideration the situation where the sewer is full, and working under pressure, or if it is semi-full how it will influence the amount of flow that will enter. If it is working as a force main, there will be no infiltration into the sewer, and probably there will be exfiltration. Also, it is assumed that the groundwater table is lower than the sewer and will therefore not affect the infiltration.

Another limitation resides on the assumption that the water leaked is presumed to get out of the study system once it has reached the volume defined between the water pipe and the sewer pipe along the width and length of the trench. This water is assumed to further infiltrate into the ground, but it could happen that it is being retained in the trench due to a lower infiltration coefficient of the soil outside the trench. As a result, the trench could be

filled up with water coming from the pipe and the model developed will no longer be accurate. Additionally, it could be transported horizontally to other pipes where it can cause infiltration, and this is not accounted for in the model.

8 Conclusions

When aiming to describe the interactions between the water distribution network and the sewer system, two main links are to be analyzed. First, the water consumed by the users that will eventually discharge into the sewer system either totally or partially. Second, leaks from the water pipes can potentially derive in I/I water into the sewer due to cracks and defects in the later.

This study has acknowledged these interactions and suggested several models to clearly define them. To begin with, two physical links where the outflows from the water distribution system would be placed as inflows into the sewer system were described: LNN (link node to node) and LTH (link through households). The latter representation aligns more closely with real-world scenarios, enhancing the accuracy of the model's outputs.

Further, the demand´s link was established by describing three possible methods. However, only two were developed, leaving the third one open to future work. The studied models include placing the demands directly to the sewer inflows (based on Lund et al. (2021)) and applying a transfer factor in order to represent losses that can occur from the moment the water is consumed until it is discharged (based on Zhang et al. (2021)). When analyzing the results of estimating the household discharges from the water demands for the case study, it was evident that improvement of the different parameters is needed to get a better approximation. However, it was possible to clearly identify a relationship between household demands and discharges, which was part of the aim for this thesis. Furthermore, the key findings highlight that accurately describing the discharges into the sewer from water consumption requires incorporating both a delay and a loss factor to achieve a more precise approximation.

Additionally, a method to describe the in-leaks into the sewer due to out-leaks from the water distribution network was suggested. The method was based on the model proposed by Karpf & Krebs (2011) for infiltration of groundwater into sewer and modified based on engineering assumptions adapt it for infiltration of water pipe leaks. This method aimed to set a basis on how in-leaks to sewer could be modelled from the out-leaks from water pipes. Furthermore, it evidenced the parameters relevant to represent the link. These parameters included the area of the sewer leak as it will limit the infiltration, together with the permeability factor of the soil surrounding the leak. Additionally, it is relevant to define the water movement through the soil once it has leaked from the water pipe as the difference between waterfront and the sewer defect plays a role on estimating the in-leaks. Further work is needed to back up the assumptions taken to describe this interaction, but it can be concluded that is possible to model it with an accurate estimation of the different parameters.

9 Future work

With regard to the household modelling, it is recommended that further work focus on adding a time lag to the demands before discharging into the sewer system. Furthermore, Model H3 (Demand pulses transformed to discharge pulses) suggested for the household demands was not implemented in this master thesis due to lack of information. However, this model seems to be a good tool to try to replicate real data from consumers where no sensor data is available, but statistic information can be collected regarding household composition and particular uses of household artifacts. Future work could implement this model to transform demands into discharge as well as the use of SIMDEUM to build the household demands in this model.

The modeling of leaks relies on assumptions that require additional validation through relevant knowledge to enhance their reliability. The proposed approach introduces a model for water transport along the waterfront through the soil, which demands further exploration. Specifically, the current model assumes a linear downstream transport of the waterfront, with a horizontal spread equivalent to the dimensions of the trench. To improve the model, it is necessary to investigate whether this assumption is accurate. A suggestion could be to adapt the equations for the spread of the waterfront during an infiltration test as presented in Alakayleh et al. (2019). Future work should also include a way to limit the flow coming into the sewer when it is full and represent what happens when the groundwater table is above the sewer level. Moreover, due to the relevance for the infiltration model, it is crucial to enhance the accuracy of estimating both the infiltration coefficient for the soil and the leak area in the sewer.

10 References

Alakayleh, Z., Fang, X., & Clement, T. P. (2019). A comprehensive performance assessment of the modified Philip-Dunne infiltrometer. *Water (Switzerland)*, *11*(9). https://doi.org/10.3390/w11091881

Bailey, O., Arnot, T. C., Blokker, E. J. M., Kapelan, Z., Vreeburg, J., & Hofman, J. A. M. H. (2019). Developing a stochastic sewer model to support sewer design under water conservation measures. *Journal of Hydrology*, *573*, 908–917. https://doi.org/10.1016/j.jhydrol.2019.04.013

Beheshti, M., & Sægrov, S. (2018a). Quantification assessment of extraneous water infiltration and inflow by analysis of the thermal behavior of the sewer network. *Water (Switzerland)*, *10*(8). https://doi.org/10.3390/w10081070

Beheshti, M., & Sægrov, S. (2018b). Sustainability assessment in strategic management of wastewater transport system: a case study in Trondheim, Norway. *Urban Water Journal*, *15*(1), 1–8. https://doi.org/10.1080/1573062X.2017.1363253

Beheshti, M., & Sægrov, S. (2019). Detection of extraneous water ingress into the sewer system using tandem methods-a case study in Trondheim city. *Water Science and Technology*, *79*(2), 231–239. https://doi.org/10.2166/wst.2019.057

Beheshti, M., Sægrov, S., & Ugarelli, R. (2015). Infiltration / Inflow Assessment and Detection in Urban Sewer System. *Vann*, *01*, 24–34.

Berardi, L., Laucelli, D., Ugarelli, R., & Giustolisi, O. (2015). Hydraulic system modelling: Background leakage model calibration in Oppegård municipality. *Procedia Engineering*, *119*(1), 633–642. https://doi.org/10.1016/j.proeng.2015.08.916

Blokker, E. J. M., Agudelo-Vera, C., Moerman, A., Van Thienen, P., & Pieterse-Quirijns, I. (2017). Review of applications for SIMDEUM, a stochastic drinking water demand model with a small temporal and spatial scale. *Drinking Water Engineering and Science*, *10*(1), 1–12. https://doi.org/10.5194/dwes-10-1-2017

Blokker, E. J. M., Vreeburg, ; J H G, & Van Dijk, J. C. (2010). Simulating Residential Water Demand with a Stochastic End-Use Model. *Journal of Water Resources Planning and Management*, *136*(1), 19–26. https://doi.org/10.1061/ASCEWR.1943-5452.0000002

Cassa, A. M., & Van Zyl, J. E. (2013). Predicting the head-leakage slope of cracks in pipes subject to elastic deformations. *Journal of Water Supply: Research and Technology - AQUA*, *62*(4), 214–223. https://doi.org/10.2166/aqua.2013.094

