Master's thesis

NTNU
Norwegian University of Science and Technology
Faculty of Engineering
Department of Civil and Environmental Engineering

Christian Skagsoset

Quantifying Infiltration and Inflow in Data-Scarce Environments

Possibilities and Co-benefits for Municipalities

Master's thesis in Civil- and Environmental Engineering
Supervisor: Franz Tscheikner-Gratl, Marius Møller Rokstad, Bardia Roghani

June 2023

NTNU
Norwegian University of Science and Technology

Christian Skagsoset

Quantifying Infiltration and Inflow in Data-Scarce Environments

Possibilities and Co-benefits for Municipalities

Master's thesis in Civil- and Environmental Engineering
Supervisor: Franz Tscheikner-Gratl, Marius Møller Rokstad, Bardia Roghani
June 2023

Norwegian University of Science and Technology
Faculty of Engineering
Department of Civil and Environmental Engineering

NTNU
Norwegian University of
Science and Technology

Abstract

Wastewater systems play a vital role in transporting wastewater away from populated areas, ensuring essential services for the global population. However, the aging infrastructure and limited capacity of sewers, coupled with the growing challenges of climate change and environmental issues, pose significant vulnerabilities for these systems. In addition, recent studies have revealed that infiltration and inflow (I/I) pose a significant problem for Norwegian municipalities, accounting for approximately 50% of the transported volumes, causing undesirable levels of stress on these systems.

Therefore, the primary objective of this study is to accurately quantify the extent of I/I, focusing on applicable approaches for small municipalities. To achieve this, novel adaptations of well-established methods for assessing I/I in municipal sewers are implemented. Specifically, the study employs various flow-based methods, including water balance and Pearson Correlation Coefficient, to select an area for evaluation by utilizing existing municipal data. In addition, the study leverages concepts such as graph theory and hydraulic modeling to quantify I/I by simulating household generated sewage flow and comparing it to wastewater flow rates. Furthermore, a mass balance approach is used to reverse engineer measured pumped volumes and estimate these wastewater flow rates.

The flow-based methods successfully managed to locate already known sources of I/I, while pointing towards new areas for evaluation. Nevertheless, some uncertainty was experienced with these methods as they lacked the desired accuracy. In the evaluation of a specific sewer zone, the effective construction of a hydraulic sewer model, provided an applicable approach to estimate sewage generated by households, subsequently enabling the quantification of I/I. Over the 40-day evaluation period, it was estimated that 76% of the total transported volume was attributed to I/I. Mainly caused by a period of significant snowmelt and rainfall, where sub-optimal manhole design and the pipes placement in ditches were possible amplifiers.

Overall, this study identified practical solutions for small municipalities to address I/I. By utilizing a hydraulic model for validation and refinement of the practical Flow-based methods further enhanced the accuracy, resulting in a combined approach. In addition, the study highlighted the efficiency of generating a sewer network graph using the municipal database, facilitating network analysis and data filling. Consequently, the municipality in this case gained valuable information for decision support, identifying areas of concern and guiding actions to enhance sewer system performance. As a co-benefit of this approach, which can be adapted for other municipalities, facilitates resource optimization and enhance decision-making in I/I management, particularly benefiting municipalities with limited resources and expertise.

Keywords: Infiltration and inflow; Data analysis; Graph theory; Hydraulic modelling; Pearson correlation coefficient

Sammendrag

Avløpssystemer spiller en avgjørende rolle i å transportere avløpsvann bort fra befolkede områder og sikre viktige tjenester for verdens befolkning. Imidlertid utgjør den aldrende infrastrukturen og den begrensede kapasiteten i kloakksystemer, kombinert med de økende utfordringene knyttet til klimaendringer og miljøproblemer, betydelige sårbarheter for disse systemene. I tillegg har nylige undersøkelser vist at fremmedvann utgjør et betydelig problem for norske kommuner, og står for omtrent 50% av de transporterte avløpsmengdene, og dermed forårsaker uønsket belastning på disse systemene.

Hovedmålet med denne studien er derfor å nøyaktig estimere omfanget av fremmedvann, med fokus på anvendelige tilnærminger for små kommuner. For å oppnå dette implementeres nyutviklede tilpasninger av etablerte metoder for å vurdere andelen fremmedvann i et kommunalt ledningsnett. Studien benytter ulike metoder basert på vannføring, inkludert en vannbalanse metode og Pearson-korrelasjonskoeffisient, for å velge et område for videre undersøkelser, ved hjelp av eksisterende kommunale data. Videre bruker studien konsepter som grafteori og hydraulisk modellering for å estimere fremmedvann ved å simulere avløp produsert av husholdninger og videre sammenligne dette med strømningsmålinger av avløp. I tillegg benyttes en massebalansemetode for å omgjøre målte videreførte volumer til å estimere disse avløpsstrømningene.

De vannføringsbaserte metodene klarte å lokalisere områder med allerede kjente kilder til fremmedvann, og pekte samtidig mot nye områder som kunne evalueres. Likevel var det et nivå av usikkerhet med disse metodene da de manglet den ønskede graden av nøyaktighet. Ved evalueringen av en spesifikk avløpssone viste det seg at å ta i bruk en hydraulisk kloakkmodell var en anvendelig tilnærming for å estimere avløpsvann produsert av husholdninger, og dermed muliggjøre estimeringen av fremmedvann. I løpet av den 40 dager lange evalueringsperioden ble det estimert at 76% av de transporterte mengdene skyldtes fremmedvann. Dette ble hovedsakelig forårsaket av betydelige mengder snøsmelting og nedbør, der mindre optimalt kumdesign og plasseringen av rør i grøfter var mulige forsterkninger av dette problemet.

Samlet sett identifiserte denne studien praktiske løsninger for små kommuner til å få oversikt over sin fremmedvanns problematikk. Ved å bruke hydrauliske modeller for validering og forbedring av de praktiske vannføringsbaserte metodene, økte nøyaktigheten betydelig, noe som resulterte i en kombinert tilnærming. Studien understrekte også effektiviteten av å generere et graf representert avløpssystem ved hjelp av kommunens database, noe som muliggjorde nettverksanalyse og utfylling av data. Dermed ble tilleggs fordelene ved denne tilnærmingen, muligheter for ressursoptimalisering og grunnlag for beslutningsprosesser innenfor håndtering av fremmedvann, spesielt til fordel for kommuner med begrensede ressurser og ekspertise.

Acknowledgements

I would like to express my deepest gratitude to the individuals who have contributed to the completion of my Master's Thesis and the fulfillment of my M.Sc. in Civil and Environmental Engineering at the Norwegian University of Science and Technology.

First and foremost, I would like to thank my supervisors, Franz Tscheikner-Gratl, Marius Møller Rokstad, and Bardia Roghani. Their expertise, guidance, and attention to detail have been invaluable throughout the entire project. Franz Tscheikner-Gratl, in particular, provided detailed advice that helped set the path for my research. Marius Møller Rokstad generously shared his knowledge and experience on related topics, enriching my understanding of the subject matter. Bardia Roghani's availability and the quality of his feedback and discussions were greatly appreciated.

I would also like to extend my appreciation to my project partner, Eirik Evjen Hagtvedt from Horten municipality. His availability and helpfulness in every aspect of this project were crucial in its success.

To my fellow students who have been part of this journey, I am grateful for your companionship and support. We have supported and inspired each other throughout this shared experience. In particular, I would like to express special thanks to my student colleague, Joakim Skjelde, for continuously providing wholesome feedback and encouragement throughout this spring.

Trondheim

June 10, 2023

Contents

List of Figures

Abbreviations

BSF Base Sanitary Flow

CCTV Closed-Circuit TeleVision

DTS Digital Temperature Sensing

DWF Dry Weather Flow

EPA U.S. Environmental Protection Agency

GIS Geographic Information System

GWI GroundWater Infiltration

I/I Infiltration and Inflow

MNF Minimum Night-Flow

MSE Mean Square Error

NSE Nash-Sutcliffe Efficiency

PE Person Equivalent

RDII Rainfall-Derived Infiltration and Inflow

SCADA Supervisory Control And Data Acquisition

SWMM Storm Water Management Model

WWP WasteWater Production

WWTP WasteWater Treatment Plant

1 Introduction

Wastewater systems provide the world's population with essential services by transporting wastewater away from populated areas (Saletti et al., 2023). Unfortunately, the suboptimal design and high level of degradation of a significant amount of these systems prevents them from functioning at their full potential (Beheshti et al., 2015; Ratnaweera et al., 2018). As the world faces increasingly challenging climate and environmental issues in the future, these sewers will become increasingly vulnerable due to their aging infrastructure and limited capacity to handle the increasing demands. The Intergovernmental Panel on Climate Change has stated that the frequency of intense rainfall is expected to increase as a consequence of climate change (Lee et al., 2023). This will result in higher proportions of infiltration and inflow (I/I) within the sewer system, thus further increasing the risk of sewer overflows, basement flooding and the amount of water transported to treatment plants (Nie et al., 2009; Saletti et al., 2023).

In the years leading up to the proposal for a revised EU Urban Wastewater Treatment Directive ("Council Directive 91/271/EEC", 1991), there has been a notable increase in research studies and monitoring campaigns focusing on wastewater (European Commission, 2022). These efforts have proven important in addressing pollution in sewer systems. Furthermore, this have compelled Norwegian municipalities to address some of the challenges associated with the water industry, which is now required to extend its pollution reporting to comply with the National Environmental Directive (Miljødirektoratet, 2023) as well as the EU directive. While several municipalities are yet to meet the existing pollution requirements, they are making efforts to do so (Jørgensen & Rostad, 2021). However, the state of the sewer systems has been a hindrance due to the high costs associated with asset renewal, resulting in a backlog of renewal for the past years (Bruaset, 2019). An estimate by Norsk Vann's report suggests that an investment of around 114 billion Norwegian kroner would be required from 2021 to 2040 to upgrade the existing sewer systems (Bruaset et al., 2021).

With the accumulation of data by Norwegian municipalities, more precise strategies can be implemented to address the most challenging areas first, reducing costs and environmental hazards. Studies by Norsk Vann and BedreVann have revealed that I/I pose a significant problem for Norwegian municipalities, accounting for approximately 50% of the conveyed volumes, while the national goal is to reduce it to less than 30% by 2030 (Jørgensen & Rostad, 2021; Scherling et al., 2020). This suggests that a large portion of Norwegian municipalities will need to take action to reduce excess water, in order to gain better I/I control, and subsequently reduce the pollution released by sewer overflows (e.g. Nitrogen, Phosphorus and Pathogens). Specifically, implementing effective stormwater management strategies, upgrading wastewater treatment facilities, and improving sewer infrastructure are crucial steps to mitigate the discharge of pollutants into the environment (Lindholm, 2017; Saletti, 2021).

At the present time, numerous studies have been conducted to address I/I in sewer systems through extensive measurement campaigns for localization or quantification (Beheshti & Sægrov, 2018; Hey et al., 2016; Lepot et al., 2017). Nevertheless, these methods can be considered time-consuming or expensive when applied across large areas. Although the most precise methods can be challenging to implement universally, they could be advantageous to implement for smaller sections. However, data scarcity in Norwegian municipalities regarding wastewater monitoring is a prevalent challenge, which results in a lack of reliable and up-to-date data on crucial parameters such as flow rates. This data scarcity poses obstacles to identifying I/I sources, assessing the effectiveness of reduction measures, and prioritizing necessary actions. By addressing data scarcity and enhancing wastewater monitoring capabilities in Norwegian municipalities, it would be possible to improve water management practices. Additionally, as municipalities are already generating data regarding sewer flows from pumping stations and stores this information in databases. This could be utilized with certain assumptions, to estimate I/I in the system (Beheshti et al., 2015; Mitchell et al., 2007).

With easily acquirable data, a comparison of sewer flows under different weather conditions can provide valuable insights into the performance and condition of the sewer system. Especially as weather is known to significantly impact flow rates and create more challenging hydraulic conditions (Hey et al., 2016; Lindholm et al., 2012; Ranck, 2017). Additionally, while a certain degree of increased volume is expected in combined sewers, a significant volume increase in a separated sewer is both "unexpected" and undesirable, as it is not necessarily designed to handle excessive water effectively. Overall, flow rate analysis can play a crucial role in identifying areas of the sewer system that require further investigation, utilizing precise methods for more accurate quantification and localization (Hey et al., 2016).

This study aims to prioritize areas for investigation by analyzing correlations between flow rate and two weather variables, namely precipitation and air temperature. By providing an accessible approach for municipalities with limited data, this research will assist in identifying focus areas for addressing I/I issues. Additionally, this study will focus on a specific area by measuring flow rates under different weather conditions and constructing an EPA-SWMM sewer model (SWMM model), leveraging graph theory to fill in data gaps. This model will serve as an initial step toward developing a comprehensive hydrodynamic model that can be utilized with the municipality's existing sewer dataset. Furthermore, the sewer model will be employed to reverse engineer sewage patterns within the system and quantitatively assess the extent of I/I by comparing the observed flow rates with the simulated sewage pattern, using the water demand as input to the model. Hence, the primary objective of this study is to quantify extent of I/I in a municipal sewer zone, by utilizing established approaches that are applicable for small municipalities.

It is important to note that this study does not provide extensive geographical assessments, nor does it consider the risks or costs associated with increased flow volumes. Additionally, the assessment of pipe conditions or examination of existing closed-circuit television-inspections (CCTV) are outside the scope of this study. This study neither involves the development of a fully comprehensive SWMM model that incorporates catchments, precipitation data, and storm sewers, as it only includes the sanitary sewer.

This thesis is structured in a systematic manner to address the research objective. The first chapter (chapter 2) presents a literature review that examines existing knowledge on I/I and its methods for quantification and localization. Moving forward, chapter 3 focuses specifically on the selected municipality for the case study, providing contextual information about the area under investigation. By highlighting key characteristics and important details, this section contributes to a deeper understanding of the local conditions. In chapter 4, the methodological framework is presented, offering a clear outline of the approach used to assess I/I. This chapter provides a detailed description of the specific procedures employed in the study. Building upon the methodology, chapter 5 presents the findings related to I/I, accompanied by a analysis and discussion of the results. Within this chapter, the implications and significance of these findings are elaborated upon. Finally, chapter 6 serves as the concluding section of the thesis, summarizing the key findings from the study and highlighting the potential co-benefits associated with implementing the proposed approach. Additionally, this chapter suggests a potential future research direction.

2 Literature Review

The purpose of this chapter is to provide an overview of I/I in sanitary sewer systems, including their causes, consequences, and components of wastewater flow (section 2.1). The chapter also discusses common methods used for quantifying I/I in section 2.2 and provides insights in the novelty of the proposed solutions for quantifying I/I in municipal sewer systems in section 2.3.

2.1 Infiltration & Inflow

I/I are factors that contribute to excess water in sanitary sewers, which can have significant consequences for the sewer system's performance. Reduced hydraulic capacity, increased treatment volume, longer pumping station operational times, and the risk of overflows or backflows are just a few examples of the negative impacts of I/I (Beheshti & Sægrov, 2018; Lindholm, 2017; Sola et al., 2019). These consequences can further result in higher operational costs, environmental degradation, and health hazards. The influence of I/I is commonly associated with aging infrastructure and precipitation, but human activities such as the construction of buildings and roads, or changes in land use can also contribute to I/I (Comeau et al., 2019). Meaning that the excess flow is caused by a variety of unintended sources of I/I, as presented in Figure 2.1.

Specifically, infiltration refers to groundwater seeping into the sewer system through sources such as broken pipes, deteriorated maintenance holes, and root intrusions, which have an impact during both dry- and wet weather conditions (Ellis & Bertrand-Krajewski, 2010; Lindholm, 2017). In general, infiltration can be affected by changes in the groundwater table or pressure, due to precipitation and can naturally vary throughout the seasons. However, human activities can also impact infiltration by altering the groundwater table or drainage patterns (Comeau et al., 2019).

On the other hand, inflow is the entry of surface water runoff into the sewer system via storm water drains, roof leaders, or other sources like land drainage and streams (Butler et al., 2018; Ellis & Bertrand-Krajewski, 2010). Furthermore, an additional significant source of inflow occurs during the transition from a separate sewer system to a combined system, where storm sewers and sanitary sewers are interconnected (Beheshti et al., 2015). These connections also occur on unintended locations, often called illicit or illegal connections. In these conditions, the separate sanitary sewers are particularly vulnerable to I/I because their pipes have smaller capacities, as described by Butler et al. (2018). Hence, I/I is therefore likely to have the greatest impact on these systems, resulting in the most detrimental consequences on the system. Furthermore, the rate of inflow is particularly high during periods of heavy rainfall or snowmelt, especially when the surface water flow exceeds the capacity of the storm sewer system (Goulding et al., 2012).