Cassa, A. M., van Zyl, J. E., & Laubscher, R. F. (2010). A numerical investigation into the effect of pressure on holes and cracks in water supply pipes. *Urban Water Journal*, *7*(2), 109–120. https://doi.org/10.1080/15730620903447613

Covelli, C., Cozzolino, L., Cimorelli, L., Morte, R. Della, & Pianese, D. (2015). A model to simulate leakage through joints in water distribution systems. *Water Science and Technology: Water Supply*, *15*(4), 852–863. https://doi.org/10.2166/ws.2015.043

DeSilva, D., Burn, S., Tjandraatmadja, G., Moglia, M., Davis, P., Wolf, L., Held, I., Vollertsen, J., Williams, W., & Hafskjold, L. (2005). Sustainable management of leakage from

wastewater pipelines. *Water Science and Technology*, *52*(12), 189–198. https://doi.org/10.2166/wst.2005.0459

Elías-Maxil, J. A., Van Der Hoek, J. P., Hofman, J., & Rietveld, L. (2014). A bottom-up approach to estimate dry weather flow in minor sewer networks. *Water Science and Technology*, *69*(5), 1059–1066. https://doi.org/10.2166/wst.2014.010

Friedl, F. (2012). *Vergleich von statistischen und physikalischen Modellen zur Berechnung der Auftrittswahrscheinlichkeit von Schadensarten auf Trinkwasserhaupt und Zubringerleitungen (Evaluation of Statistical and Physical Models for Failure Mode Prediction of Transmission Mains)*. Dissertation (PhD). Graz, University of Technology.

Fuchs-Hanusch, D., Steffelbauer, D., Günther, M., & Muschalla, D. (2016). Systematic material and crack type specific pipe burst outflow simulations by means of EPANET2. *Urban Water Journal*, *13*(2), 108–118. https://doi.org/10.1080/1573062X.2014.994006

García, V. J., & Cabrera, E. (2006). *THE MINIMUM NIGHT FLOW METHOD REVISITED*.

Geological survey of Norway. (n.d.). Retrieved April 6, 2023, from https://geo.ngu.no/kart/minkommune/?kommunenr=5001

Gupta, H. V., Kling, H., Yilmaz, K. K., & Martinez, G. F. (2009). Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *Journal of Hydrology*, *377*(1–2), 80–91. https://doi.org/10.1016/j.jhydrol.2009.08.003

Jaurena, M. (2023). *Python Codes - Master Thesis*. https://drive.google.com/drive/folders/1Ie2fZkd6CecNlRL1KDHdYwvToSHCIdW7?usp=sharing

Karpf, C., & Krebs, P. (2011). Quantification of groundwater infiltration and surface water inflows in urban sewer networks based on a multiple model approach. *Water Research*, *45*(10), 3129–3136. https://doi.org/10.1016/j.watres.2011.03.022

Kragset, B. T. (2019). *Modelling of the Interaction Between Storm and Foul Sewers in a Joint Model and Assessment of Infiltration and Inflow (I/I) A Case Study at Risvollan, Trondheim*.

Lambert, A. (2001). What Do We Know About Pressure: Leakage Relationship in Dirstribution Systems. *Proc. IWA System Approach to Leakage Control and Water Distribution Systems Management*.

Langeveld, J. G., De Haan, C., Klootwijk, M., & Schilperoort, R. P. S. (2012). Monitoring the performance of a storm water separating manifold with distributed temperature sensing. *Water Science and Technology*, *66*(1), 145–150. https://doi.org/10.2166/wst.2012.152

Li, J., Yang, X., & Sitzenfrei, R. (2020). Rethinking the framework of smart water system: A review. *Water (Switzerland)*, *12*(2). https://doi.org/10.3390/w12020412

Lund, N. S. V., Kirstein, J. K., Madsen, H., Mark, O., Mikkelsen, P. S., & Borup, M. (2021). Feasibility of using smart meter water consumption data and in-sewer flow observations for sewer system analysis: A case study. *Journal of Hydroinformatics*, *23*(4), 795–812. https://doi.org/10.2166/hydro.2021.166

Lundblad, U., & Backö, J. (2014). *Svenskt Vatten Utveckling Juridisk och ekonomisk hantering av tillskotts vatten som sker till spill vattenförande ledning innanför förbindelsepunkt*. www.svensktvatten.se

Muleta, M. K., & Boulos, P. F. (2008). Analysis and calibration of RDII and design of sewer collection systems. *World Environmental and Water Resources Congress 2008: Ahupua'a - Proceedings of the World Environmental and Water Resources Congress 2008*, *316*. https://doi.org/10.1061/40976(316)642

Nash, J. E., & Sutcliffe, J. V. (1970). River flow forecasting through conceptual models part I - A discussion of principles. *Journal of Hydrology*, *10*(3), 282–290. https://doi.org/10.1016/0022-1694(70)90255-6

Norwegian Water. (2018). *The water services in Norway*. https://norskvann.no/wp-content/uploads/The_water_services_in_Norway_2018.pdf

Overview — pyswmm 1.2.0 documentation. (n.d.). Retrieved January 29, 2023, from https://pyswmm.readthedocs.io/en/stable/overview.html

Overview — WNTR 0.6.0dev documentation. (n.d.). Retrieved January 29, 2023, from https://wntr.readthedocs.io/en/latest/overview.html

Pieterse-Quirijns, E. J., Agudelo-Vera, C. M., & Blokker, E. J. M. (2012). *Modelling sustainability in water supply and drainage with SIMDEUM®*.

Pieterse-Quirijns, I. (2014). *Manual SIMDEUM Pattern Generator*.

Romano, M., & Kapelan, Z. (2014). Adaptive water demand forecasting for near real-time management of smart water distribution systems. *Environmental Modelling and Software*, *60*, 265–276. https://doi.org/10.1016/j.envsoft.2014.06.016

Rossman, L. A. (2015). *Storm Water Management Model User's Manual Version 5.1*.

Shi, Y., Xu, J., & Du, W. (2019). Discussion on the New Operation Management Mode of Hydraulic Engineering Based on the Digital Twin Technique. *Journal of Physics: Conference Series*, *1168*(2). https://doi.org/10.1088/1742-6596/1168/2/022044

Sola, K. J., Bjerkholt, J. T., Lindholm, O. G., & Ratnaweera, H. (2018). Infiltration and inflow (I/I) to wastewater systems in Norway, Sweden, Denmark, and Finland. *Water (Switzerland)*, *10*(11). https://doi.org/10.3390/w10111696

Sorge, H.-C. (2007). *Technische Zustandsbewertung metallischer Wasserversorgungsleitungen als Beitrag zur Rehabilitationsplanung [Condition Assessment of Metalic Water Supply Pipes to Support Rehabilitation Planning]*. Dissertation, Bauhaus University of Weimar.

Steffelbauer, D. (2021). *Optimization Part II - Course notes TVM4174-Hydroinformatics for Smart Water Systems*.

Storn, R., & Price, K. (1997). Differential Evolution-A Simple and Efficient Heuristic for Global Optimization over Continuous Spaces. In *Journal of Global Optimization* (Vol. 11). Kluwer Academic Publishers.

swmmio — swmmio 0.6.2 documentation. (n.d.). Retrieved April 7, 2023, from https://swmmio.readthedocs.io/en/v0.6.2/

Van Zyl, J. E., & Cassa, A. M. (2014). Modeling elastically deforming leaks in water distribution pipes. *Journal of Hydraulic Engineering*, *140*(2), 182–189. https://doi.org/10.1061/(ASCE)HY.1943-7900.0000813

WSAA. (2003). *Common Failure Modes in Pressurised Pipeline Systems*. Water Services Association of Australia.

Zhang, Q., Zheng, F., Jia, Y., Savic, D., & Kapelan, Z. (2021). Real-time foul sewer hydraulic modelling driven by water consumption data from water distribution systems. *Water Research*, *188*. https://doi.org/10.1016/j.watres.2020.116544

11 Appendix

11.1 Appendix A: Inflows in SWMM

Inflows into a sewer model using SWMM can be introduced in several ways according to the type of inflow we want to model and it´s characteristics. SWMM defines three main types: "Direct", "Dry Weather" and "RDII". This appendix aims to explain how the different inflows are computed by SWMM. The information provided is mainly taken from SWMM´s user manual (Rossman, 2015) unless another source is stated.

To specify the temporal history of direct external flow entering a sewer/drainage system node, the **"Direct"** tab in the inflows section in SWMM is used. Both a constant and a time-varying component is used to describe these inflows. The inflow in this case is computed as:

$$\begin{aligned} &Direct\ Inflow\ at\ time\ t \\ &\quad = (baseline\ value) \times (baseline\ pattern\ factor) \\ &\quad + (scale\ factor) \times (time\ series\ value\ at\ time\ t) \end{aligned} \qquad (11\text{-}1)$$

Where:

- Baseline values: Provides information on the value of the constituent's inflow's constant baseline component.
- Baseline pattern: is an optional Time Pattern that can be used to modify the baseline inflow hourly, daily, or monthly (depending on the type of time pattern specified).
- Time series: it contains inflow data over time.
- Scale factor: it is used to alter the values of the constituent´s inflow time series.

An ongoing source of dry weather flow entering a drainage system node can be specified using the **"Dry Weather"** Inflow. It is computed in SWMM as follows:

$$Dry\ weather\ Inflow = (Average\ value) \times (Pattern\ 1) \times (Pattern\ 2) \ldots \qquad (11\text{-}2)$$

Where:

- Average value: is the average value of the dry weather inflow.
- Patterns: are the patterns used to vary the dry weather flow change on a regular basis either by month of the year, day of the week, time of the day, etc.

Finally, **"RDII"** inflow refers to the rainfall dependent infiltration/inflow. Two inputs are required for this inflow: the unit hydrograph and the sewer shed area. To create a time series of RDII inflows per unit area during the simulation, the unit hydrographs in the group are combined with the group's designated rain gage. The sewer shed area, represents the area that will contribute to the RDII in the node (usually a smaller area than the total sub catchment that contributes to the node´s surface runoff). The unit hydrographs are added in SWMM as sets of three representing short-, medium- and long-term response. They are defined by three parameters:

- R = fraction of rainfall volume that enters the sewer system
- T = time from the onset of rainfall to the peak of the UH in hours

- K = ratio of time to recession of the UH to the time to peak.

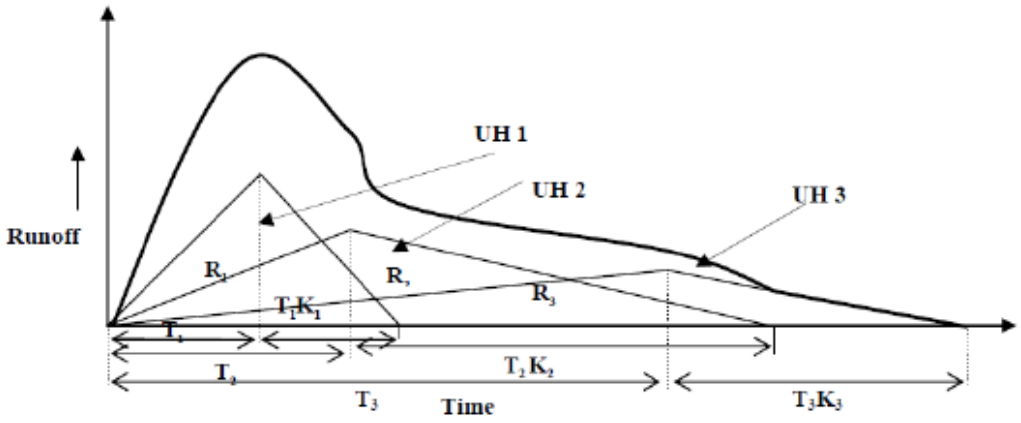


Figure 11-1: RTK-parameters in the unit-hydrograph. Extracted from Muleta & Boulos (2008).

11.2 Appendix B: Study by Lund et al. (2021)

The study done by Lund et al. (2021) evaluates the possibility of estimating the sewer discharges from the smart water meters placed in the water distribution network. In order to evaluate the results, they calculate the difference between the observed and simulated flows as:

$$Q_{res}(t) = Q_{obs}(t) - Q_{sim}(t) \qquad (11\text{-}3)$$

The residuals can be either positive or negative, exhibit a constant, diurnal or seasonal variation. Several reasons for the deviations are stated in Table 11-1.

Table 11-1: Potential reasons for the deviation between observed and simulated sewage flow (extracted from Lund et al. (2021)).

Deviation type	Potential reason	Expected residual pattern
Observed flow smaller than simulated flown (negative residuals, Equation (11-3))	Consumed water not discharged to the sewer	Diurnal
	Exfiltration	Seasonal
Observed flow larger than simulated flow (positive residuals, Equation (11-3))	Unaccounted for consumers	Diurnal
	Pumping of groundwater to the sewer system	Constant or seasonal
	Rainwater harvesting	Diurnal
	Snow melting (only in winter)	Temperature-dependent
	Infiltration	Constant-, seasonal- or rainfall dependent
	Sedimentation	Pipe geometry-dependent
General reasons	Sedimentation	Pipe geometry-dependent
	Erroneous smart meters, data transmission or data handling	Constant or diurnal
	Wrong conceptualization of the sewer system	Diurnal
	Erroneous in-sewer sensors	Constant or diurnal

11.3 Appendix C: Differential evolution algorithm

Evolution algorithms are based on the principle behind evolution of a species. The underlying concept is that the individuals that are more likely to survive, are those that are best adapted to the environment. The differential evolution algorithm was proposed by Ken Price and Rainer Storn in Storn & Price (1997).