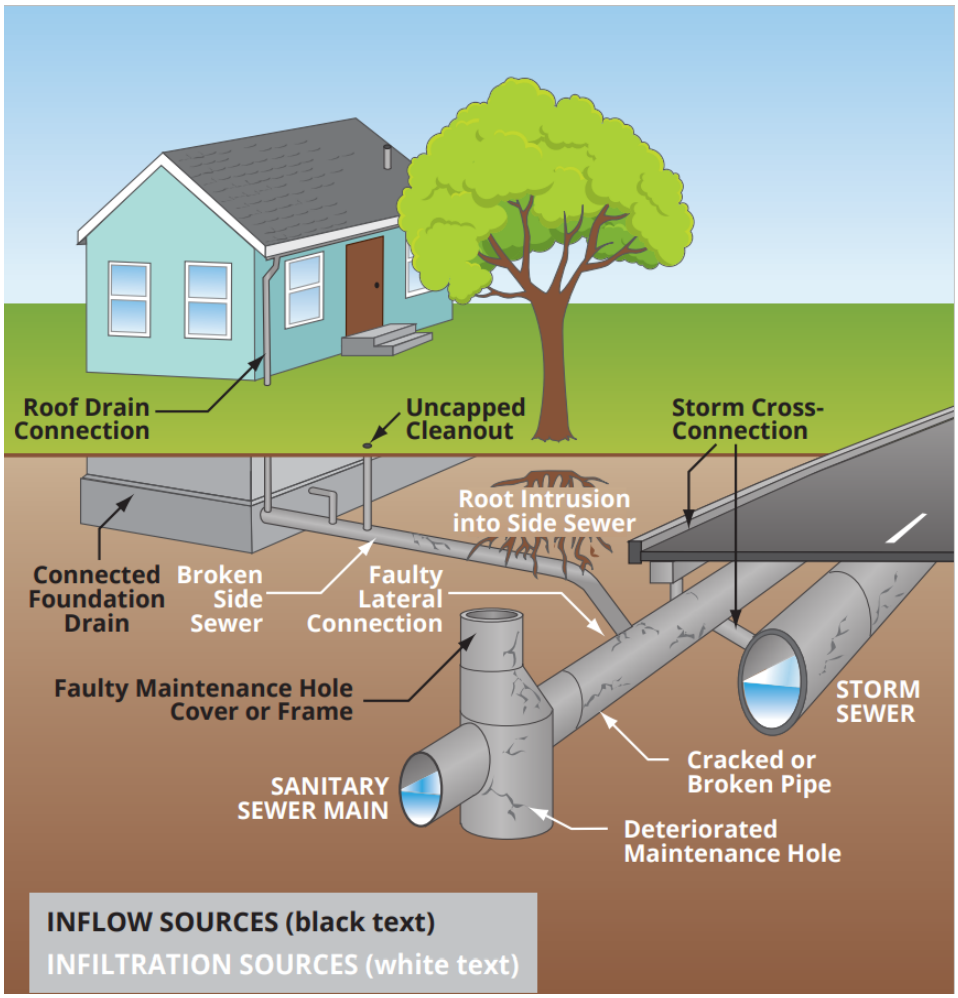


Figure 2.1: Illustration of I/I Sources (King County, 2021)

2.1.1 Components of Wastewater Flow

Wastewater flow can be categorized into three primary components, as presented in Figure 2.2. The first component is base sanitary flow (BSF), originating from residential, commercial, institutional, and industrial sources and follows a daily diurnal pattern (Comeau et al., 2019). BSF in residential areas is generally highest in the morning and evening, lower at night and early morning, and vary noticeably on weekdays, weekends, and holidays. More about diurnal wastewater fluctuations can be seen in Appendix C.

The second component, known as groundwater infiltration (GWI) or base infiltration, contributes to the slowly varying baseflow throughout the year (Comeau et al., 2019). This component typically increases towards the end of the wet season and is less distinct after an extended dry period. Flow rates consisting of only BSF and GWI are known as dry-weather flow (DWF), which occurs when there is no precipitation affecting the flow rates. However, accurately determining these components is challenging. Therefore, to estimate and separate flow components, commonly used methods involve making assumptions about water consumption return rates and wastewater composition during the early morning hours. As shown in the studies of Mitchell et al. (2007) and Bogusławski et al. (2022), there are several simple techniques that can be used to estimate GWI components which utilizes these assumptions (e.g. WWP-method,

Stevens-Schutzbach method and Minimum Flow Factor method).

The final component of wastewater flow is rainfall-derived infiltration and inflow (RDII), also known as the wet-weather component, which refers to the increase in flow rates due to rainfall (Comeau et al., 2019). Bäckman et al. (1993) and Saletti (2021) categorized this flow component into two response times: slow runoff response and fast runoff response. The slow runoff response primarily consists of increased infiltration rates due to rising groundwater tables, while the fast runoff component is the direct inflow response, often caused by storm sewer connections. Furthermore, flow rates with the inclusion of the RDII component is defined as wet-weather flow (WWF). Hence, RDII is often estimated based on the difference between dry- and wet-weather flow. To effectively evaluate the consequences caused by excess RDII, it's essential to have a comprehensive understanding of sanitary sewer flow in both dry and wet weather conditions (Comeau et al., 2019). This requires conducting a thorough baseline characterization of wastewater.

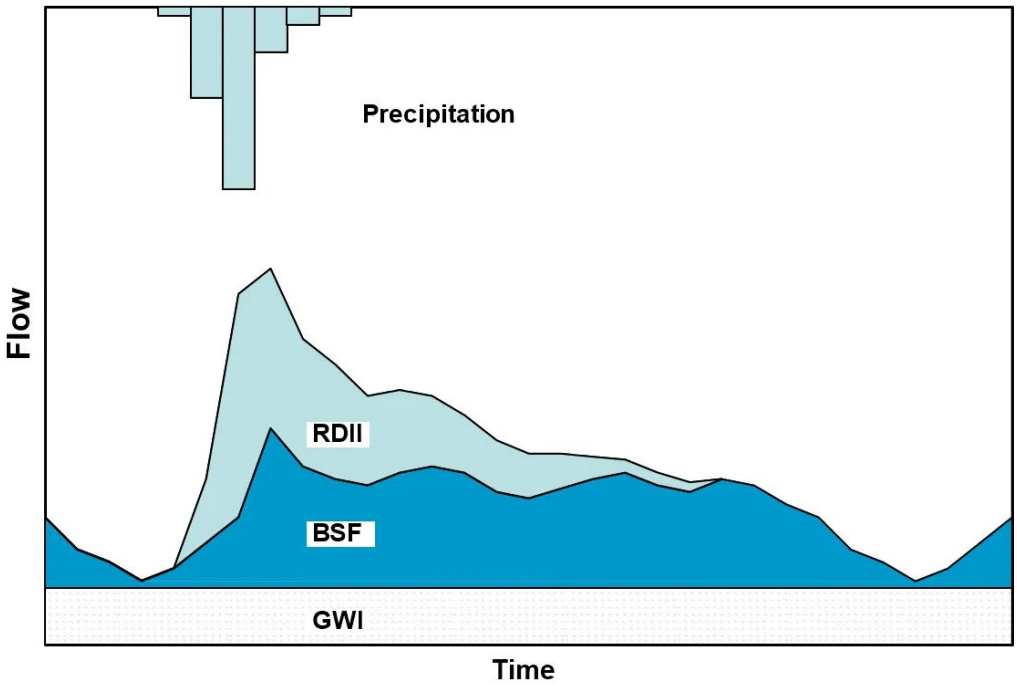


Figure 2.2: Wastewater Flow Components (Rossman & Huber, 2016)

2.2 Methods for Quantifying and Localizing I/I in Sewer Systems

There have been a range of methods developed to address the challenge of excess water caused by I/I in sewer systems (Hey et al., 2016). These methods can broadly be categorized into localization and quantifying methods. The quantity methods, such as flow monitoring or tracer methods, are typically limited to measuring the flow rate of excess water. On the other hand, localization methods are aimed at determining the source and location of excess water, and often involve visual inspections or temperature difference analysis (e.g. DTS or physical inspection). Saletti (2021) further categorized these methods based on their specific application, which is summarised in Figure 2.3

and examples of implementations are elaborated below. For detailed descriptions of the different methods, refer to the study of Saletti (2021).

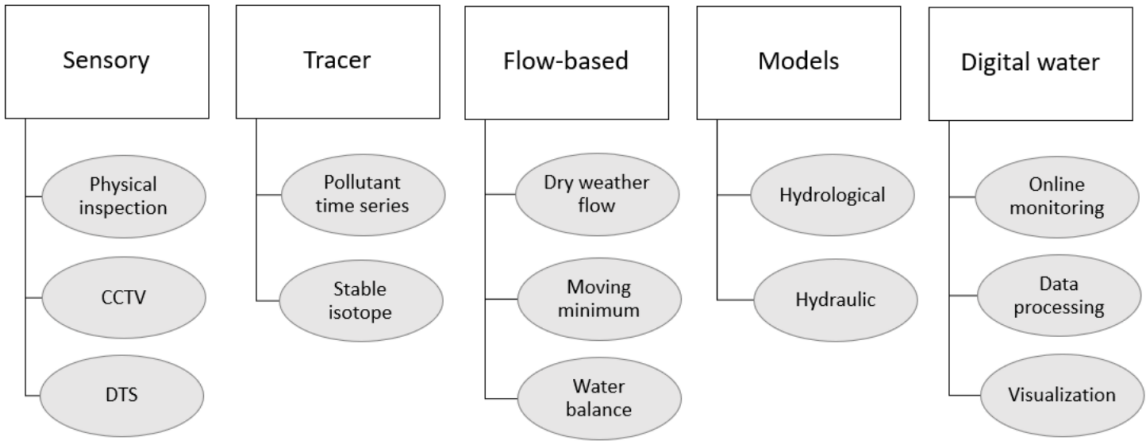


Figure 2.3: Methods for Quantifying and Localizing I/I (CCTV - Closed-Circuit Television, DTS - Digital Temperature Sensors) (Saletti, 2021)

Several studies have utilized sensory techniques, specifically Beheshti and Sægrov (2018) and Hoes et al. (2009) showed that digital temperature sensors (DTS) were a feasible and accurate method for measuring extraneous water from individual I/I sources in separate sewer networks. The method was demonstrated to be robust and practical in I/I measurement during wet weather conditions, as it identifies areas with large temperature variations. However, the studies also identified limitations, such as the difficulty in detecting I/I water in low-flow conditions and the need to monitor air temperature. The study of Beheshti et al. (2015) also reported that investment cost could be an issue. Furthermore, while the use of cameras has been identified as a promising method for detecting lateral connections in sewer systems and measuring water depth and velocity, this technique is not yet widely adopted (Lepot et al., 2017; Meier et al., 2022). CCTV and physical inspection are often seen as laborious and uncertain, due to their high degree of subjectivity (Catalfio et al., 2007; Fugledalen et al., 2021).

The tracer methods used in Kracht et al. (2008) and Heiderscheidt et al. (2022) successfully proved the stable water isotope and pollutant methods to be reliable and accurate for GWI detection in sewage networks. These methods utilizes pollutant and isotope concentration in wastewater to quantify I/I. However, it should be noted that these methods are costly and especially the stable water isotope method requires comprehensive hydrologic and hydrogeological investigations (Ellis & Bertrand-Krajewski, 2010; Kracht et al., 2008).

Flow-based methods, such as DWF and water balance, are popular for their simplicity and effectiveness (Beheshti et al., 2015; Saletti, 2021). However, these methods often rely on simplified assumptions regarding wastewater production (WWP) that limit their accuracy. In a comparison study, Mitchell et al. (2007) evaluated several common flow-based methods, including WWP-method and Minimum Flow Factor method, against more reliable alternatives, such as Dead Lowest Flow and Stevens-Schutzbach method. The study revealed the limitations of these common methods, but also proposed adjust-

ments to improve the accuracy of the WWP-method.

The use of I/I models, which are representations of the distribution of I/I in a sewer system, can estimate and assess the I/I in the system trough flow simulations (Saletti, 2021). I/I models generally require substantial amount of labor for setup and calibration. According to Ranck (2017), RDII models can help managers understand and quantify the impact of RDII on sewer system hydraulics. However, the models can become obsolete, so practitioners must maintain them to reflect system changes and evolving technology.

Pereira et al. (2019) implemented a system that integrated flow measurement devices with a SCADA system in the sewer network. This enabled for real-time flow rate measurements, consequently providing a more effective monitoring of illegal connections. Furthermore, Davalos et al. (2018) presented a methodology for analyzing and quantifying I/I through SCADA Flow Data analysis. The study determined a selection of pumping stations that required rehabilitation, which resulted in substantial reductions in I/I and considerable cost-savings.

All of the methods used to address I/I in sewer systems are subject to certain assumptions and limitations (Beheshti et al., 2015). As a result, there is no standard method that is universally applicable, and the choice of method should be based on the specific conditions and limitations of each case. The use of multiple methods in combination may help to reduce uncertainties and provide more accurate and reliable data regarding the I/I problem.

2.3 Utilizing a Combination of Methods for Municipal Implementation

By utilizing a combination of methods, it is possible to achieve an effective approach to assess the I/I in sewer networks. Where simple methods can be used to target the most pressing areas, and precise methods can validate and accurately quantify I/I sources. Specifically for this study, the selection of methods prioritizes practicality and accessibility for municipal implementation, aiming to reduce the need for expert personnel and increase the efficiency. In relation to this study's objective, the proposed methods should provide municipalities with a practical procedure to assess their own sewer networks and information to address the most pressing conditions.

To achieve this, flow-based methods offer a simple and practical approach that is highly applicable for municipalities. For example, Sola et al. (2019) successfully implemented the water balance approach, which involves comparing measured total volumes to expected volumes based on water consumption. The study further used principal component analysis, in order to evaluate the significance of a parameter subset for several wastewater treatment plants (WWTP). In the means of adapting this approach, pumping stations can be used to evaluate flow from different sewage sheds by separating them into distinct zones. By conducting basic investigations, such as examining correlations with climate conditions, municipalities can obtain the straightforward assessment they require to select the most pressing areas.

To improve the accuracy of these findings, hydraulic modeling of the sewer network can provide a comprehensive evaluation and enable comparison between situations. For instance, SWMM modeling, as demonstrated by Choi and Schmidt (2023), can simulate the sewer system and allow for a comparison of DWF to WWF with different RDII responses. Consequently, enabling the opportunity to quantify I/I similarly to the approach employed by Bentes et al. (2022), which utilized precipitation hydrographs to distinguish WWF and DWF. Additionally, measurements can be easily obtained using the method described by Van Assel et al. (2023), whereby the outflow volumes from pumping stations can be used to calculate the inflowing volumes to the pumping station.

Traditionally, generating a SWMM model involves using the modeling software's graphical user interface, but an improved and more efficient method involves using geographic information system (GIS) software to create a network from a sewer dataset (Schilling & Tränckner, 2022). However, the GIS approach has limitations in handling missing information, which requires pre-processing of the dataset before implementation. Filling in the missing information can be a laborious manual process, requiring high levels of expertise. Nevertheless, there are alternative methods for constructing a sewer data model. As demonstrated by Rokstad and van Laarhoven (2022), graph theory has been used to optimize the design of water distribution networks while considering drinking water requirements. Furthermore, the study of Turan et al. (2019) utilized graph theory to optimize sewage network design by identifying the path with the lowest cost to develop the minimum spanning tree. Both approaches showed promising results in utilizing and establishing network connectivity within their constraints. The use of this approach further enables the implementation complex shortest-path and optimisation algorithms, like Dijkstra and Prim's algorithms, to reduce the manual processes (Furqan et al., 2018; Lawande et al., 2022). A brief introduction to graph theory is provided in Appendix C.

This study proposes the use of graph theory on a sewer network provided by a municipality, to fill in the gaps of missing information and to automatically create a representative SWMM-model. The model will then be used in comparison with actual water demand and sewer flow rates to evaluate the sewer network and impact of I/I. Subsequently, quantify the extend of I/I by decomposing the measured flow rates into the distinct BSF, GWI and RDII components. The initial selection of area for investigations should be based on the simplistic methods, water balance and correlation analysis. The implementation of this framework is further elaborated in chapter 4.

3 Study Area

This chapter covers a description of Horten municipality in section 3.1, as well as the condition of the sewer system. Additionally, this section outlines the motivation behind the study from the municipal perspective, along with some of the goals of the watershed coalition. In section 3.2, a specific area within Horten is described, as it is the focus area for this study.

3.1 Horten

This case study was carried out in collaboration with Horten municipality, with the aim of quantifying I/I as information in decision support for future rehabilitation plans. Hence, the results from this study have the potential to benefit the investments in upgrading Horten's sewer network. As the results could point out the most effective locations for Horten's future investments (subsection 3.1.2).

Horten is located at the South-East Coast of Norway and has a 40 km long coastal line (Hansen, 2022). It consists of two smaller cities, Horten and Åsgårstrand, as well as three smaller districts, Nykirke, Borre, and Skoppum, as depicted in Figure 3.1 with the municipality's border marked by a red dotted line. According to SSB (2023), the municipality has a total population of 27,600 and covers an area of 70.96 km^2. As of 2022, there are a total of 13,514 housing units in the area.

According to Klimaservicesenter (2022), the municipality receives an average yearly precipitation of 952 mm, with an annual temperature averaging at 7.5°C. Nevertheless, projections indicate that by the year 2100, the annual precipitation will rise by approximately 10%, which could coincide with increased frequency and intensity of precipitation events. The municipality typically experiences mild winters in terms of temperature, and the months of October and November are described to have the most significant amount of precipitation.

Figure 3.1: Boundary of Horten Municipality (Outlined in Red Dotted Line) (Google, 2023)

3.1.1 Description of the Sewer Network

The sewer system in Horten consists of three unconnected municipal WWTPs, serving different sections of the municipality along with private septic tanks in rural areas. This study focuses on the Falskenstein treatment plant located in the city's northern part, which handles wastewater from Borre and Skoppum, pumped through pipes throughout Horten city. The central city of Horten consists mainly of combined sewer, with overflow discharging into the ocean, whereas the outer part mostly consists of separate sewer.

According to Horten, the total share of combined sewer is 13%, while the rest is separate sewer. Compared to the rest of Norway, this is a fairly low share according to Jørgensen and Rostad (2021) and Lindholm et al. (2012). Additionally, the sewer system can be described as widely spaced out with a low share of combined sewer, with a large portion of the pipes recently upgraded, as can be seen in Figure 3.2. The sewer network in Horten consists of 22,401 pipe segments, with a total length of 436 km.