The core principle of DE is that it creates new candidate solutions by multiplying a weighted difference vector between two population agents by a third agent (Storn & Price, 1997). The process can be seen as a sequence of mutation, recombination, and selection operators as shown in Figure 11-2. The information presented below is taken from Storn & Price (1997) and D. Steffelbauer (2021) unless it is specified otherwise.

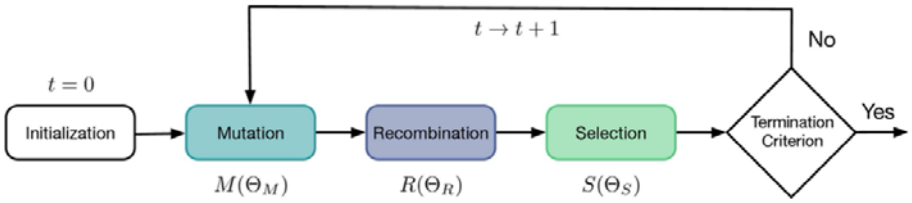


Figure 11-2: Sketch of DE algorithm. Extracted from (Steffelbauer, 2021)

The initialization phase consists of generating a random initial population of solutions, that are within the boundaries of the parameter space. Then, a new parameter vector is generated by the mutation operator. The vector is built by adding to a member of the population, a weighted difference vector (with weights Fi) between at least two other population members x_{r1} and x_{r2} which are chosen randomly. The first member can be selected either randomly from the initial population, is the same vector as used for recombination or is the best solution from the actual population. The method used depends on the mutation strategy, where several are presented in Table 11-2.

Table 11-2: Mutation strategies (extracted from Storn 1996)

Strategy	**Donor Formula ($\boldsymbol{v}_j = \ldots$)**
DE/rand/1	$\boldsymbol{x}_{r1} + F_1(\boldsymbol{x}_{r2} - \boldsymbol{x}_{r3})$
DE/best/1	$\boldsymbol{x}_{best} + F_1(\boldsymbol{x}_{r2} - \boldsymbol{x}_{r3})$
DE/rand to best/1	$\boldsymbol{x}_{r1} + F_1(\boldsymbol{x}_{r2} - \boldsymbol{x}_{r3}) + F_2(\boldsymbol{x}_{best} - \boldsymbol{x}_{r1})$
DE/current to best/1	$\boldsymbol{x}_j + F_1(\boldsymbol{x}_{r2} - \boldsymbol{x}_{r3}) + F_2(\boldsymbol{x}_{best} - \boldsymbol{x}_j)$
DE/rand/2	$\boldsymbol{x}_{r1} + F_1(\boldsymbol{x}_{r2} - \boldsymbol{x}_{r3} + \boldsymbol{x}_{r4} - \boldsymbol{x}_{r5})$
DE/best/2	$\boldsymbol{x}_{best} + F_1(\boldsymbol{x}_{r2} - \boldsymbol{x}_{r3} + \boldsymbol{x}_{r4} - \boldsymbol{x}_{r5})$

The following process is used by the recombination operator to combine the donor vector v_j and the trial vector u_j to create a more diverse set of parameter vectors:

$$u_{i,j} = \begin{cases} v_{i,j} & if\ rand_{i,j} \leq CR \\ x_{i,j} & else \end{cases} \quad (11\text{-}4)$$

Where:

- $rand_{i,j}$ = uniform random number withdrawn from the interval [0,1]
- CR = crossover probability determined by the user, within the interval [0,1]
- x_j = target vector
- j_{rand} = is a randomly drawn integer, which assures that one parameter at least differs from the trial vector and the target vector.

This implies that if for the i^{th} element of the vectors, $rand_{i,j}$ is smaller than CR, the i^{th} element will be relocated from the donor to the trial vector. On the contrary, if it is higher than CR, the i^{th} element will be taken from the target vector. This recombination process is explained in Figure 11-3.

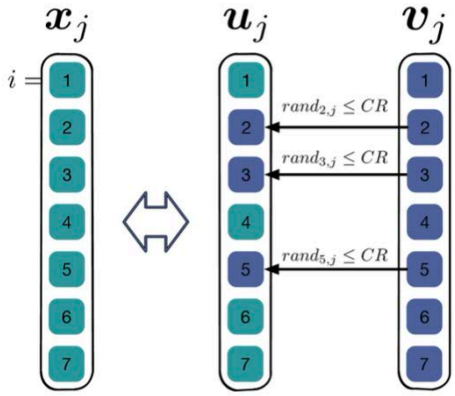


Figure 11-3: Recombination process for the DE algorithm. Extracted from (Steffelbauer, 2021).

Last, the selection operator compares the fitness of x_j and u_j in iteration k after they have been evaluated. The fitter individuals will continue to the next iteration. Mathematically, this can be expressed as:

$$x_j^{k+1} = \begin{cases} u_j^k & if \quad f(u_j^k) < f(x_j^k) \\ x_j^k & else \end{cases} \qquad (11\text{-}5)$$