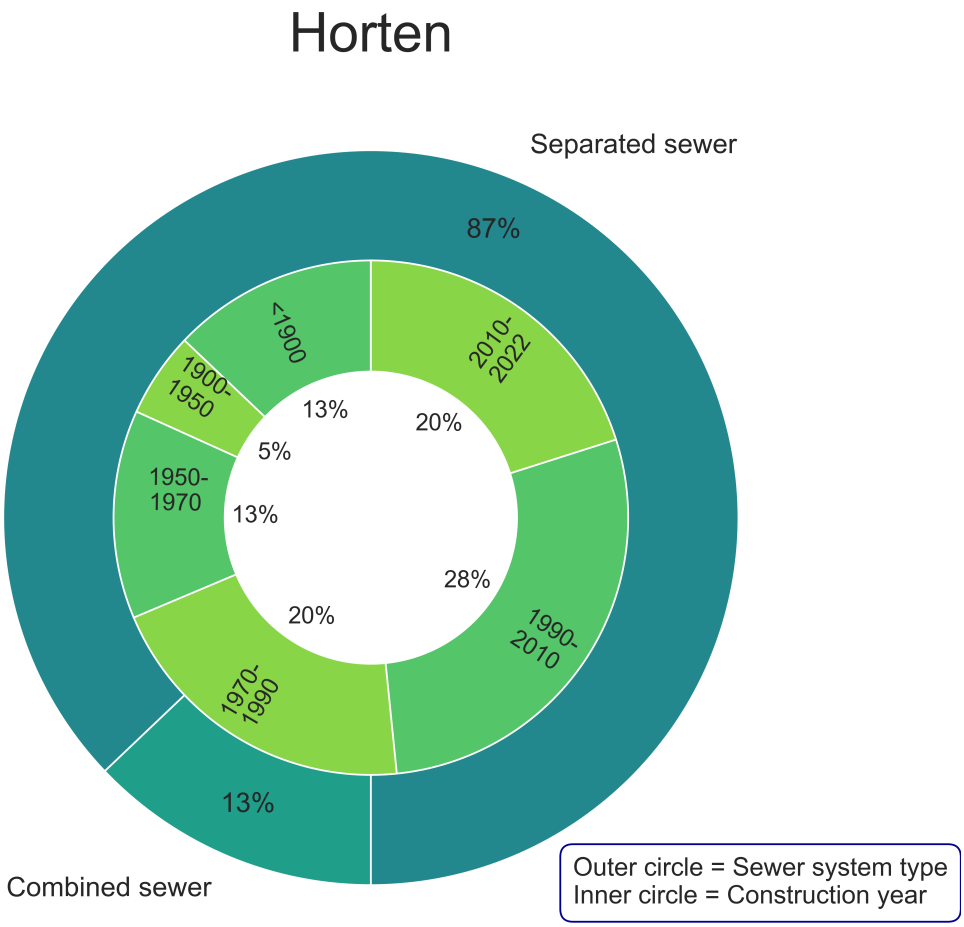


Figure 3.2: Description of Horten's Sewer System (Skagsoset, 2022)

Despite being built as separate pipes, both the sanitary and stormwater systems have intentional overflows that allow for the discharge of sewage between them, resulting in a combined sewer system during periods of high flows. This can be observed in Figure 3.3, where an intentional overflow is depicted. These overflows are typically installed in areas where it is not feasible to discharge into a nearby recipient and occur when the water level reaches the separation wall height between the systems.

Figure 3.3: Intentional Overflow in Separated Sewer

3.1.2 Motivation for Environmental Protection Measures

Horten is a member of the Vannområdet Horten-Larvik watershed coalition, which aims to meet the environmental criteria set in the regional plans ("Vannportalen", 2023). The latest 6-year plan (2022-2027) proposes 246 measures with a total cost of 2,350 million NOK to achieve these goals, including reducing annual phosphorus release by 16.5 tons to maintain ecological conditions. Whereby, wastewater is identified as one of the primary sources contributing to the excess phosphorus. Hence, upgrading the sewer network to a separate system and upgrading the existing WWTP Falkensten, or alternatively constructing a new WWTP, are critical measures proposed for Horten. As excess water caused by I/I entering the sewer system can overload WWTPs, it can cause untreated sewage to discharge directly into recipients. Accurate assessment of I/I is therefore necessary to achieve the coalition's environmental protection objectives.

3.2 Holtandalen

In this study, Horten was divided into several sewer zones, one of which was Holtandalen. This zone is located at the endpoint of a network of sewer branches, which means that it does not receive additional sewer volumes from other zones (Skagsoset, 2022). Furthermore, Holtandalen is situated in a suburban area of the municipality located inland, as shown in Figure 3.4, and is therefore not influenced by seawater. Although only 1.7% of its sewer network is combined, Holtandalen has a large number of combined manholes with separation walls for overflowing possibilities, as depicted in Figure 3.3.

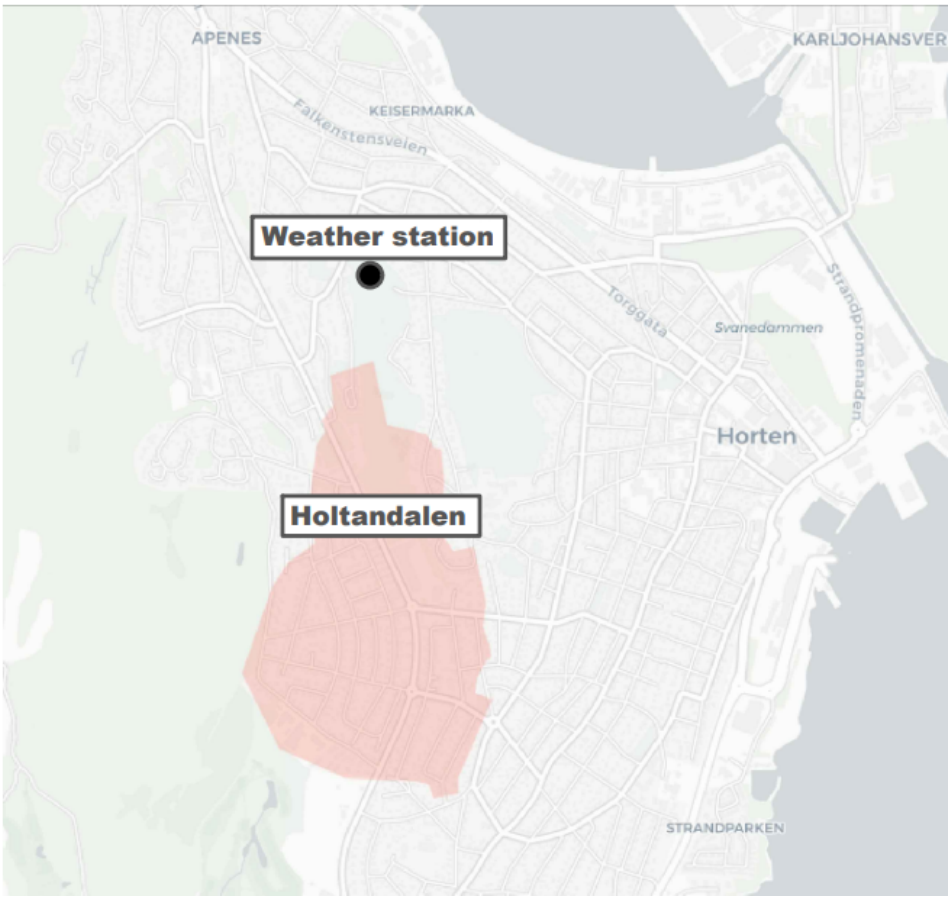


Figure 3.4: Holtandalen Sewer Zone and Horten Weather Station

In Holtandalen, the current practice of placing the sewer pipes have been at the bottom of the ditch, as depicted in Figure 3.5. This poses potential challenges and unintended consequences, such as an increased risk of infiltrated volumes in the sanitary sewer. Specifically, this arrangement creates a more accessible pathway for groundwater or surface water to enter the sewer pipes whenever the ground is saturated, potentially leading to increased volumes of infiltration. In addition, the zone is located in an area with soil that has high potential infiltration capacity, which increases the probability of undesired infiltration as the ground can store more water and thus, an increase in transported volumes (NGU, 2022).

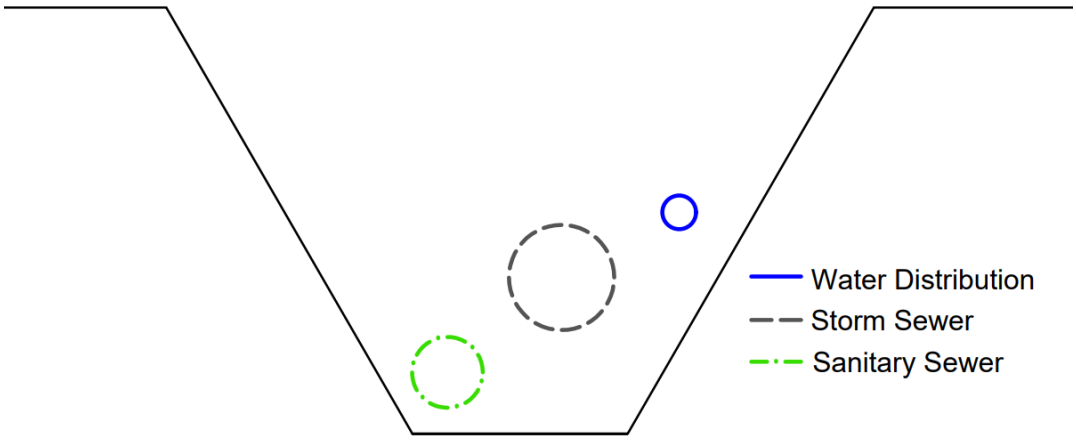


Figure 3.5: Placement of Sewer Pipes in Holtandalen

Previous research by Zaidan (2018) examined the stormwater sewer capacity in Holtandalen. A SWMM model was developed by manual imputation of missing information, incorporating sub-catchments and the stormwater sewer. The model simulated various rainfall scenarios and compared the results with data on basement flooding and flow depths. The study identified a bottleneck that contributed to localized flooding, but due to some inaccuracies it was recommended to conduct flow measurements for improved model accuracy.

4 Methodology

In this chapter, the purpose is to provide a detailed description of the specific procedures employed in the study, which aimed at quantifying the extent of I/I in a municipal sewer zone. To achieve this, several necessary steps were completed before accurately quantifying the I/I. As illustrated in Figure 4.1, this process was divided into three major phases: Preprocessing, Model Preprocessing, and Model Execution, each with its own separate sub-objectives.

The Preprocessing phase was utilized to select a suitable area for evaluation. This involved dividing the sewer system into smaller zones (section 4.1), comparing easily estimated flow rates to climate conditions (subsection 4.1.1), and considering expected flow rates in a water balance approach (subsection 4.1.2).

The objective of the Model Preprocessing phase was to create a functional sewage model that could simulate the sewage flow rate based on water demand. This phase included measuring the flow rates and distinguishing different flow rate conditions (section 4.2), utilizing the municipal database to construct a graph network to fill in missing information (section 4.3), and transforming this network into a calibrated SWMM model (section 4.4).

The final phase, Model Execution, involved running the calibrated model in various weather conditions to quantify the extent of I/I for the selected area (section 4.5). Thus, fulfilling the main objective of this study.

Python was utilized extensively throughout the study and the related scrips are included in Appendix F. Various modules were employed to facilitate data manipulation, analysis, visualization, and specific tasks. Commonly used modules such as NumPy, pandas, datetime, seaborn, and Matplotlib were employed throughout the methods to handle data structures, perform computations, manage time-related operations, create visualizations, and more. In addition to these general-purpose modules, specific modules relevant to the study were utilized for specific tasks and analyses, such as GeoPandas, Shapely, NetworkX, rasterio and SWMM-API.

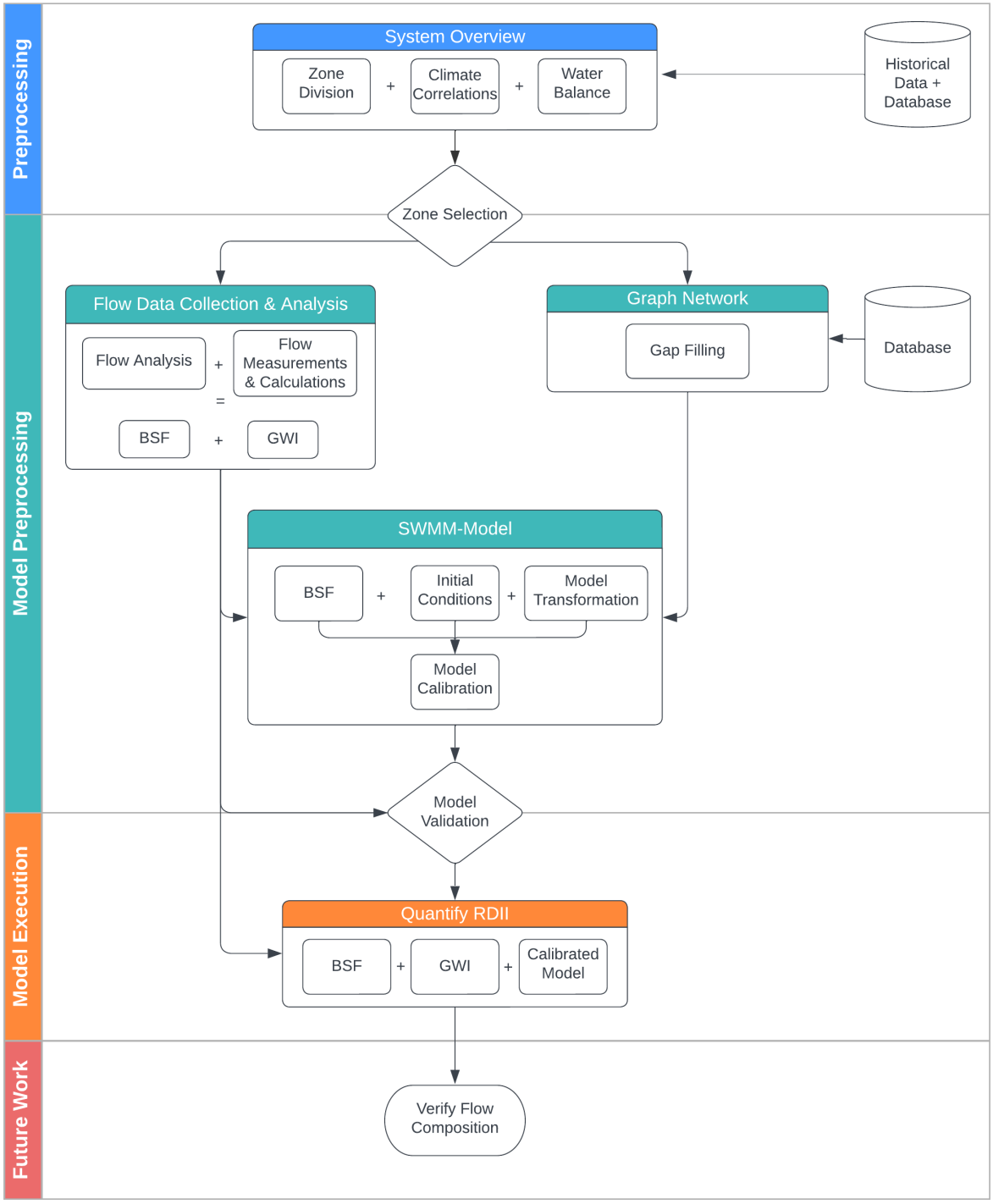


Figure 4.1: Flowchart Illustrating the Methodological Framework for Quantifying I/I (BSF - Base Sanitary Flow, GWI - Groundwater infiltration, RDII - Rainfall derived infiltration and inflow)

4.1 System Overview

The first step in evaluation of the sewer network involved obtaining an overview of the system, which was accomplished through the preprocessing phase outlined in the methodological framework (refer to Figure 4.1). This phase encompassed the division of the sewer system into distinct zones. To achieve this, the study utilized the Gemini VA platform to export sewer pipelines, which were subsequently separated manually in AutoCAD into catchments. The separation process involved attributing branches in flow directions and using either the pumping stations or the WWTP as boundary points, as illustrated in Figure 4.2.

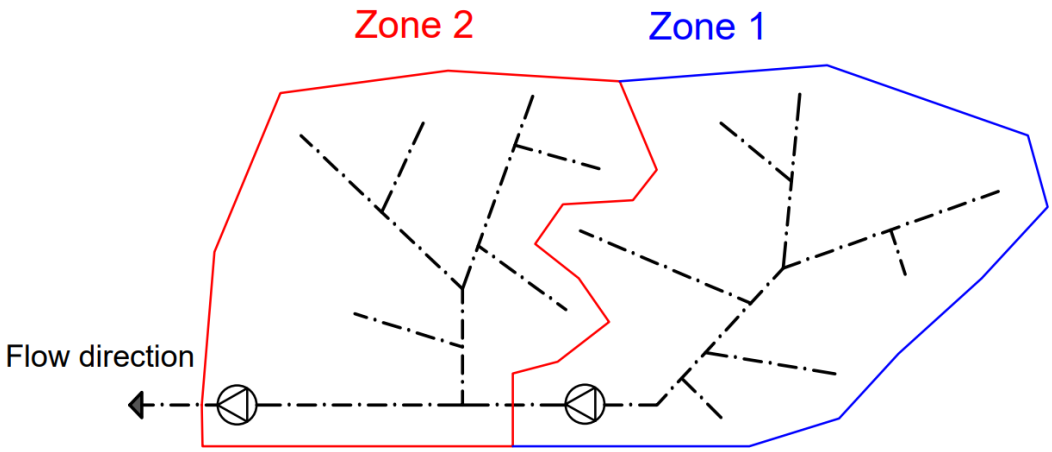


Figure 4.2: Method for Separating Sewer Zones (Skagsoset, 2022)

Pipe coordinates were then compared to the coordinates of each zone, and evaluated with GeoPandas intersection function to either include or exclude them in each zone (See Figure 4.3) (Jordahl et al., 2020). The zones were then split into separate data frames containing attributes such as dimensions, sewer type, and construction year, for each pipe within the zone.

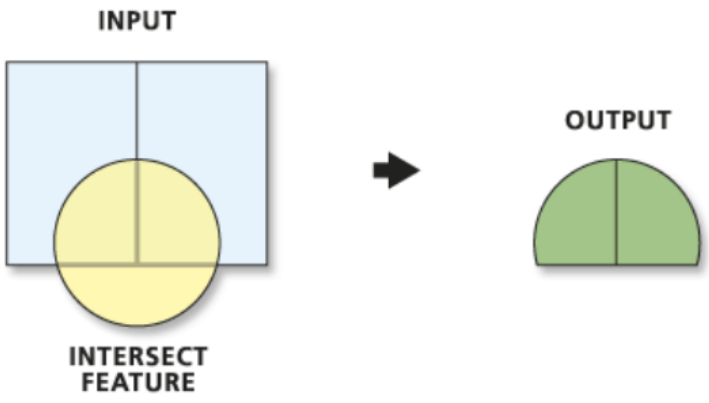


Figure 4.3: Intersection Method (Esri, 2022)

4.1.1 Comparing Pump Data to Climate Conditions

Further in the preprocessing phase, the analysis of the sewer system's response to climate conditions, including precipitation and air temperature, required comparing the flows under different impacts of these conditions. Hence, this step required a longer time series to identify trends in the sewer system. For this study, the historical data used spans from June 2, 2015, to October 27, 2022, during which daily values for the number of hours the pumps ran and the number of starts were recorded by the municipality. Nevertheless, due to limitations in data availability, the analysis of the sewer system's response to climate conditions was restricted to daily median air temperature and daily accumulated precipitation.

Moreover, as the recorded pumped volumes was influenced by the upstream sewer zones, the inflowing sewage had to be excluded from the downstream sewage flow. This was accomplished by using estimates of the pumps nominal flow rate, which was provided by the municipality, and multiplying it by the number of hours the pumps were running. However, as pumping stations typically had multiple pumps running individually or simultaneously with varying flow rates, these factors had to be taken into account during the flow calculations through a separation of these conditions. In addition, overflows within the pumping stations were also included in the calculations.

To incorporate this, the flow generated by individual pumps were calculated using Equation 4.1. Specifically, the total number of hours that the pumps have run is summed (1P), while the hours during which multiple pumps were running simultaneously (2P) and the hours during which the sewer was overflowing (O) are subtracted. Additionally, the single pump flow rate (F_1) is multiplied by the value obtained from the previous calculations while also converting the units (l/s to m^3/day). To calculate the flow generated by two pumps ($Q_{2\text{pump}}$), the flow rate of two pumps (F_2) is multiplied by the hours during which two pumps are running simultaneously, as presented in Equation 4.2. This value is then added to the flow generated by a single pump and subtracted by the flow from the upstream pumping stations (Q_{inflow}), as presented in Equation 4.3.

$$Q_{1\text{pump}}[m^3] = \left(\sum(1P_i) - 2P - O\right) \cdot F_1 \cdot \frac{3600}{1000} \tag{4.1}$$

$$Q_{2\text{pump}}[m^3] = 2P \cdot F_2 \cdot \frac{3600}{1000} \tag{4.2}$$

$$Q_{\text{total}}[m^3] = Q_{1\text{pump}} + Q_{2\text{pump}} - Q_{\text{inflow}} \tag{4.3}$$

Furthermore, the calculated flows were analyzed with data from two weather stations located in Skoppum and Horten city, as these were located at different ends of the municipality. The Pearson correlation coefficient was used as an indicator to assess the linear relationship between variables. Additionally, it provided information on the direction of the variables' linear relationship. Person correlation coefficient is further described in Appendix C.

4.1.2 Water Balance

Encompassing the last step in the preprocessing phase, the water balance offers a straightforward approach for identifying noticeable outliers in reported flow rates. As a result, it plays an important role in the system overview by facilitating the identification of areas that are particularly susceptible to I/I. In this specific case, four extended periods without precipitation were selected, each lasting between 10 and 15 days, to serve as representative instances of daily transported dry weather volumes. Utilizing the daily values calculated as described in subsection 4.1.1.

In the calculations the average value between these dry weather periods served as the transported volume for comparison. These volumes were then compared to the anticipated wastewater production, calculated using Equation 4.4, which assumes a daily production of 200 liters ($Q_{assumed}$) per person equivalent (PE) for each sewer zone in order to encompasses a wide range of cases. Additionally, the value was divided by 1000 to make the units comparable with calculated flow rates (subsection 4.1.1).

$$Q_{Dayly}\,[m^3] = \frac{Q_{assumed}[\frac{l}{day \cdot PE}] \cdot PE}{1000[\frac{l}{m^3}]} \tag{4.4}$$

4.2 Flow Data Collection & Analysis

Initiating the model preprocessing phase, a sub-objective after selecting the zones was to obtain data with higher resolution than used in the correlation analysis in subsection 4.1.1. A measurement period was therefore conducted in this zone to gather the required data, which would then be used for further analysis and calibration of a sewer model (see section 4.4). Because of the accessibility to measurement locations and available measurement devices, an Ultrasonic doppler device was used to measure the pumped volumes from Holtandalen pumping station. The device used was a FlowPulse Ultrasonic Clamp-on Flow Sensor model A-800-0396-A, with a FlowPulse Handheld as the control unit (Pulsar, 2023). For more information about ultrasonic doppler measurment devices refer to section C.3, while technical specifications of the utilized ultrasonic doppler device and control unit is provided in Appendix D.

It was determined that the optimal location for the non-intrusive measurement device, which required pressurized flow, was on the pipe downstream from the pump manifold. This pipe, which was a DN 200 stainless steel pipe, was located on a short straight pipe section near a sewer storage tank and a 30 degree bend, as presented in Figure 4.4. The measurement device was placed above the sump and pumps, as shown in Figure 4.5. At the pumping station, three pumps operated in cycles based on the sewage level in the sump. To mitigate pressure changes caused by the pump cycles changes, frequency converters had been installed. In addition, an overflow pipe was connected to divert excess water back to the sump when outgoing capacity was reached.

Figure 4.4: A Mounted FlowPulse Ultrasonic Clamp-on Flow Sensor with the Handheld Controller, Used for Measuring the Sewage Flow Rates.

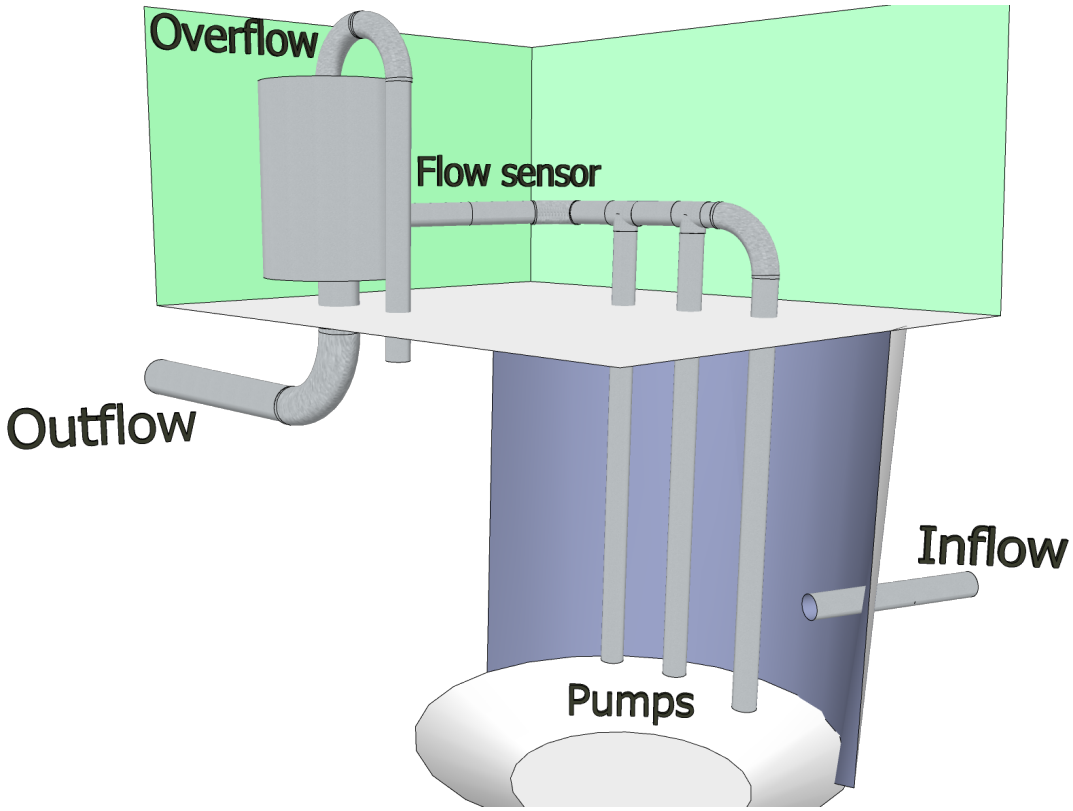


Figure 4.5: Visualization of the Holtandalen Pumping Station and the Flow Sensor Location

Prior to the measuring period, the device was configured by adjusting the specific parameters described in Table 4.1. A calibration factor was set to account for turbulence caused by the short distance to disturbances (bend, pumps, tank). The device was calibrated to match the reported nominal flow rate of a single pump. To validate the flow measurements, the flowmeter was initially tested at another pumping station that already had an electronic flowmeter (Krohne Optiflux 2000F, DN 200) installed. The measurements of the two devices were compared to ensure accuracy of the Ultrasonic Flow sensor.

Table 4.1: Parameters Configured on the Ultrasonic Doppler Device

Parameter	Value	Notes
Sensitivity	5	Signal sensitivity (0 - 10)
Damping	3	Smoothing fluctuations (0 - 10)
Calibration Factor	0.7	70% of the measured flow (0 - 1)
Pipe Internal Diameter	200	Used for flow calculations (mm)
Pipe Type	1	Adjustment factor (Stainless steel)
Density	2	Effluent flow (1 - 4)
Flow Unit	Liters per second	Logged units

During the 40-day measuring period from March 3rd to April 12th, the device was set to record flow data every 30 seconds. This recording frequency was chosen to ensure that the local storage capacity was not exceeded, while also providing data with sufficient resolution for calculating inflow to the pumping station (see subsection 4.2.1). At the end of the measuring period, the logged data was exported as a CSV file to a computer for the further calculations.

4.2.1 Calculation of Inflow to the Pumping Station

To further analyze sewage flow and responses from weather events, it would require knowledge about the inflowing volumes to the zones pumping station. As this data was not possible to acquire for this project, the measurements conducted in this study has the purpose of extrapolating this information, in the sense of back calculating the sewage flow. The measured volumes, from the ultrasonic device, had to be used in combination with the recorded water levels of the sewage sump, in order to calculate inflowing volumes. The sump levels were exported from Gemini VA with a minute resolution. The sump geometry consists of a circular cross-section with a narrowing section at the bottom, as visualized in Figure 4.5.

According to Van Assel et al. (2023), the inflowing volumes to a pumping station can be calculated by separating the data into different cycles: filling and emptying cycles. Missing data and changes between cycles had to be excluded from the flow calculations to prevent errors. This proposed method can be simplified to the mass balance approach presented in Equation 4.5, similarly to the liquid level procedure described by Coughanowr and LeBlanc (2009). Here, Q_{inflow} is the inflow to the pumping station, $Q_{outflow}$ is the measured outgoing flow, h is the reported water level of the sump, and A is the cross-sectional area of the sump. Where the cross-section was assumed to be constant, as the water level was not expected to reach the narrow section of the bottom in this case (see Figure 4.5).

$$Q_{inflow} = \frac{dh}{dt} \cdot A + Q_{outflow} \tag{4.5}$$

The calculation method involved iterating through the reported water levels and categorizing the cycles based on the measured flow for each time step. For the filling cycles, the sewage flow was calculated between every time step, and for the emptying cycles the initial water level was compared to the water level of the cycle end, this was then added to the pumped volume of the entire cycle and then divided by the duration of the cycle. Providing a constant average flow rate throughout the filling cycle. Missing values and values between cycles were linearly interpolated between cycles. A more detailed description of the calculation procedure can be seen in Appendix B.

4.2.2 Analysis of Wastewater Flow Rates

Concluding the final step of the flow Data collection and analysis within the model preprocessing phase, it was essential to differentiate between various flow rates. Therefore, once the sewage flows were calculated, they were classified by the respective weather conditions of each day to simplify data analysis. This was because there is a differ-

ence in sewage flow between weekdays and weekends, as well as between dry- and wet weather.

Additionally, as mentioned in section 2.2, WWP-method is a well established method for calculating rough estimates of the GWI. In the WWP-method, the GWI is calculated as the difference between the daily average DWF and the daily average WWP, which is determined by subtracting the minimum night flow (MNF) from DWF and dividing the result by a fraction of WWP (X), as shown in equations Equation 4.6 and Equation 4.7 (Mitchell et al., 2007). Further explanation of DWF, GWI and MNF can be seen in Appendix C

The value of 'X' is based on the assumption that the minimum water usage accounts for approximately 12% of the total daily water usage (Mitchell et al., 2007). Hence, 88 % is the fraction of the daily average DWF generated by wastewater, where 12 % of the daily average DWF occurs at the MNF. Nevertheless, the findings of Mitchell et al. (2007) suggests that the value for X should be lower than this, and that it overestimates the GWI for large catchments and underestimates for smaller catchments. It is suggested to be within the span of 70 % to 75 %. Thus, it was set to 75 % for this study as it was utilized for a small catchment.

$$WWP = \frac{DWF - MNF}{X} \tag{4.6}$$

$$GWI = DWF - WWP \tag{4.7}$$

To validate the calculated GWI values, these were compared to the daily minimum rate of water consumption, specifically the minimum water demand. In this study, this comparison method is referred to as the Minimum Night-Flow Method (MNF-method). The MNF values were estimated by determining the daily minimum demand (generally around 4 AM), and then adjusting the daily lowest wastewater flow rate to match this. Hence, assuming that water consumption was equal to the WWP, where the deficit at this time was attributed to a constant GWI for each day.

4.3 Creating a Sewage Network Graph

Organizing network information was necessary in order to fill in missing data and to facilitate the creation of a sewage model. According to Meijer et al. (2018), Graph Theory can be utilized for this purpose, as it defines the connectivity through vertices and edges. Referring back to the framework of this study (see Figure 4.1), this step provides a preprocessing of the database before it can be transformed to a SWMM model.

In general, due to frequent missing information, Norwegian sewage datasets are often incomplete, making it difficult to fully define the connectivity of a graph using the data alone. This is also true for Horten, where pipe and manhole exports from Gemini VA were used as the primary dataset for graph generation. To overcome the issue of missing

information, proximity analysis was conducted using the buffer zone functionality of GeoPandas (Jordahl et al., 2020). Connections were defined by creating buffer zones around the pipe lengths and ends, by using GIS coordinates and the Shapely module's buffer functionality (Gillies et al., 2023). The resulting geometries were used to identify connections where they overlap, as shown in Figure 4.6. Furthermore, the documented connections were also incorporated into the process of constructing the base graph, in which each vertex represented a manhole or a theoretical node with its corresponding downstream pipe, and each edge represented a connection to that manhole or pipe. In the transformation to a directed graph, the flow direction of each pipe was taken into consideration, which was based on the start and end coordinates of the pipes. Additionally, pipes that were found to have different flow directions than their adjacent pipes were flipped accordingly.

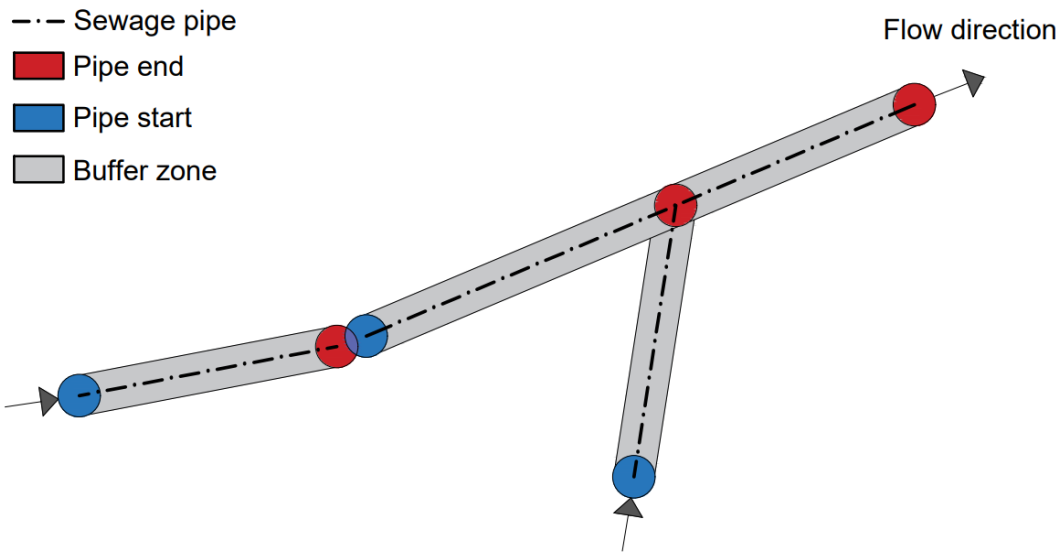


Figure 4.6: Buffer Zones Around Pipe's GIS-Coordinates

Using the directional graph, the root node could be easily identified using the NetworkX algorithm "ancestors" (Hagberg et al., 2008), which determines the vertices upstream of each vertex. This involved sorting the pipes by the number of vertices upstream and selecting the vertex with the most vertices as the root node. The root node was then used to filter its connected vertices. Knowledge about the root node and having a directed graph opened up the possibility for gap filling, which is described in subsection 4.3.1. Furthermore, the directional graph could then be converted into a minimum spanning tree using only the paths producing the shortest distance to the root node, a methodology similar to that of Turan et al. (2019). In this case, only the shortest path was kept by disconnecting the edges producing longer paths to the root node. The minimum spanning tree was further used to create a sewage model, described in section 4.4.