11.4 Appendix D: Outflow results for LTH method for both model H1 and H2

11.4.1 Results for models H1

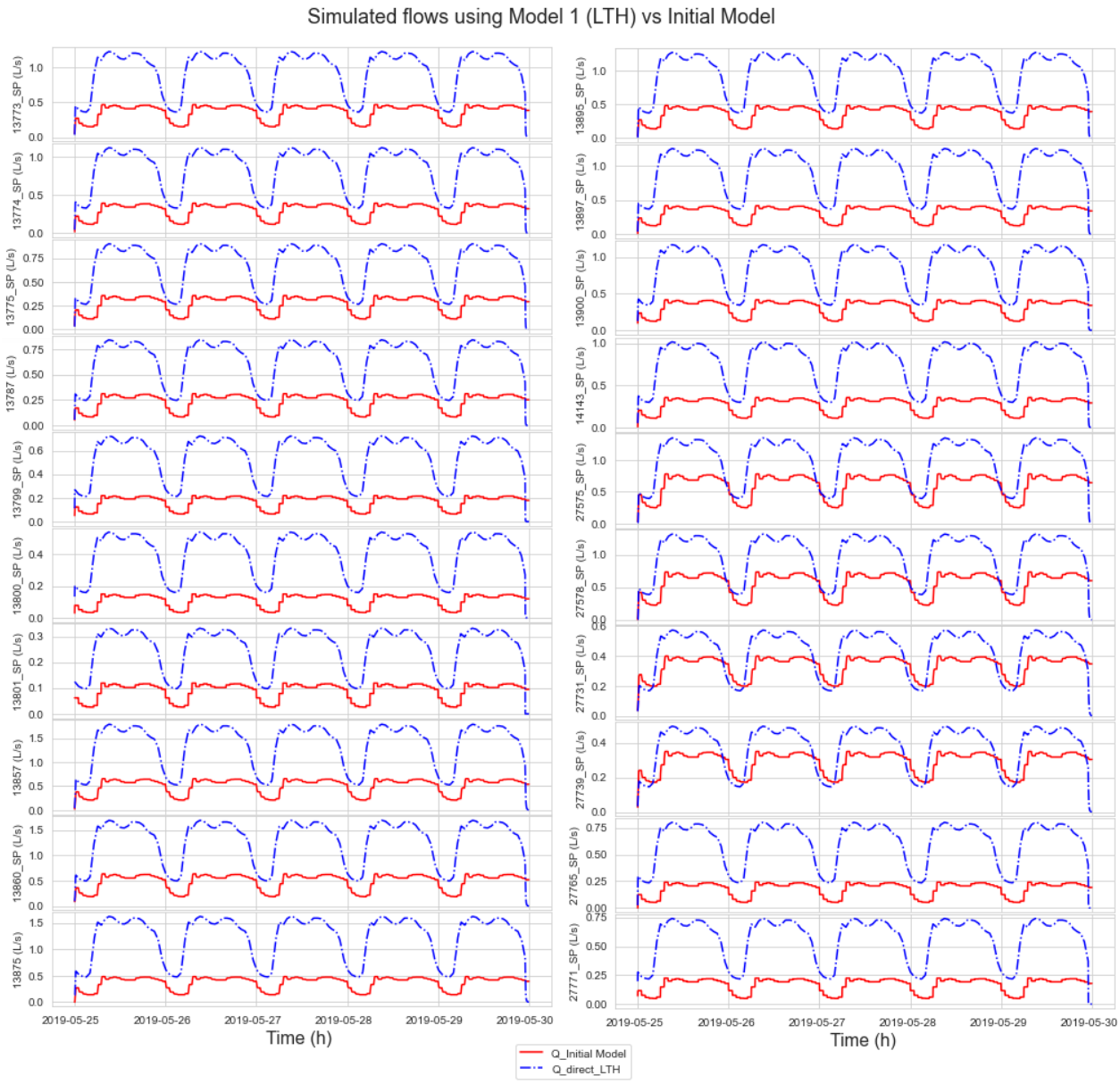


Figure 11-4: Comparison of flows using model H1 (direct flow) with link method LTH and the flows in the initial model (1).

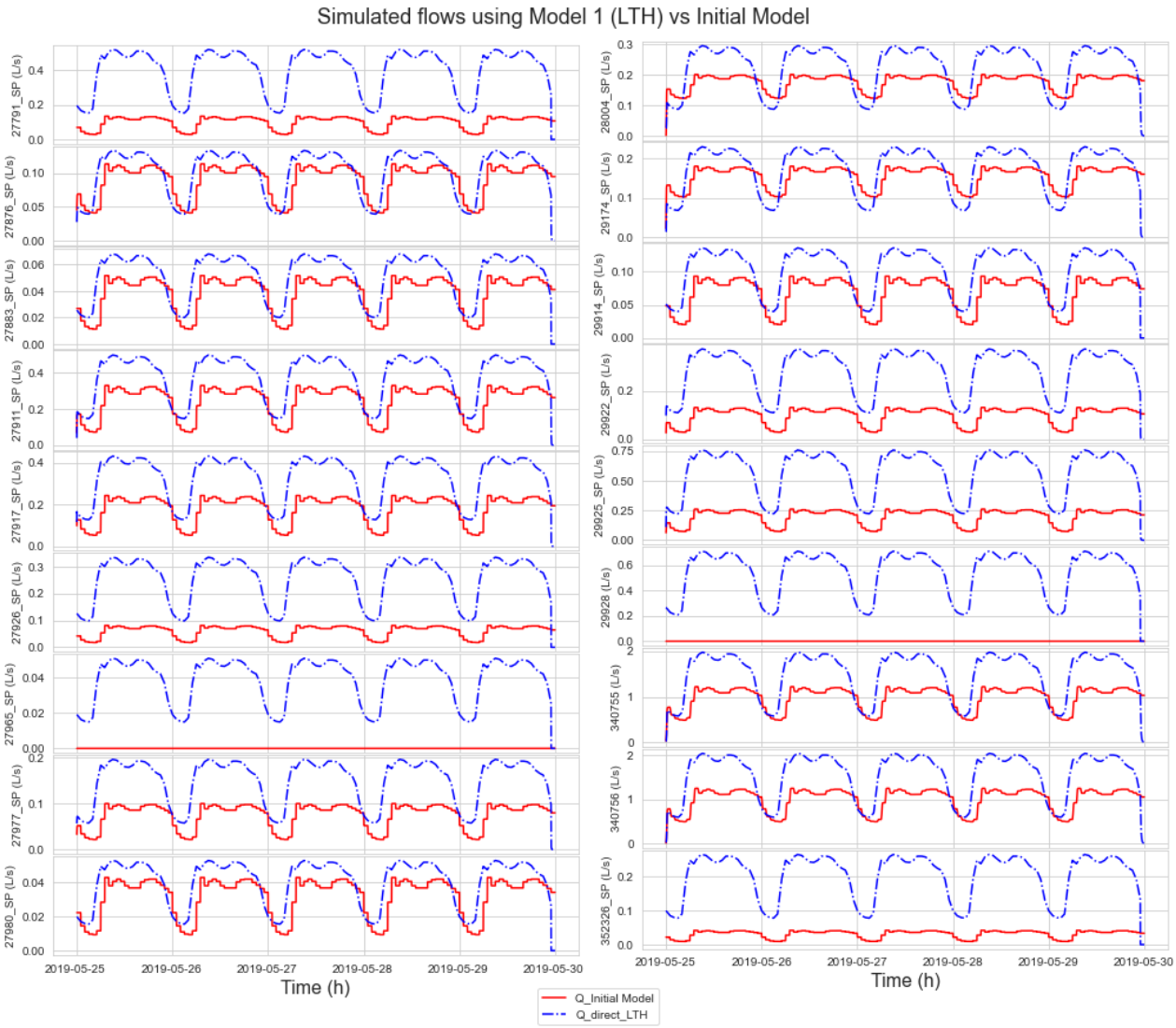


Figure 11-5: Comparison of flows using model H1 (direct flow) with link method LTH and the flows in the initial model (2).