4.3.1 Gap Filling with a Sewage Network Graph

In a sewage dataset, it is common to have missing information for multiple parameters, such as diameter, length, material, roughness, elevation, and manhole depth. To fill in these gaps, this study employed several methods that takes advantage of the connectivity

of the graph.

- To fill in missing **diameters**, the largest upstream pipe diameter was used as the default value. If the maximum diameter of the upstream pipe was less than the minimum acceptable diameter (110 mm) or if there was no upstream pipe available, then the downstream pipe diameter was used.

- The missing **length** was calculated by applying the Pythagorean theorem to the start and end coordinates of the pipe.

- To handle missing values in the **material** variable, a method inspired by mean imputation (Kabir et al., 2020), referred to as 'most common variable' imputation, was employed. This approach involved determining the most common material type based on pipes with the same diameter in the network.

- To estimate the missing **elevations**, a Digital Terrain Model (DTM) provided by Kartverket was utilized (Kartverket, 2023). The DTM supplied elevation data for the corresponding node coordinates, and the rasterio module was employed to extract the necessary elevations (Gillies, 2023).

- When determining **manhole depth**, the elevations and depths from upstream and downstream nodes were utilized through a linear interpolation. Additionally, to ensure the validity of the depth, the height difference between the pipe's elevations was checked against the maximum and minimum elevations, as a minimum depth of 0.5 meters (d_min) and a minimum of 10 ‰ slope from the downstream node (d_max). If the height difference exceeded these limits, it was set to the limit. This method is visualized in Figure 4.7

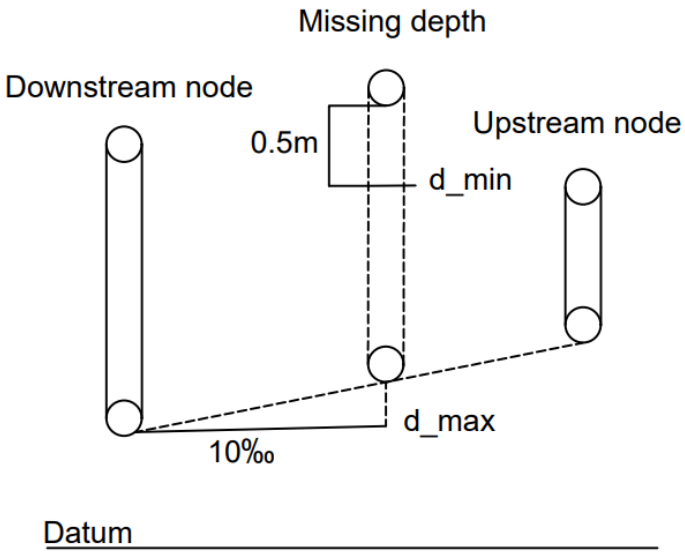


Figure 4.7: Visualization of Method for Filling in Missing Node Depths, Where d_max and d_min Represents the Depth Limits

4.4 Transforming the Graph to a SWMM-Model

The combination of the flow measurements and the graph network was used to create a sewage model. As visualized in Figure 4.1, this step would conclude the model preprocessing phase and its objective. Furthermore, to evaluate how well the sewage system performed under different operating conditions, such as varying flow rates, wastewater loads, and I/I conditions, EPA-SWMM was used to create this model and simulate the system (Rossman & Simon, 2022). Although primarily designed as a stormwater tool, EPA-SWMM is also useful for sewage system evaluation and is available as a public domain software, making it easily accessible for this study.

Transforming the network graph into a SWMM-model was therefore a necessary step that would enabled simulations of wastewater flow through the network and the evaluation of the system's performance. The process involved defining network components, characterizing wastewater flow, and calibrating the model to ensure accuracy. In this study, the SWMM-API module (Pichler, 2022) was utilized as a method for constructing an EPA-SWMM input file, providing the necessary capabilities for simulating and evaluating the sewage system's performance.

In addition, a simplification was implemented to facilitate the transformation of the graph into a SWMM-model, particularly concerning lateral connections. The simplification involves redirecting the flow from lateral connections to pass through the entire pipe to which it is connected. By employing this approach, the desired lateral connections can be incorporated without the necessity of creating additional nodes and splitting of pipes. Specifically, without this simplification, the model would have required a significantly larger number of nodes and pipes. This is because the pipes would have needed to be cut at each connection point to create a new node at that location.

4.4.1 Initial Conditions

To initialize the SWMM-model, it was necessary to set the initial conditions for wastewater flow in the network model. These conditions consist of defining a flow pattern with its base value component and assigning values for pipe roughness throughout the model. For the initial state of the model, only DWF was included as flow input to the model.

To determine the distribution of household connections, a dataset of registered persons addresses was provided. Using the central coordinates of each house, the number of PE was assigned to the closest pipe within a certain distance through a proximity analysis with a buffer zone, similarly to the method described in section 4.3.

The flow dynamics in the sewage zone were simulated by uniformly distributing the daily reported water demand, which is summarized in Figure 4.8. Specifically, a time series was created as input for each node with associated PE, providing flow throughout the selected time period for the simulation. The values in the time series were based on the pattern from the daily demand curve, with a universal base value multiplied with the number of PE. The initial base value for the daily production was set based on the assumption that each PE produces 140 liters of wastewater per day. Additionally, in order to accurately simulate the wastewater flow, the model was pre-run five days before starting to report flow rates, these days were run with the pattern of the initial day of

the simulation.

Lastly, the roughness was set based on the material types of the pipes. The Mannings coefficients from Engineering ToolBox (2004) was used as the initial estimates of the distributed parameter value for roughness.

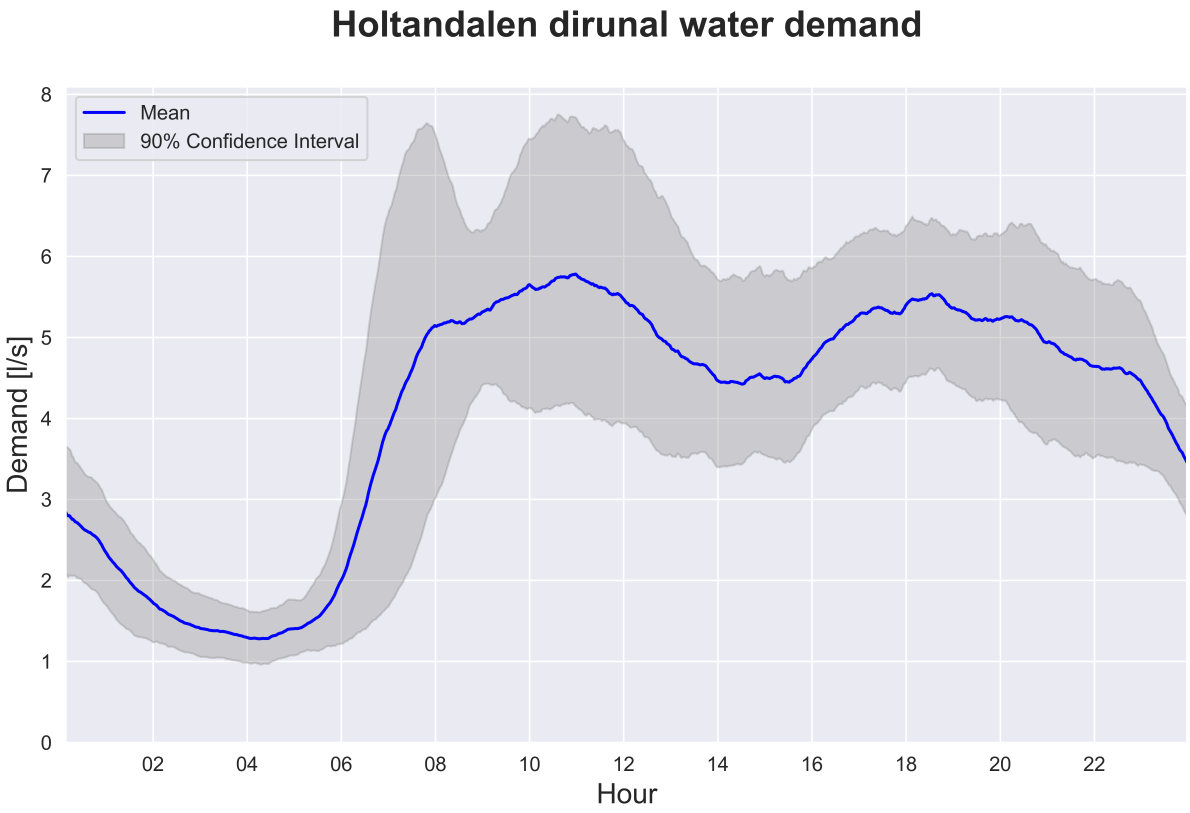


Figure 4.8: Holtandalen Diurnal Water Demand, Excluding Weekends (3rd of March to the 12th of April 2023, Confidence Interval = = 5th to 95th percentile)

4.4.2 Model Calibration & Validation

Following the establishment of the initial conditions of the SWMM model, it was necessary to calibrate the model specifically to the sewer zone. Model calibration generally involves adjusting its parameters to improve accuracy and reliability. In this case, the calibration process involved modifying the base value for the input time series and the pipe roughness. Regarding the base value parameter, it was adjusted for each node based on their assigned PE values. Essentially, a universally calibrated value was multiplied by the respective PE to fine-tune the base values. As a result of these adjustments, the flow rates in the model were tuned while maintaining the same flow pattern as provided by the water demand data. To address the roughness parameter, the initially set values served as the baseline during the calibration process. A universally applied multiplier was used to adjust the roughness, ensuring a distributed roughness across the entire system rather than calibrating a single value. This approach aimed to maintain consistency and uniformity in the roughness parameter adjustments throughout the model.

In order to accurately evaluate the model and make precise parameter adjustments, the

calculated flow rates from subsection 4.2.1 were used as a benchmark for the simulations. Thus, by only providing the water demand time series as input, the calibrated model would be capable of predicting wastewater flow. However, as mentioned in subsection 4.2.2, there was an anticipated GWI that needed to be excluded from the flow measurements to ensure an accurate comparison between the simulated and measured sewage flow rates. Since the GWI was expected to vary over time, the MNF method was utilized to calculate its value for each day, as it demonstrated less subjectivity compared to the WWP-method. Subsequently, the GWI was subtracted from the daily flow measurements accordingly, and then compared to the simulated flow rates from the outlet node of the model.

As described by Bennett et al. (2013), visual performance analysis of parameters is a suitable method for calibrating a chaotic system, with the primary objective being to capture the pattern rather than precisely reproduce the observed time-series. However, instead of relying solely on visual adjustments of the simulation curve, multiple goodness-of-fit criteria were utilized. This approach facilitated a more precise comparison of parameter adjustments. Nash-Sutcliffe Efficiency (NSE) and Mean Square Error (MSE) were selected as the goodness-of-fit criteria in this case, primarily due to their well-established nature. Additionally, MSE was chosen for its ability to provide a comparative measure of residuals across simulations, while NSE indicated the model's ability to predict variance in the observations (Bennett et al., 2013). A further description of MSE and NSE is provided in section C.5, with their respective mathematical expressions.

The time period for model simulation was determined by utilizing the initial dry weather period of the measurements, and by visual inspection, the days with the least amount of noise was selected, which further involved splitting the time series into two separate segments. These periods were March 8th to March 10th, 2023, and 16th of March, 2023. Additionally, these periods were chosen based on their comparable time series, which exhibited minimal errors in both the measurements and water demand. These selection criterion made them the most suitable time frames for achieving accurate representations of the BSF component of DWF. Moreover, the model was calibrated using the initial period and subsequently evaluated using the latter period. This evaluation served as a performance assessment to verify the validity and accuracy of the model.

4.5 Quantify I/I

To address the objective of the model execution phase, and the main objective of this study, the calibrated sewer model for the Holtandalen sewer zone was employed to accurately quantify the I/I within the corresponding area, serving as the final phase in the methodical framework visualized in Figure 4.1. This involved executing the model throughout the measurement period, where the simulated flow rates provided wastewater estimates, and the deficit to the measurements was assumed to be contributed by I/I, which consisted of a combination of RDII and GWI.

Furthermore, to distinguish between these components, a DWF period was employed similarly to the study of Bentes et al. (2022). In this case, it was assumed that an 8-day dry period, with no precipitation, preceding the measurement period was sufficient

to represents the flow as DWF. During this period, the MNF-method, in combination with simulated flow rates, was utilized to estimate BSF and GWI values for each day. However, for WWF periods, this method was not applicable. Consequently, the GWI was extrapolated from the most recent known day with DWF. Regarding the RDII component, it represented the difference between the total daily measured flow and the combined value of the extrapolated GWI and simulated BSF.

Additionally, to verify the results, the simulation was executed on specific days with WWF, and subsequently compared the simulated results with the measured data. As the model setup had already been established, this step only required implementation of a time series representing the water demand for the given date and plotting the results for a direct comparison. The assessment of this step was conducted through visual inspection of the results, considering factors such as precipitation. The selection of these days, specifically the 17th and the 21st of March 2023, was to ensure a comprehensive evaluation of the system's performance during different WWF conditions.

5 Results & Discussion

This chapter presents the outcomes of the study, which aimed to achieve three main sub-objectives: (1) selecting a suitable area for evaluation (section 5.1), (2) creating a functional and calibrated sewage model that represented the selected area, and (3) running this model in various weather conditions to quantify the extent of I/I (section 5.5). The following sections detail the findings obtained through the completion of these sub-objectives and discuss their possible interpretations and implications. Similar to the structure depicted in Figure 4.1, the results leading to the fulfillment of the second sub-objective consist of three key steps: obtaining wastewater flow rates (section 5.2), creating a sewer network graph (section 5.3), and transforming the graph into a SWMM model (section 5.4).

Furthermore, this study builds upon the findings of a project thesis conducted in the fall of 2022 (Skagsoset, 2022), which laid the foundation for the current study as the project thesis primarily focused on the preprocessing phase. Hence, the results from the project thesis are incorporated in section 5.1, and the project thesis is included in Appendix E.

5.1 System Overview

As an initial step in the preprocessing phase, the sewer system was divided into 21 distinct zones, as visualized in Figure 5.1, based on the WWTP and pump locations. The majority of the outgoing connections from these zones were single pipelines, transporting sewage to downstream zones, making the zones easily separable. However, separating a few zones was more challenging due to connections that had multiple downstream paths to different pumping stations. To determine the flow path, an evaluation of the elevations was necessary to define the connection. As a result, a few zones would have a bypass to the next downstream pumping stations whenever a certain threshold was reached in these connections. Nonetheless, due to the location of these connections, it was not expected to significantly influence the results, especially since the locations were generally high up in the system and did not affect the majority of the flow. Therefore, the flow of sewage passing through these pumping stations was considered representative of the volumes generated by the sewer zones in the system for this study.

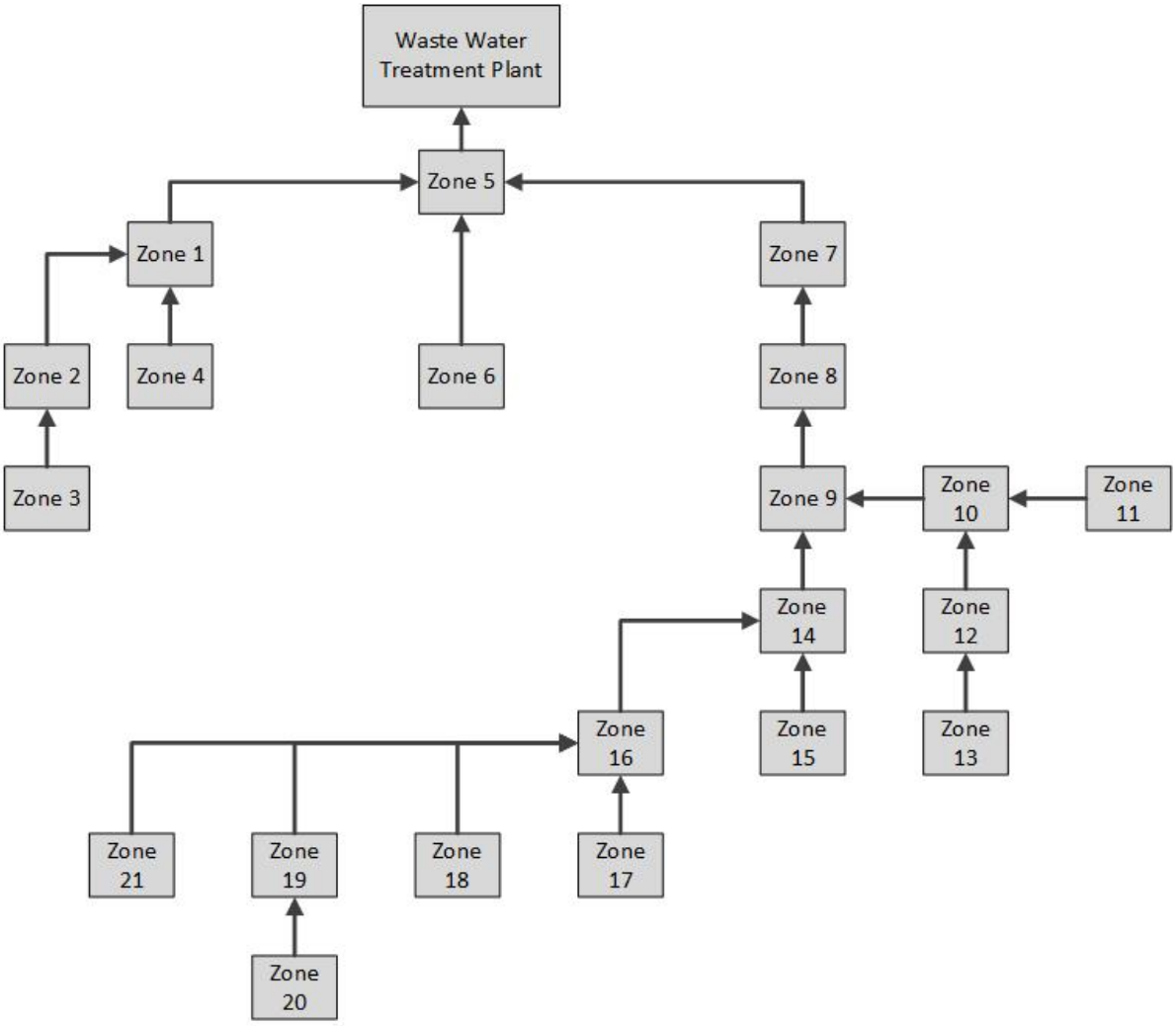


Figure 5.1: System Overview, zone names are not shown to prevent sensitive information disclosure about Horten's sewer system (Skagsoset, 2022).