11.4.2 Results for models H2

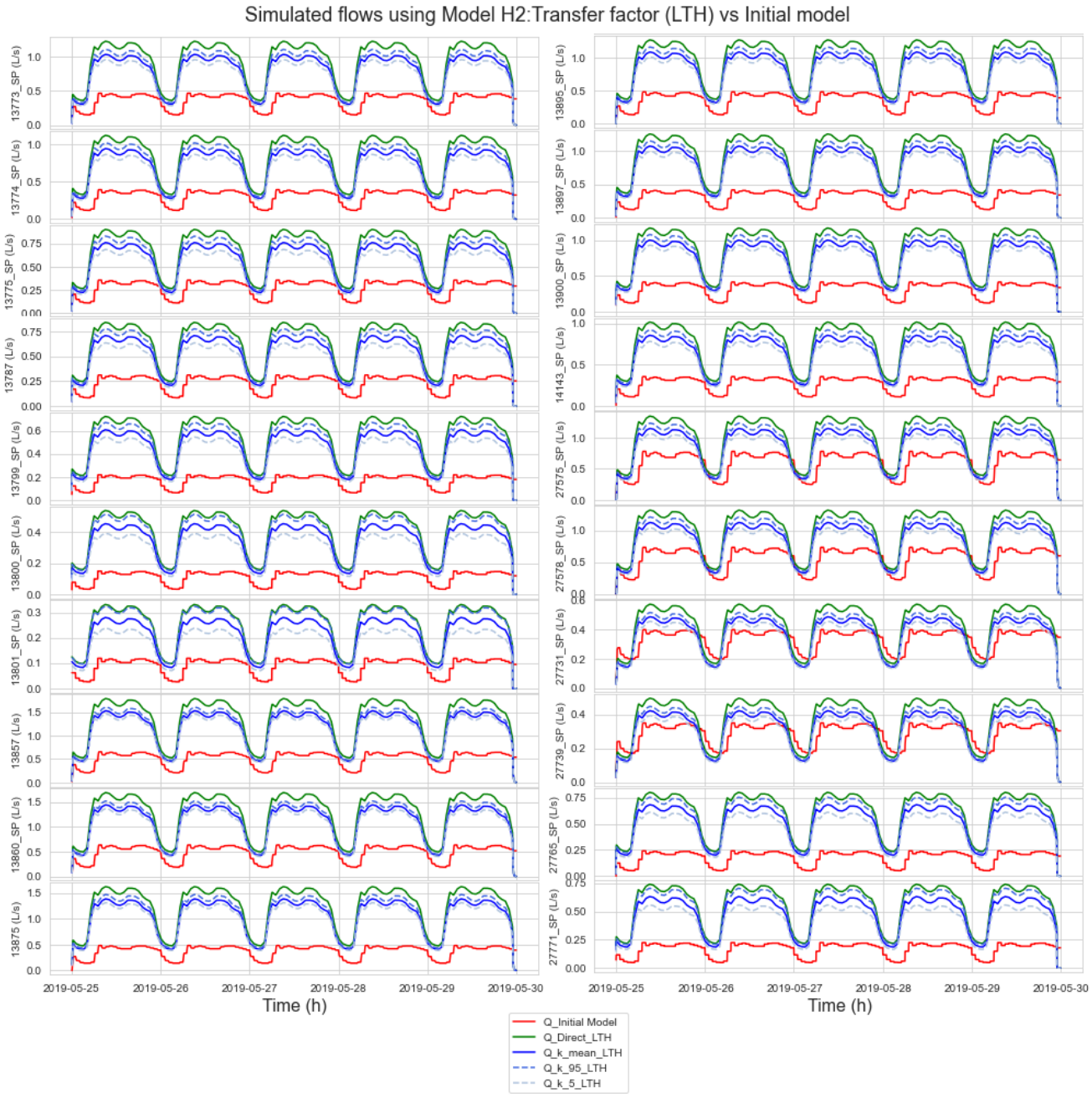


Figure 11-6: Comparison of flows using model H2 (transfer factor) with link method LTH and the flows in the initial model (1).

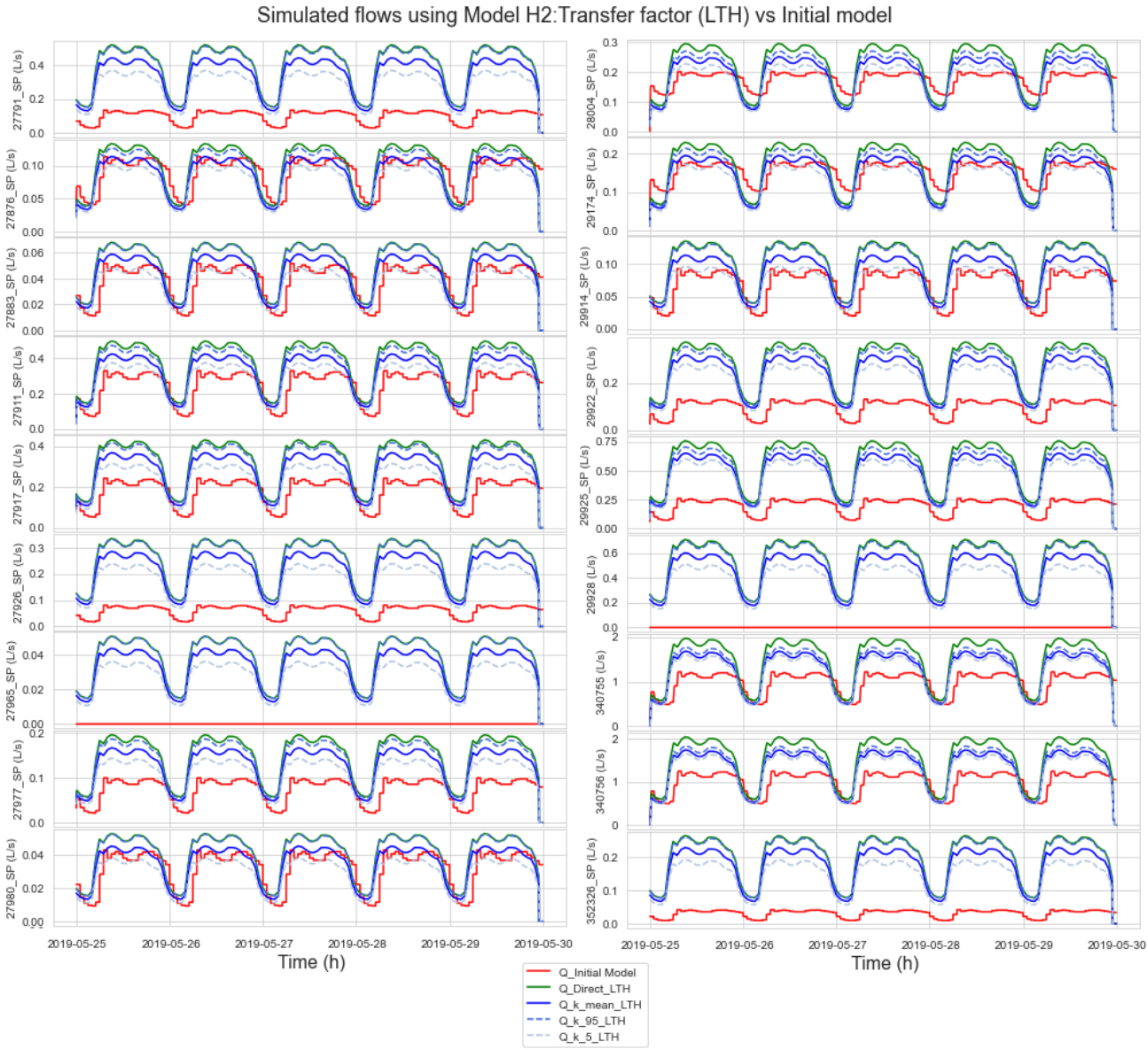


Figure 11-7: Comparison of flows using model H2 (transfer factor) with link method LTH and the flows in the initial model (2).