Overall, dividing the sewer system into distinct zones was a crucial part of the preprocessing phase to gain a better understanding of the overall system. This step provided insights into the scale of each sewer zone that contributed sewage to the pumping stations. As a co-benefit for the municipality, the creation of the flowchart visualization has simplified the understanding of Horten's large-scale sewer system, providing an accessible overview for new employees and anyone working with the system. This method of structuring the system and visualization has greatly improved the accessibility of the system's overall structure and flow, making it easier to understand and manage.

5.1.1 Climate Correlations

Wastewater flow rates were calculated and compared through correlation investigations, as a part of the preprocessing. The aim of these investigations was to explore the relationships between various parameters affecting the flow of sewage in the studied system. Specifically, analyzing the correlations between sewage flow, precipitation and temperature, as described in subsection 4.1.1. The results from this correlation investigation are presented as a heatmap, displayed in Figure 5.2, where a dark red color represents a positive correlation and a dark blue color represents a negative correlation. The pumping stations are indicated in the figure with their corresponding share of combined sewer, while in some pumping stations this information was not available and are marked as (-). Additionally, a map in Appendix A visually represents the correlation strength of major zones, utilizing color strength to provide insight in the location of these occurrences.

Furthermore, due to limited data availability, certain pumping stations did not provide data of sufficient quality and quantity for analysis. This was due to a large degree of errors in the data or significantly shorter time series that were not compatible with the rest of the pumping stations. While some pumping stations did not provide adequate data for analysis, data from other minor pumping stations within a few of these zones were easily available. This data provided valuable information about the flows generated within these zones and, to a certain degree, still represented effects of the sewage shed. These zoned does not include representative share of combined sewer, but are marker with a capital letter after the zone number. Therefore, some zones were excluded from the study, while other minor pumping stations were included to examine their effects. Consequently, the dataset used for the correlation analysis does not precisely align with the initially defined sewage zones.

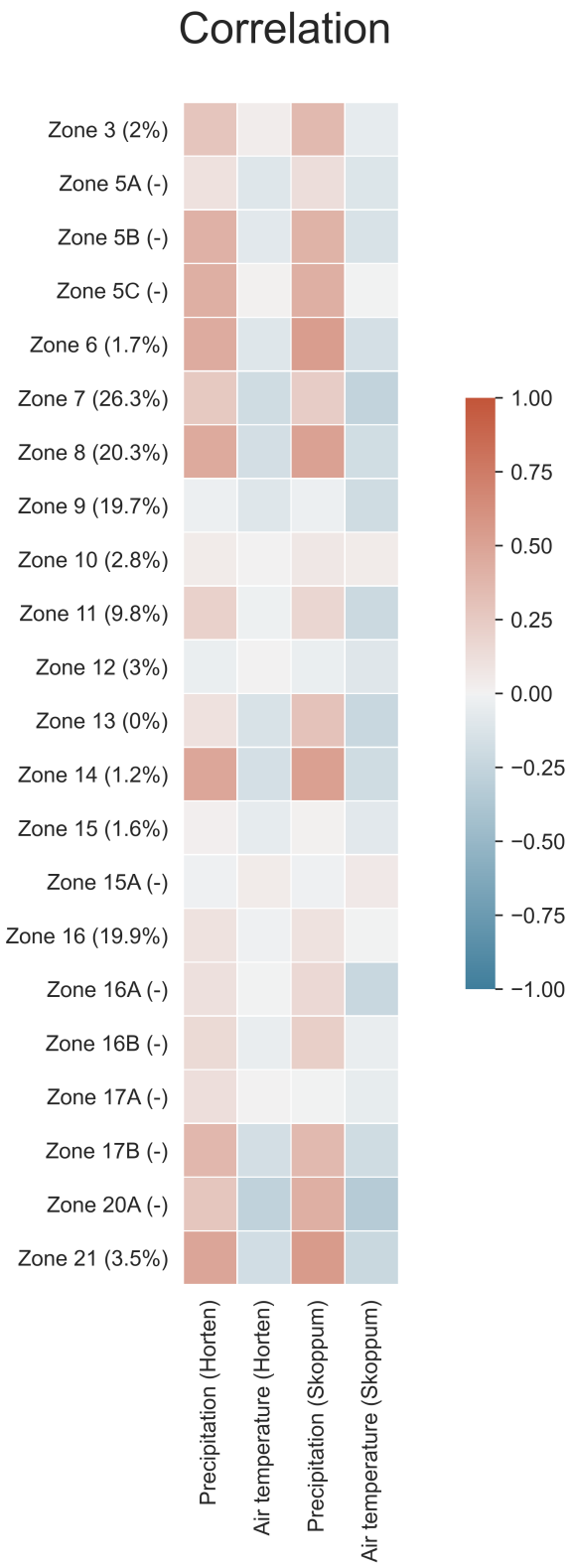


Figure 5.2: Heatmap Showing Person Correlations Coefficients, using historical data from the 2nd of June 2015 to the 27th of October 2022 (Skagsoset, 2022)

The negative correlation observed between air temperatures and flows is not entirely clear, but it is possible that seasonal variations may play a role. In Horten, where the dominant rain season occurs in the fall, increased infiltration could contribute to

the observed impacts of I/I in the system. Additionally, during the spring when soil thawing and snow melt is prevalent, there is an expectation of a significant increase in transported volumes. In contrast, during the summer months when temperatures are higher and rainfall is generally lower, the influence of I/I may be less pronounced. This could mean that the flows correlates with seasonality, rather than air temperature alone. Thus, a weak negative correlation may align with the expected rainy season in the fall and the snowmelt in the spring.

Upon analyzing the correlation between flows and precipitation, several interesting relationships were identified, although less than expected. As shown in Figure 5.2, there was a generally positive correlation between precipitation and flow, which supports the anticipated influence of RDII in the system. This correlation was noticeable in zones with a high share of combined sewer ($> 10\%$), where RDII was expected to have a significant impact on the transported volumes. The results revealed only correlation coefficients below 0.5, which may seem lower than anticipated for zones with high shares of combined sewer. However, it is important to note that other factors not considered in this study influence these flow rates. For instance, peak cut-offs resulting from sewage overflows in the largest zones with combined sewers could prevent the detection of stronger correlations, whereas other factors such as infiltration by high groundwater tables or tides could also have an impact on the flow rates.

Furthermore, as can be seen in Figure 5.1, a significant amount of sewage in Horten flows through multiple zones before reaching the WWTP. As a result, the sewage volumes are recorded multiple times as they traverse the system, which increases the complexity of flow evaluation and introduces uncertainty in several zones. Hence, flow results from zones with no upstream influence have a less degree of uncertainty and are generally more reliable. For the majority of zones, sewage production from upstream zones had to be excluded from the calculation to prevent the flow produced in these zones from being obscured by inflowing volumes. By using the calculation based on the nominal flow rate and the hours pumps were running, the results provided reasonable daily volumes and method worked as intended. However, as sewage takes time to move through the system and the municipality only provided daily values, this approach proved problematic for zones with large distances between them. This issue was particularly evident in two large sewer zones, Zone 9 and 16, where the correlation coefficients were close to zero despite their high shares of combined sewer.

The zones with a low percentage of combined sewer ($< 10\%$) are particularly interesting as they indicate areas that may be susceptible to unintended sources of RDII. Identifying and quantifying these sources would be of significant interest to the municipality, as their focus is on reducing such occurrences. As shown in Figure 5.2, Zones 6, 14 and 21 emerged as potential candidates for further investigations due to their relatively high correlations with precipitation.

While it's important to note that correlation doesn't necessarily imply causation, in this particular system there were documented reports of the influence of rainwater, and there were already established expectations of a correlation between flow and precipitation. Therefore, a strong correlation between these two factors in this case could indicate where I/I are most likely occurring within the system. Overall, the correlation analysis presented in this study provides valuable insights into the relationship between precip-

itation, seasonality and flow in the sewer system. Providing the municipality with the co-benefit of identifying some of the zones that are most susceptible to I/I.

5.1.2 Water Balance

In this study, the water balance approach was employed to compare the expected volumes with the measured volumes. The approach provided a straightforward method for identifying outliers in the transported volumes through the calculation procedure described in subsection 4.1.2. However, due to the sensitivity of the information regarding the sewer system, the specific comparison results are withheld to avoid disclosing potentially harmful details.

The comparison faced several challenges primarily due to data quality issues concerning PE. The available data only accounted for the number of inhabitants and did not consider BSF from industrial sources. Consequently, there were gaps in the information, particularly in areas with substantial sewer flows but no population. Hence, information about Zones 9 and 14 were not comparable with their measured volumes. Moreover, the Zones 7 and 8 exhibited significantly higher pumping volumes than the estimates indicated. Subsequent discussions with the municipality unveiled that these zones experienced considerable infiltration of seawater, thus confirming the initial findings. Furthermore, the investigations identified additional zones with noteworthy discrepancies between pumped volumes and expectations, suggesting a potential susceptibility to I/I. Horten confirmed that Zones 6 and 21 were indeed generating more flow than anticipated. These findings provided the municipality with the additional benefit of reassurance, as it identified these known issues as the largest outliers in their system.

5.1.3 Zone Selection

To accomplish the objective of the preprocessing phase, a specific sewer zone was selected for further measurements and modeling. This would enable the possibility to focus on simpler systems and increase the precision for quantifying I/I, providing the intersection between the preprocessing and model preprocessing phases (refer to Figure 4.1). The criteria for zone selection were based on accessibility to the zone's pumping station, the complexity of the sewer network, and the correlations with climate conditions. In this case, simpler systems with no upstream inflow were given higher weighting to simplify the study.

Based on the correlations presented in Figure 5.2 and Table 5.1, Zones 6, 14 and 21 emerged as promising candidates for further investigation, considering their correlation with precipitation. However, Zone 14 presented greater challenges due to its predominant industrial connections, making it difficult to compare expectations with measurements and predict variations accurately (subsection 5.1.2). Consequently, Zone 14 was not chosen for further work. Zones 6 and 21, on the other hand, demonstrated similarities in terms of size, correlation, accessibility and the presence of combined sewer systems. Ultimately, Zone 6 was selected as the focus area for measurements because it solely comprises residential consumers and represents a sewer system that is slightly simpler compared to Zone 21. Hence, a specific zones was selected for further analysis, namely Holtandalen, which described in section 3.2.

Table 5.1: A Selection of the Most Interesting Correlation to Precipitation (Horten)

Pumping station	Correlation to precipitation
Zone 6 (1.7%)	0.45
Zone 14 (1.2%)	0.47
Zone 21 (3.5%)	0.49

5.2 Flow Data Collection & Analysis

The model preprocessing and execution involved comparing a model to real data, which required the collection of field measurements. Consequently, flow measurements were conducted at the Holtandalen pumping station, as a part of the model preprocessing phase. This was measured as described in section 4.2, and was conducted over a period characterized by significant variability in the sanitary sewer flow. This variability was influenced by the climate conditions during the period, which included a combination of snowmelt and rainfall, as well as an initial dry period with temperatures below zero. The measurements were conducted with minor interruptions, ensuring continuous data collection. Notably, a maintenance service was performed on the pumping station, and a data integrity verification was carried out within the period to confirm the logging of data. Utilizing the collected data, the inflowing sewage rates were reverse engineered, providing insights into the flow coming from the connected sewage shed. However, it is essential to acknowledge that within the dataset, a total of six days was identified as having faulty measurements. Consequently, these days were excluded from the analysis to uphold the integrity and reliability of the dataset.

The flow rates were calculated as described in subsection 4.2.1 and were based on liquid level calculations described by Coughanowr and LeBlanc (2009) and the procedure of Van Assel et al. (2023). The results from these calculations are shown in Figure 5.3, including the recorded precipitation, snow depth and air temperatures throughout the flow measurement period. Upon analyzing the figure, it became evident that the flow rates demonstrated an upward trend when temperatures became positive. This finding suggested that the presence of the snow pack in Horten had a significant impact on flow rates as a result of snow melt. Furthermore, the flow rates displayed an increase during days with rainfall, aligning with initial expectations of correlation to precipitation (see subsection 5.1.1). Moreover, it is worth noting that the flow rates remained relatively stable at the start of the measurement period, characterized by subzero temperatures and an absence of rainfall for a duration of eight days before the measurements. This observation indicated that the flows recorded in this period corresponded to DWF.

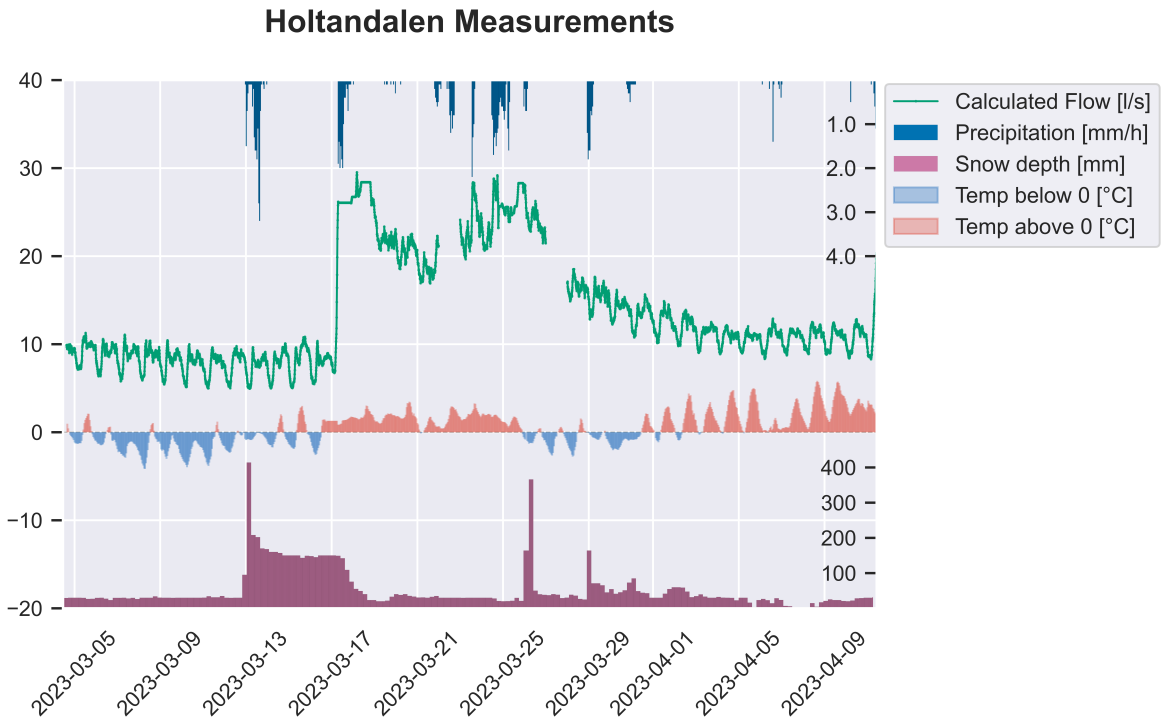


Figure 5.3: Comparison of Flow Measurements at Holtdandalen with Weather Station Data (Horten), from the 3rd of March to the 12th of April 2023. The plot shows the relationship between flow rate and air temperature, with both variables displayed on the left axis. Precipitation and snow depth are represented on separate axes displayed on the right side. For a more detailed view of the plotted values, please refer to Appendix A.

As the measurements were limited to a specific location, it subsequently imposed constraints and introduced uncertainties to the results. Spesifically, due to the uncertainty regarding the device's placement and the presence of flow disturbances caused by nearby equipment, the flow rates were calibrated to align with the nominal flow rate. This calibration process resulted in a more reliable estimation of flow dynamics. However, it should be noted that the pumped volumes in this particular case may not precisely align with these theoretical rates, as these are just theoretically estimated flow rates. Nevertheless, the theoretical flow rates provided a good estimate of the flow rate domain. Hence, a validation of these values could further improve the reliability of the measurements.

Furthermore, it is important to acknowledge the presence of uncertainty in the flow calculations due to overflowing volumes, which were not considered in these calculations and measurements. Accurately calculating the flow rate posed a challenge due to the disparity between the recorded sewer overflowing volumes, measured in hours per day, and the measured flow rates in liters per second. Additionally, the flow rates tended to remain constant whenever reaching this condition, primarily because of the calculation method used, as the flow rates were set as the mean value during the filling cycles. This issue arises from a constant filling cycle and thus the uncertainty increased when high flow rates were encountered. Therefore, this study recognizes the difficulties involved in obtaining precise measurements under such circumstances. Specifically, acknowledging the uncertainty associated with the observed high flow rates, noting that some of the

reported flow rates in reality are expected to be slightly lower with increased variance due to these particular situation.

It is worth noting that the mass balance approach commonly incorporates fluid density in the calculations, accounting for changes in mass within inhomogeneous fluids (Coughanowr & LeBlanc, 2009). However, due to practical constraints, specifically related to measuring the fluid density, it was not included in the calculations. As a result, the calculations were based on volume balance rather than mass balance. This limitation should be considered when interpreting the results, and future studies may explore methods to address this aspect and evaluate its importance. Nevertheless, as presented in Figure 5.3 the initial flow rates, posing a similarity to typical diurnal flow rates of wastewater in residential areas, served as a validation of the flow calculations. They align with the initial expectations regarding BSF rates within the specified sewer shed. Therefore, the flow calculations are still seen as reliable.

It was assumed that the significant variance between daily flow rates can be attributed to multiple factors, as mentioned in subsection 3.1.1 and section 3.2, which present various possibilities for large excess flow rates in the sewer system. Notably, an interesting observation relates to the distinction in sewer pipe placement between Holtandalen and conventional ditches, where storm sewers are typically situated at the bottom. In Holtandalen, the sewer pipes are located at the bottom of the ditch, potentially influencing flow dynamics and increasing infiltration rates. Additionally, the sewer system in Hotandalen has a large portion of combined manholes, potentially resulting in significantly increased flow rates whenever the capacity of the storm sewer is reached. Therefore, it was expected that these factors contributed to the large variations observed in the flow rates.

An intriguing question arises as to whether this methodology could be implemented using existing pump data in the municipal SCADA system. As mentioned in section 4.2, Van Assel et al. (2023) made a similar attempt with promising outcomes, indicating the potential applicability of this approach. Additionally, the results from Davalos et al. (2018) demonstrated the possibilities of utilizing flow data from the SCADA system for a similar tasks. For municipalities, this could involve utilizing the reported power consumption within the SCADA system along with a power-to-flow rate curve typically provided by pump manufacturers. Nevertheless, these findings align with the observations made by Van Assel et al. (2023), who also experienced the limitations of the method when flow rates consistently exceeded the pump capacity over extended periods. Despite this challenge, the method demonstrates its effectiveness on days with heavy rainfall, functioning as intended for the majority of the day until the capacity is continuously exceeded. Van Assel et al. (2023) further emphasized the advantage of gaining valuable insights into pump performance through accurate calculation of individual pump flow rates when integrated into the SCADA system. Investigating this possibility could offer an alternative approach to flow estimation and pose as a potential co-benefit for municipalities with limited opportunities for flow measurements.

5.2.1 Analysis of Wastewater Flow Rates

Upon analyzing the calculated flows, it was possible to manually differentiate between various flow rate conditions, primarily because the flow rates, as displayed in Figure 5.3, exhibited connections with temperature and precipitation. Specifically, during the initial phase, the flows exhibited distinct characteristics typically associated with DWF, such as periods without precipitation and minimal day-to-day variance. However, as the measurement period progressed, differentiating the flows became considerably more challenging due to the significant impact of snow melt on flow rates. The conventional assumption of DWF occurring after three to seven days without precipitation could not be applied since I/I caused by snow melt would still impact the flow rate (Lindholm, 2017). Although these days were not directly influenced by rainfall induced inflow (fast RDII component), the flow rates could still be doubled as a result of increased volumes associated with snow melt and infiltration (slow RDII component).

The calculation of GWI was done based on the WWP and MNF methods described in subsection 4.2.2. Notably, when calculating the GWI during the initial dry weather period, it became evident that substantial infiltration occurs even during this period. Both the WWP-method and MNF-method estimates suggest a significant fraction of the DWF experiences a considerable level of GWI. However, despite the differences between the calculated values from these two methods, they are relatively close to each other, validating the calculation results. Specifically, the WWP-method yielding slightly lower values compared to the MNF-method.

In Figure 5.4, the DWF is further divided into weekends and weekdays, including the mean values for both the WWP and MNF estimates (averaged over all dry weather days). Noticeably, these two curves align with the assumption that early morning wastewater production is delayed during weekends. In Figure 5.5, an overview of wastewater flow outside the dry weather period is provided. This includes days with rainfall and snow melt, highlighting the significant difference in flow rates compared to the DWF. For more detailed information, please refer to Appendix A, where all the measured curves are presented.

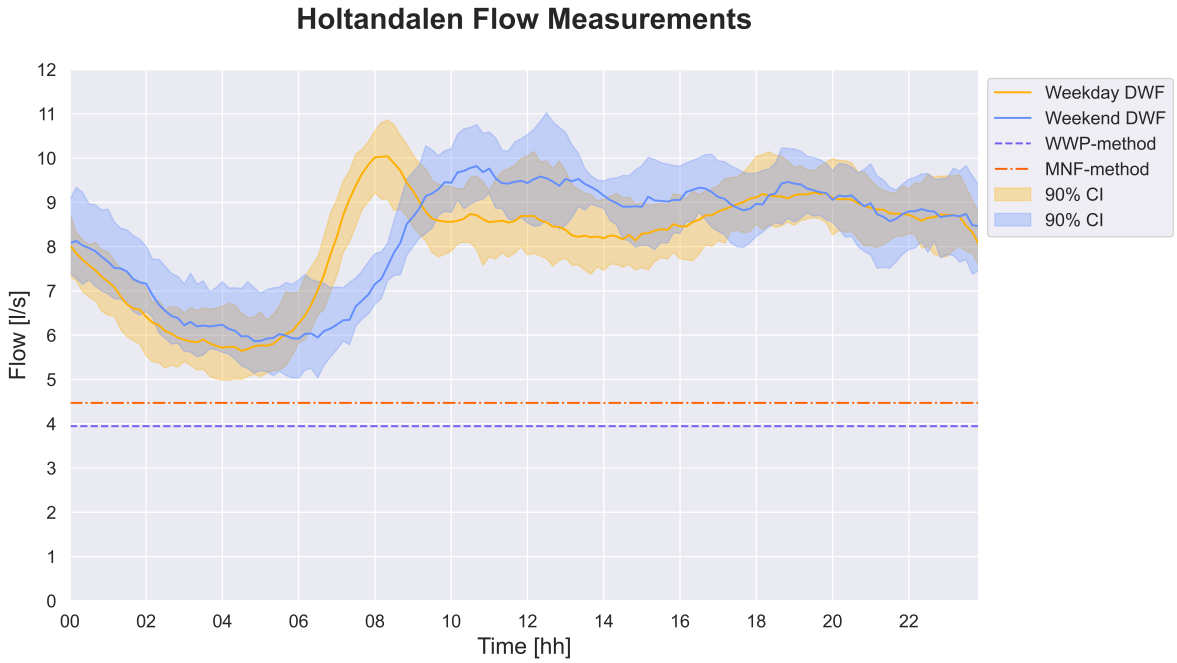


Figure 5.4: Holtdandalen DWF: Weekends and Weekdays with Simplified Groundwater Infiltration Rates (MNF and WWP methods), from the 3rd of March to the 12th of April 2023 (Confidence Interval [CI] = 5th to 95th percentile)

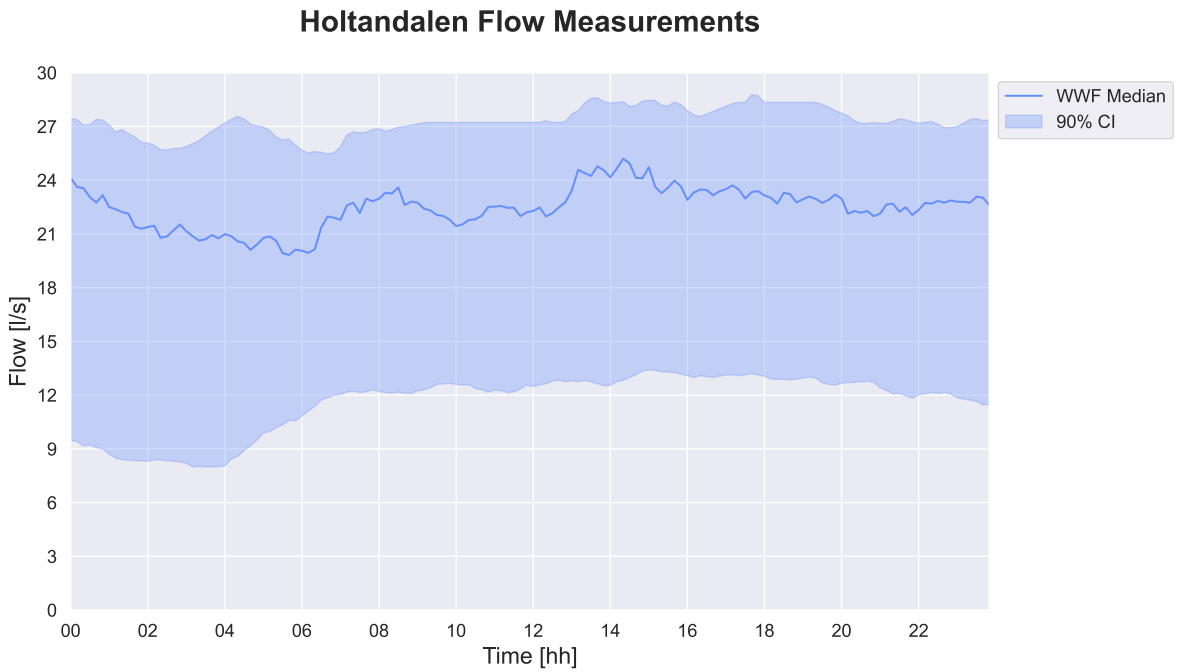


Figure 5.5: Holtdandalen WWF: Flow Rates Including Weekends, Weekdays with High Infiltration and Inflow Rates, from the 3rd of March to the 12th of April 2023 (Confidence Interval [CI] = 5th to 95th percentile)

The manual method employed for distinguishing different flow conditions is inherently associated with a degree of uncertainty as it relies on subjective identification of flows that align with the expectations of DWF and WWF. However, the evident flow correlations with snow melt and rainfall from subsection 5.1.1 support the validity and usefulness of this approach.

Utilization of the WWP and MNF methods in this study provides a straightforward and accessible approach for estimating GWI flow rates for municipalities. However, it is essential to acknowledge that these methods are associated with a level of uncertainty, as they rely on assumptions concerning water consumption and the daily water use. Specifically, the MNF method makes an assumption that the water consumed is equivalent to the wastewater generated. However, this assumption does not consider the presence of leakages in the water distribution pipes, which is a common occurrence in reality. Despite this, it is worth noting that these pipes are often located alongside sewer pipes, and any leaked volumes could potentially infiltrate in to the sewer system.

Additionally, it should be noted that the WWP method introduces additional uncertainty through the selection of an X factor. In this study, the chosen X factor falls within the range described by Mitchell et al. (2007), which was deemed reasonable compared to the other methods investigated in that study. Despite these inherent uncertainties, the results obtained from the WWP method demonstrate reasonable agreement with the water demand curve and therefore the MNF. This alignment indicates that the estimated wastewater flow rates correspond to the expected patterns of water usage.

5.3 Graph Network

In the model preprocessing phase, as visualized in Figure 4.1, the creation of a sewer network graph was a necessary step to address missing information and establish the foundation for a SWMM model. This involved performing computations using a Python script in order to follow the procedure described in section 4.3. This computation used Holtandalen's dataset with 1376 pipes and took approximately 3 minutes to compute with an Intel i5-8250U 3.4GHz processor. The outcome of this graph construction process is illustrated in Figure 5.6, where teal lines represent the connected pipes in the graph, while orange lines indicate the pipes that could not be connected. It is worth noting that all municipal pipes were successfully connected, whereas the missing connections were attributed to incomplete documentation of privately owned pipes. Nonetheless, the sewer network was possible to present as a direction graph and a minimum spanning tree, with its root node as the inlet pipe to the pumping station of Holtandalen. This provided the flow path trough the network, which would simplify the further process of creating a sewer model.

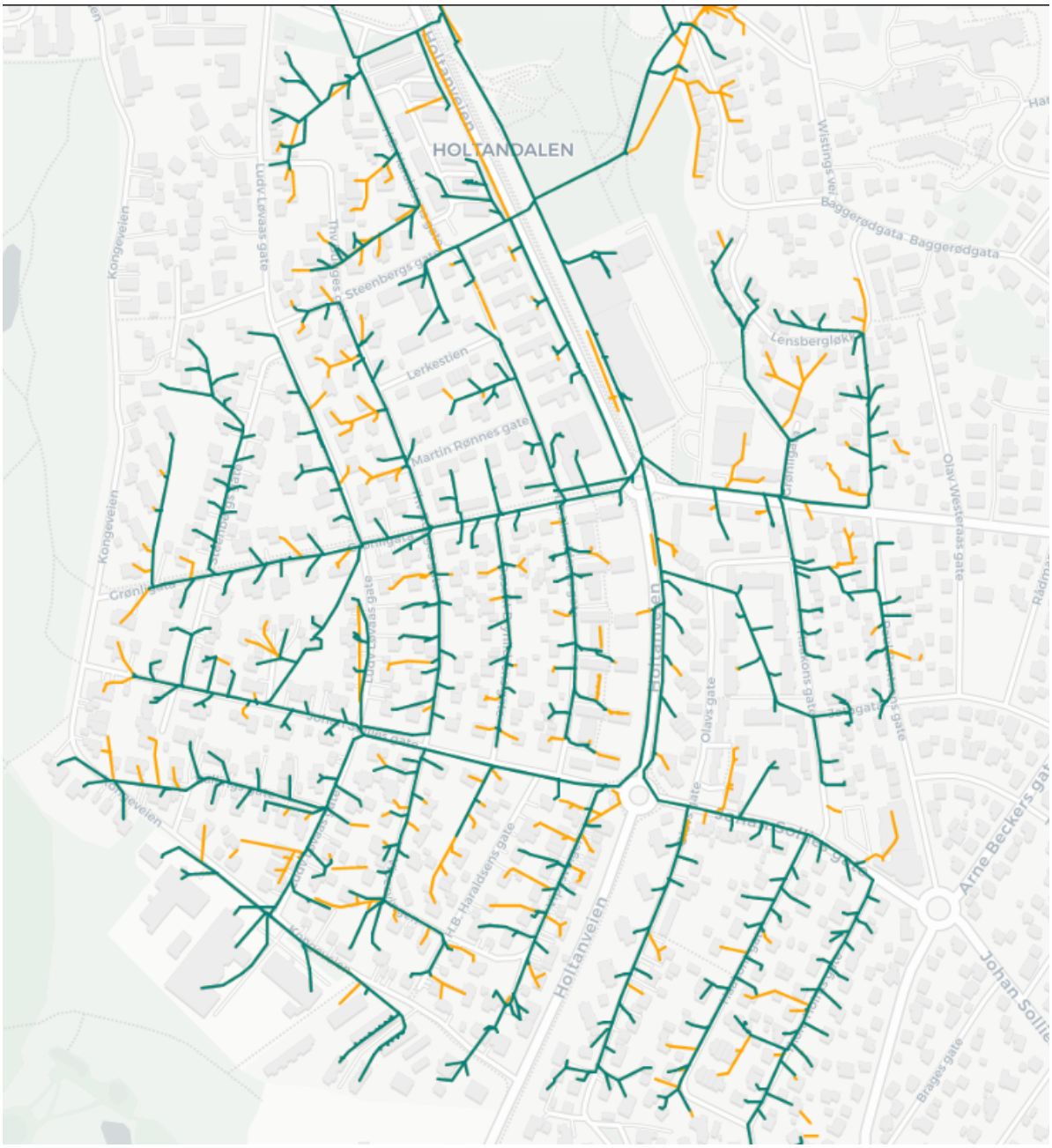


Figure 5.6: Holtandalen Sewer System, teal lines represent pipes connected in the graph, while orange represent unconnected pipes

The graph creation involved utilizing buffer zones to connect nodes. Because of this, it is important to note that there is a level of uncertainty associated with the accuracy of these buffer zones. This uncertainty arises due to the nature of how pipes were recorded in the documentation. Specifically, there was the occurrence of multiple pipes being represented on top of each other whenever they were located in the same ditch. To address this issue, the graph construction considered only pipes with the endpoint falling within the buffer zone, which was defined with a radius of 0.3 meters. The selection of this radius also introduced uncertainties as it was manually chosen. However, after several attempts with different radius values, the value of 0.3 meters emerged as a suitable distance for identifying connections without including false connections.

Furthermore, due to this approach reducing the inclusiveness, a significant fraction of unconnected pipes exists in the graph, as seen in Figure 5.6. This is because these pipes are recorded differently in terms of direction and lack information about the nodes they are connected to. However, it is expected that this issue will become less prevalent in the future as the documentation is expected to be updated and improved. It's still worth noting that this issue primarily affects outer branches and specifically private connections. Its impact on the overall results is not significant, except for ensuring the correct connection of PE to the respective pipes. Therefore, to ensure the inclusion of most PE, the radius of acceptable PE associations was set to 30 meters. This adjustment helped incorporate the majority of the PE into the graph without including unreasonable associations between pipes and PE.

As depicted in Figure 5.7, there are additional sources of errors observed in the unconnected pipes. The main cause of this issue is the unclear placement of pipes and the lack of detailed documentation regarding their connections to the rest of the sewer system. Consequently, it was not feasible to establish connections for these particular pipes with the graph. While a larger buffer radius could have established connections to these pipes, it would also introduce false connections, thereby limiting the reliability of this approach.

Figure 5.7: Infeasible Connections, due to the large distances. Teal lines represent pipes connected in the graph, while orange represent unconnected pipes

This graph creation did not incorporate the stormwater system due to the specific construction design of the manholes. This choice was primarily driven by the difficulty of distinguishing connections within the combined manholes, which exhibited variability based on the flow rate in the system. The design of these manholes are similar to the example presented in Figure 3.3. Consequently, this aspect was intentionally excluded from this study, and the model created by Zaidan (2018) was not incorporated in this study.

The graph creation process considers the limitations and challenges related to buffer zones and addresses specific connectivity within the sewer system. Additionally, as a co-benefit for the municipality, this process enhances data management by organizing and structuring the sewer system representation, including the documentation of pipes, manholes, and connections. This improved data management has the potential to contribute to more efficient operations and maintenance activities, by utilizing advanced graph theory algorithms (e.g. Dijkstra's Algorithm, Ancestors/Descendants, and Prim's Algorithm).