11.4.3 Results of changing transfer factor to k=0.4

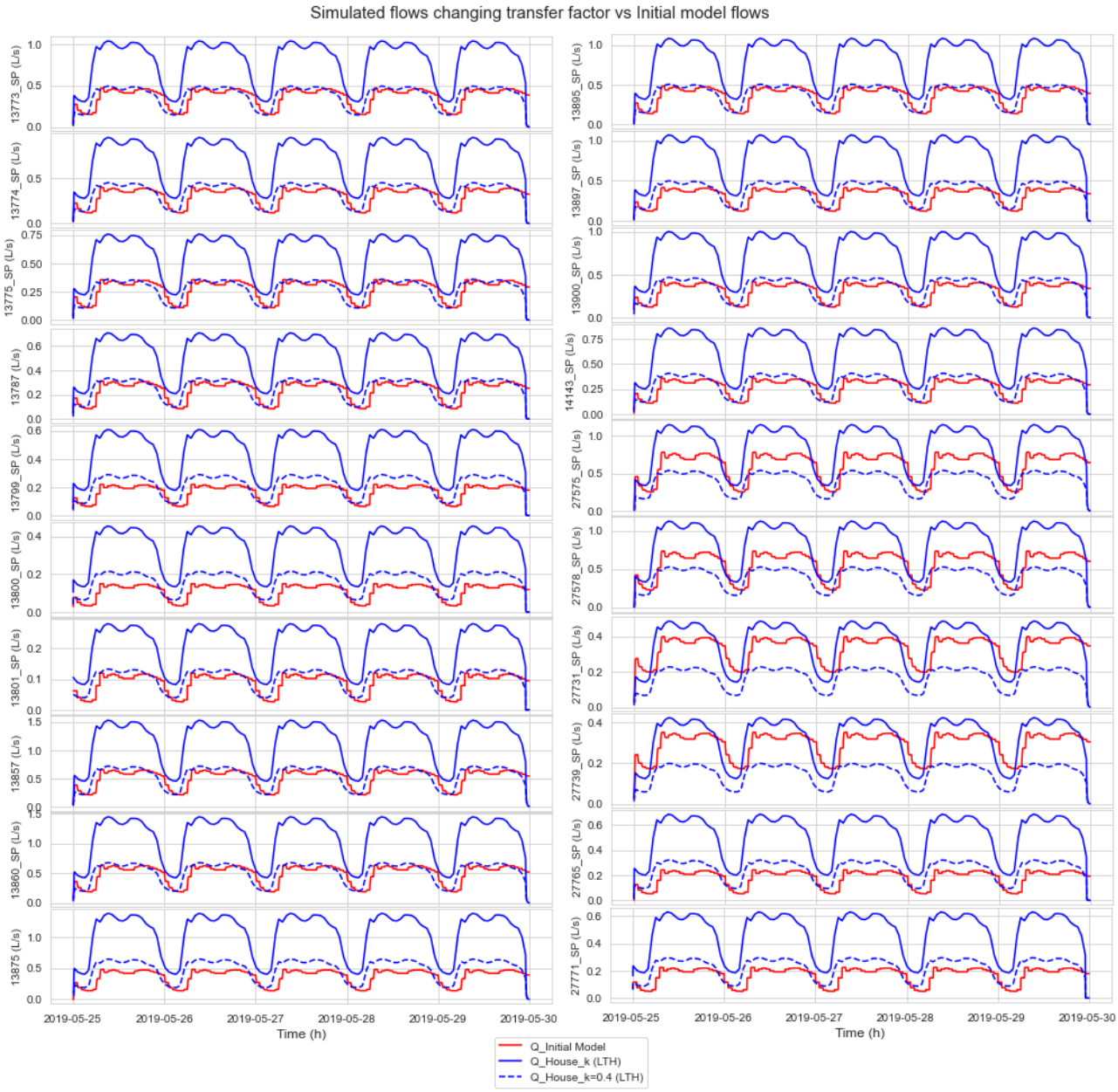


Figure 11-8: Comparison of flows using model H2 using transfer factor k=0.4 and the mean k factor obtained in section 6.1.2.2, with link method LTH and the flows in the initial model (1).

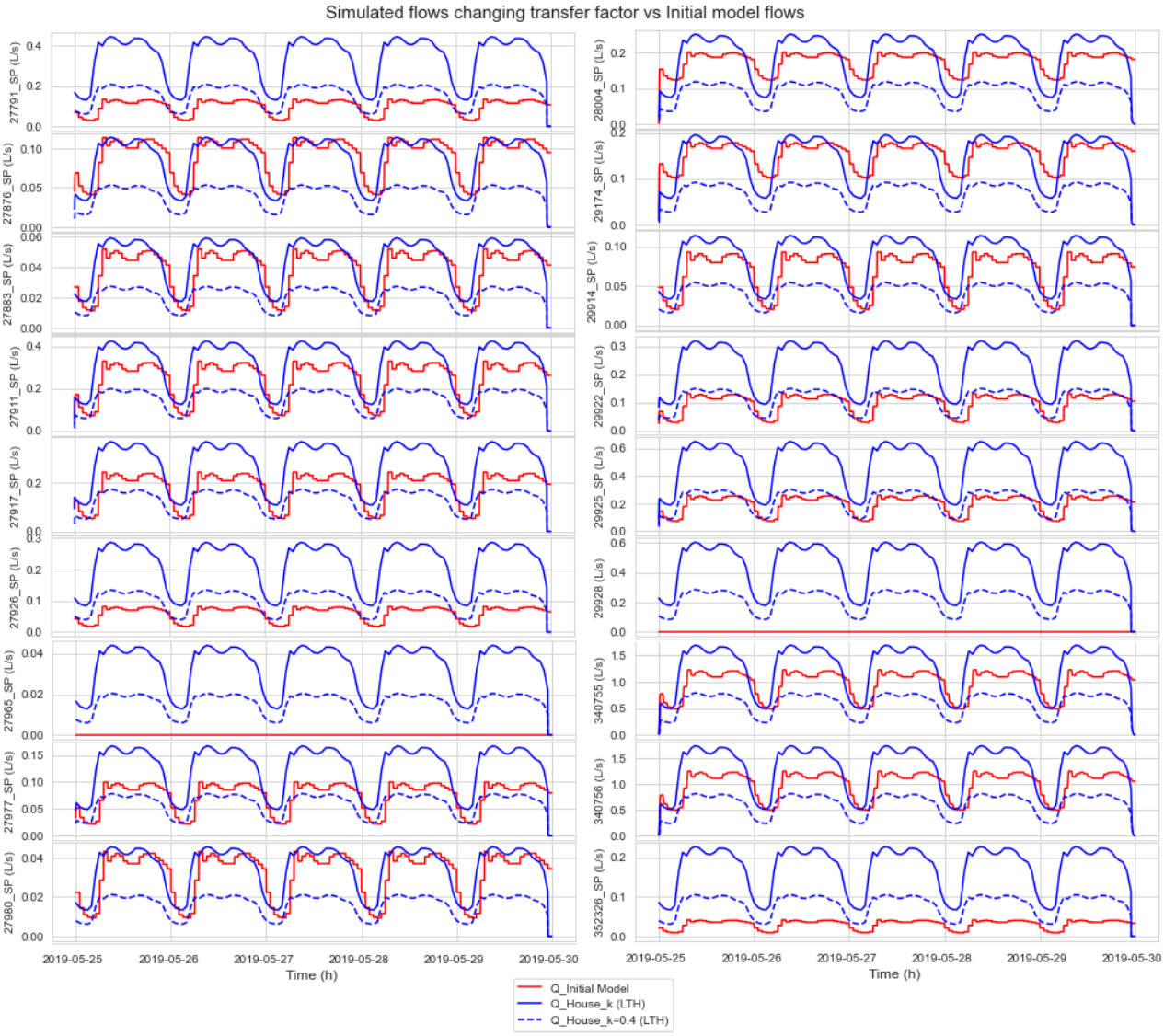


Figure 11-9: Comparison of flows using model H2 using transfer factor k=0.4 and the mean k factor obtained in section 6.1.2.2, with link method LTH and the flows in the initial model (2).

11.5 Appendix E: Performance evaluators for LTH model

Model 1 : Directly to sewer (LTH)

	MAPE	R^2	NSE	KGE
13773_SP	1.001307	0.543341	-31.417182	-1.453859
13774_SP	1.001328	0.541702	-37.911131	-1.7081
13775_SP	1.001301	0.542696	-26.780239	-1.280364
13787	1.001322	0.541308	-29.61989	-1.430339
13799_SP	1.001359	0.537579	-57.34887	-2.332108
13800_SP	1.001393	0.537047	-61.899246	-2.624501
13801_SP	1.001343	0.53637	-29.206207	-1.48411
13857	1.001316	0.544435	-34.95097	-1.586628
13860_SP	1.001314	0.545594	-30.296267	-1.431441
13875	inf	0.538148	-59.490715	-2.404726
13895_SP	1.001314	0.545505	-28.441657	-1.366807
13897_SP	1.00134	0.541062	-44.568199	-1.935887
13900_SP	1.001326	0.540451	-36.56509	-1.663906
14143_SP	1.001327	0.540878	-39.630696	-1.757466
27575_SP	1.001198	0.547058	-5.970532	-0.14241
27578_SP	1.001214	0.545915	-6.996689	-0.219235
27731_SP	1.001108	0.554358	-3.474101	-0.16157
27739_SP	1.001105	0.550152	-3.144132	-0.11742
27765_SP	inf	0.539095	-48.841994	-2.216922
27771_SP	1.001376	0.537852	-48.105333	-2.191835
27791_SP	1.001405	0.53637	-74.927123	-2.991144
27876_SP	1.001052	0.539122	-0.075772	0.535281
27883_SP	1.00115	0.536373	-0.636142	0.499646
27911_SP	1.001193	0.540971	-2.190544	0.260365
27917_SP	1.001239	0.537707	-5.428878	-0.093556
27926_SP	1.001417	0.536371	-88.791223	-3.347939
27965_SP	NaN	NaN	NaN	NaN
27977_SP	1.001265	0.537124	-8.416868	-0.344523
27980_SP	1.001127	0.536363	-0.163784	0.592747
28004_SP	1.001112	0.536455	-6.994032	-0.799705
29174_SP	1.001064	0.540318	-2.115211	-0.203181
29914_SP	1.001183	0.53637	-1.722324	0.32466
29922_SP	1.001347	0.536809	-31.265501	-1.569502
29925_SP	1.001336	0.537684	-38.406879	-1.748453
29928	NaN	NaN	NaN	NaN
340755	1.001159	0.549739	-5.298861	-0.15161
340756	1.001163	0.55077	-5.533823	-0.165954
352326_SP	1.001489	0.53637	-257.526052	-6.437622

Model 2: Transfer factor (LTH)

	MAPE	R^2	NSE	KGE
13773_SP	1.001267	0.54455	-17.450374	-0.855175
13774_SP	1.00129	0.542867	-21.63731	-1.062813
13775_SP	1.00126	0.543707	-14.553477	-0.705724
13787	1.001284	0.541957	-16.585763	-0.830454
13799_SP	1.001324	0.537868	-34.188773	-1.58356
13800_SP	1.00136	0.537223	-38.011204	-1.835367
13801_SP	1.001308	0.536379	-16.857581	-0.888387
13857	1.00128	0.546236	-20.64428	-1.009019
13860_SP	1.001278	0.547189	-17.672905	-0.873678
13875	inf	0.539682	-37.275004	-1.702639
13895_SP	1.001278	0.54685	-16.671349	-0.82675
13897_SP	1.001307	0.542127	-27.343151	-1.312316
13900_SP	1.001292	0.541121	-22.018876	-1.081955
14143_SP	1.00129	0.541884	-22.681804	-1.105309
27575_SP	1.001147	0.548985	-2.330725	0.213176
27578_SP	1.001166	0.547789	-2.951259	0.150371
27731_SP	1.001051	0.556516	-0.81182	0.174907
27739_SP	1.001047	0.551613	-0.691365	0.209755
27765_SP	inf	0.539532	-30.52077	-1.539659
27771_SP	1.001345	0.537979	-29.990552	-1.517012
27791_SP	1.001376	0.536379	-48.453713	-2.203907
27876_SP	1.001009	0.539472	0.345674	0.688915
27883_SP	1.001098	0.536383	0.20624	0.67548
27911_SP	1.001142	0.54134	-0.460452	0.533143
27917_SP	1.001195	0.537827	-2.247492	0.252779
27926_SP	1.001388	0.53638	-57.814044	-2.502078
27965_SP	NaN	NaN	NaN	NaN
27977_SP	1.001223	0.537326	-3.99069	0.048742
27980_SP	1.001072	0.536373	0.417353	0.720648
28004_SP	1.001057	0.53785	-2.453622	-0.37527
29174_SP	1.00101	0.54135	-0.716334	0.117903
29914_SP	1.001125	0.53638	-0.13409	0.599076
29922_SP	1.001314	0.537002	-18.761031	-0.990683
29925_SP	1.001304	0.53804	-23.719926	-1.169873
29928	NaN	NaN	NaN	NaN
340755	1.001105	0.552011	-1.775391	0.199635
340756	1.001109	0.553238	-1.903389	0.188734
352326_SP	1.001465	0.53638	-181.847557	-5.241471

Figure 11-10: Performance evaluators for Model 1 (left) and Model 2 (right) for household demand, using spatial link LTH.

DEPARTMENT OF INFRASTRUCTURE AND ENVIRONMENTAL ENGINEERING
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2023

www.chalmers.se