5.3.1 Missing Information in the Dataset

Given the incomplete Holtandalen dataset, additional data management was necessary to address the missing information. A summary of the missing data is presented in Table 5.2 as percentages of the total dataset with 1376 pipes, this also includes the percentage of unconnected pipes resulting from the graph creation process and the number of PE not associated to pipes.

As presented in Table 5.2, filling in the missing information for diameters, lengths, and materials did not require significant effort and could have been manually completed. However, the challenge was the large amount of missing information regarding depths and elevations. The high number of missing elevations and depths was due to the absence of manholes at every connection points, resulting in a lack of information for these intermediate locations.

As discussed earlier in this section regarding the share of unconnected pipes and PE, the presence of unconnected pipes primarily contributes to the share of associated PE. This introduces the issue that there are no pipes in proximity to the house coordinates to which the PE can be associated to.

Table 5.2: Summary of Missing Information in the Graph

Variable	Share of Missing Information
Unconnected Pipes	29.7 %
Person equivalents	3.7 %
Diameter	14.2 %
Length	1.1 %
Material	2.8 %
Elevations	77.4 %
Depths	84.2 %

In general, the methods utilized for missing lengths and elevations are expected to produce reliable data, as these assumptions rely on GIS data. The DTM utilized for the elevations is presented in Figure 5.8, where the overlay of Holtandalen sewer shed is included. The methods utilized for the other variables are further elaborated.

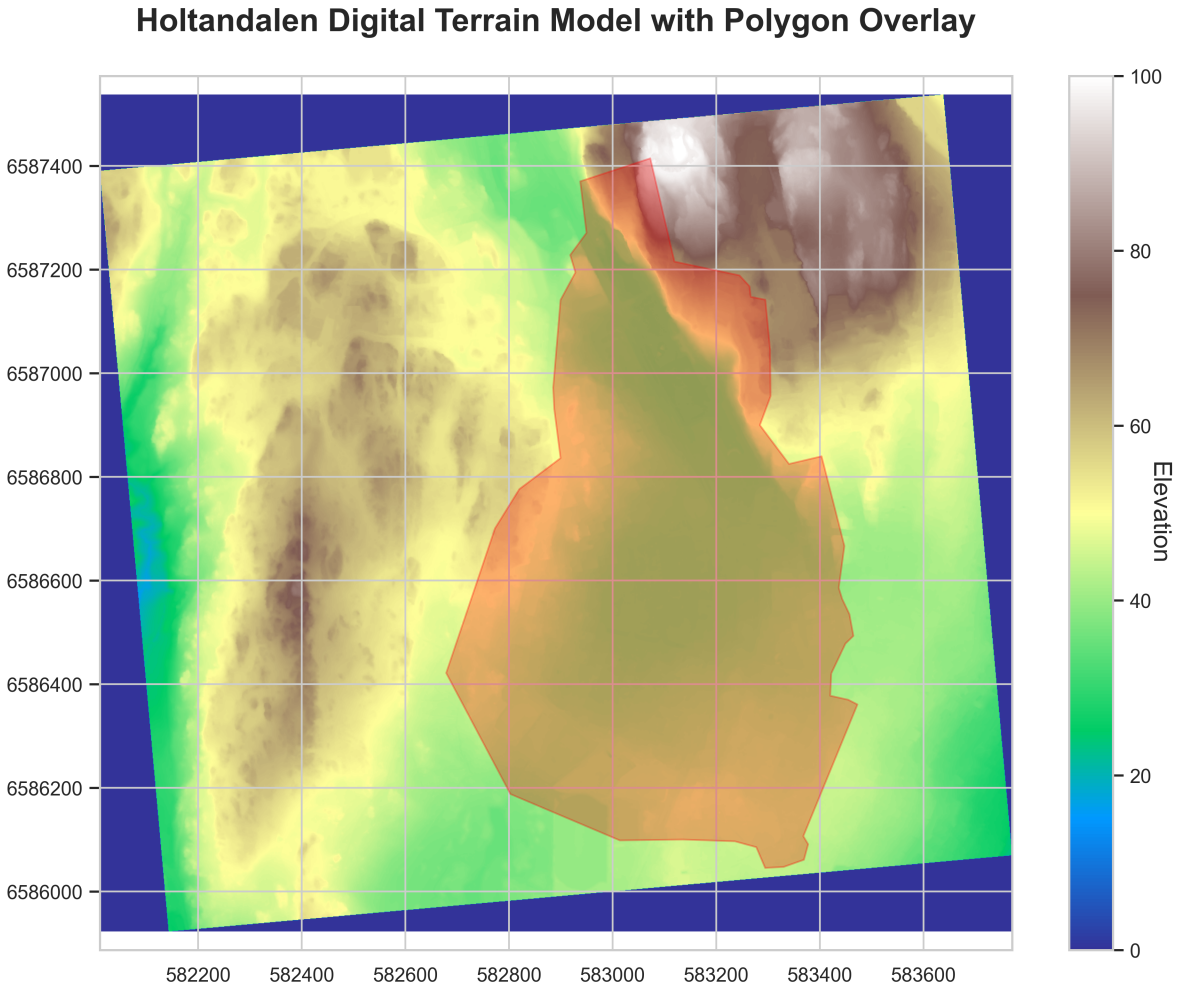


Figure 5.8: Holtdandalen DTM with it's Respective Sewer Zone as a Polygon Overlay

To fill in the diameters, the largest upstream diameter was used as a default value. While this approach may lead to some incorrect assumptions, it is expected to minimize the incorrect assumptions. Additionally, by considering the surrounding diameters it is expected to result in assumptions that closely approximate the actual diameter, even if not entirely accurate for all cases.

In instances where materials were missing, the "most common variable" by dimension was employed. However, it should be noted that this approach may yield strange effects and be somewhat unreliable when dealing with a significant number of missing diameters or materials. Nevertheless, due to the relatively small proportion of missing data, it is anticipated to produce reliable results. Alternative approaches such as incorporating street sections or conducting proximity analysis could be considered to determine the material. Additionally, other statistical techniques could be utilized for this purpose. However, since this parameter is not considered crucial for this particular study and is

solely used to establish pipe roughness, which is further calibrated within the model, the focus on precision is not as critical.

The depth in the sewer network can be subject to inaccuracies and uncertainties, especially when there is limited available information. In this study, as described in Table 5.2, there is a significant amount of missing information, making it challenging to predict accurate depth values. Therefore, it is not expected that the exact depth for each pipe will be correct. However, the method employed in this study ensures that the overall slope throughout the system remains reasonable by utilizing an interpolation method between known depths. By including minimum and maximum limitations, the occurrence of unusual effects such as negative slopes and excessively shallow depths when determining the depths is also reduced. Nonetheless, these limitations can become problematic in cases where the system is actually poorly designed with shallow depths or slopes less than 10‰.

Generally for all adjustments, in cases where there is a significant amount of missing information in a municipal database, challenges may arise during the creation of the sewer network graph. The completeness and quality of the data directly affect the reliability and usefulness of the resulting graph. Furthermore, pipes in the dataset that lack flow direction or contain false information can have a negative impact on the accuracy and reliability of the sewer network graph. Therefore, careful consideration and quality control were necessary to minimize the influence of such issues on the results. In this case, this was done by visually searching and validating the outliers by plotting the sewer system attributes similarly to Figure 5.6.

In contrast to the GIS approach employed by Schilling and Tränckner (2022), who developed an open-source plugin to utilize geospatial data, the utilization of a graph to generate a SWMM-model has presented a superior alternative for addressing the data gaps. This approach allows for customization to meet the required assumptions regarding information completion, while the connectivity facilitated by the graph greatly enhances the available options and possibilities for data filling. Providing the municipality with the flexibility to fill in data gaps according to their specific assumptions or requirements through minor adjustments to the script.

5.4 SWMM-Model

To achieve the objective of the model preprocessing phase, which was to create a calibrated SWMM model representing Holtandalen's sewer system, the method described in section 4.4 was utilized. This consisted of generating a SWMM model that would estimate the sewage flow rates, based on the known water demand for the given area. Providing a more accurate measure of the BSF, which would facilitate the quantification of the I/I (section 5.5).

The effectiveness of the approach was demonstrated by successfully incorporating the entire dataset from the graph network described in section 5.3 into the model, as depicted in Figure 5.9. This was possible to achieve within seconds of computation time and underscores the applicability and robustness of method in handling and utilizing complex datasets. This was facilitated by the SWMM-API Python module (Pichler,

2022), which incorporated adaptations from both GIS and sewer network graphs. The module not only enabled the creation of the sewer model, but also offered functionalities for simulations, adjustments and network visualization. Additionally, this approach simplified the setup of initial conditions, as customizable Python scripts were employed to automate the process. Notably, the inclusion of time series data was seamlessly integrated into the model creation process, allowing for the direct export of flow data from the municipal database to be added to each node associated with specific PE. This provided the municipality with a simple approach to create a sewer model in EPA-SWMM, using only their information about pipes, manholes and peoples residency, with the continuously measured water demand superpositioned throughout the model.

Figure 5.9: Holtandalen EPA-SWMM model: Black Dots Represent Network Nodes and Lines Represent Pipes

The implementation of the sewer network within an EPA-SWMM input file enabled the simulation of the sewer flow through the model. This simulated flow could be visualized in 2D plots, as depicted in Figure 5.10, allowing for a visual representation of the simulated flow patterns throughout the simulated period.

By examining Figure 5.10 and similar plots, it became evident that the model exhibited backwater effects due to negative slopes in certain sections. Whereby, these occurrences were mainly attributed to the documented elevations and depths. Although the accuracy of this representation in reflecting real conditions remains uncertain, it offers a possibility for the municipality to detect and resolve localized occurrences if they indeed exist. Additionally, it is important to mention that the limited number of instances was unlikely to have a significant impact on the model's simulations. In fact, if these instances correspond to real-world conditions, they would benefit the accuracy of the simulations.

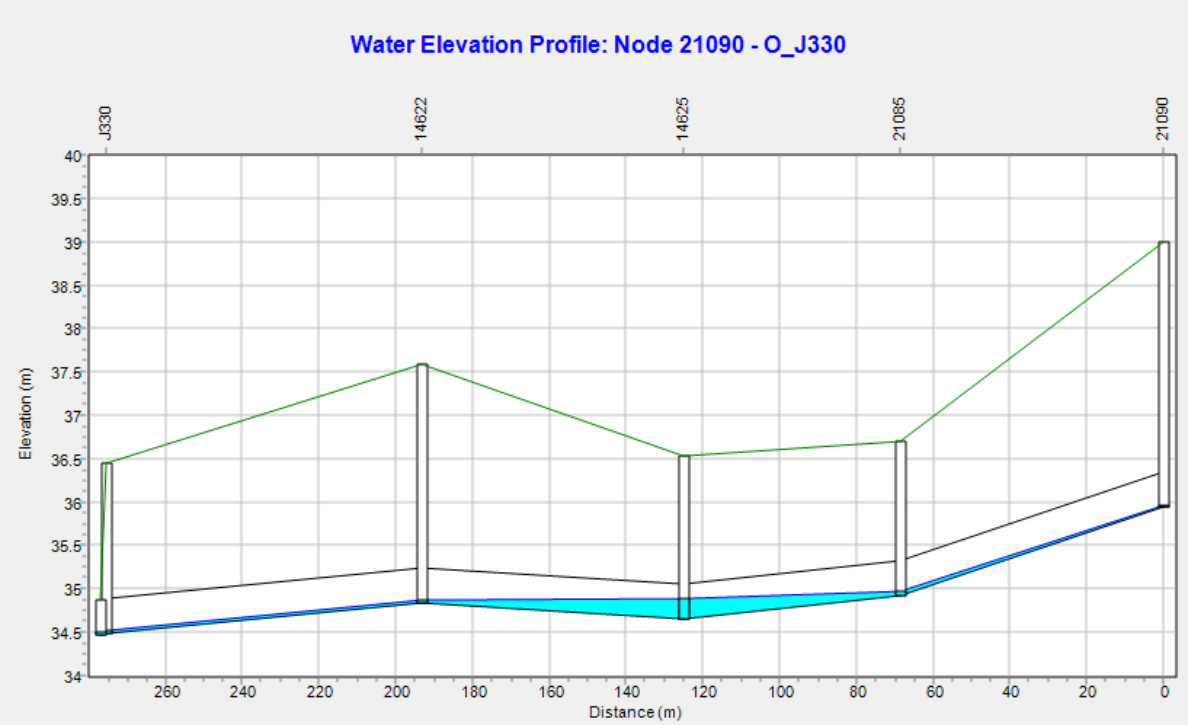


Figure 5.10: Sewage Flow Trough the EPA-SWMM Model

Moreover, the dataset used to generate the graph in section 5.3 lacked details regarding the elevation offsets of pipes within the manholes. Consequently, if a pipe was connected at a higher level than the manhole floor, it would not be properly integrated into the SWMM-model. This issue could potentially affect the accuracy of the slope calculations. Hence, influence the backwater effect seen in Figure 5.10, but without access to the pipe offset information, it was not possible to verify this aspect. Furthermore, it is worth noting that the common practice in documentation aligns the bottom of the manhole with the bottom of the sewer pipe, suggesting that this is not a significant concern.

Another notable simplification in the model refers to the procedure employed for connecting nodes to each other, particularly with regards to lateral connections. As described in section 4.4, instead of creating new connections at the designated locations,

an alternative solution was implemented by diverting the connection through the upstream node. Consequently, this approach resulted in a longer flow path for wastewater in certain locations, potentially leading to minor effects on the slope between these nodes. However, it was anticipated that this would only introduce some additional delay in the flow from the specific nodes.

Regarding issues related to the slope in these cases, it was not expected to pose a significant problem since the slope at these locations was estimated by interpolating known elevations further upstream, as these were not expected to produce significant flow rates. Specifically, if the connection represented a large municipally owned sewer pipe, it would most often have been connected in the manhole with reported elevations.

5.4.1 Model Calibration

To fulfill the final criterion of the model preprocessing phase, a calibration of the SWMM model was performed. As shown in Figure 5.11, the calibration can be considered successful as it was possible to reconstructed the flow rate patterns in the SWMM model. Specifically, by visual inspection, the simulated pattern closely resembles the measured pattern during the specified time period. Additionally, although the MSE and NSE values may not be an perfect fit (NSE = 1, MSE = 0), they indicate a good fit of the model. Scatter plots showing the relationship between the measurements and the simulation results are provided in Appendix A.

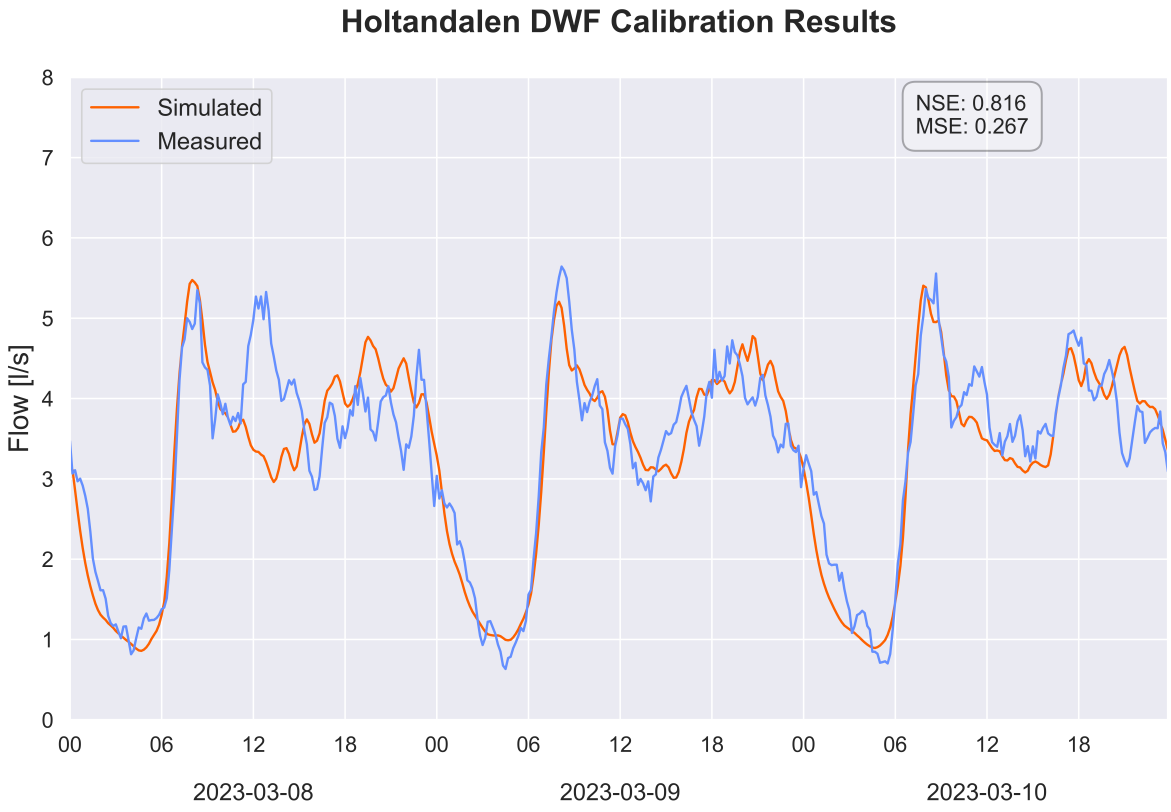


Figure 5.11: Holtandalen Calibration Results

In order to validate the goodness-of-fit achieved during the calibration process, the model was tested on an additional time series. As depicted in Figure 5.12, the simulated curve displayed a remarkable alignment, indicating a strong agreement between the measurement and simulation values. This agreement is further supported by the goodness-of-fit measures, confirming the high level of correspondence between these patterns.

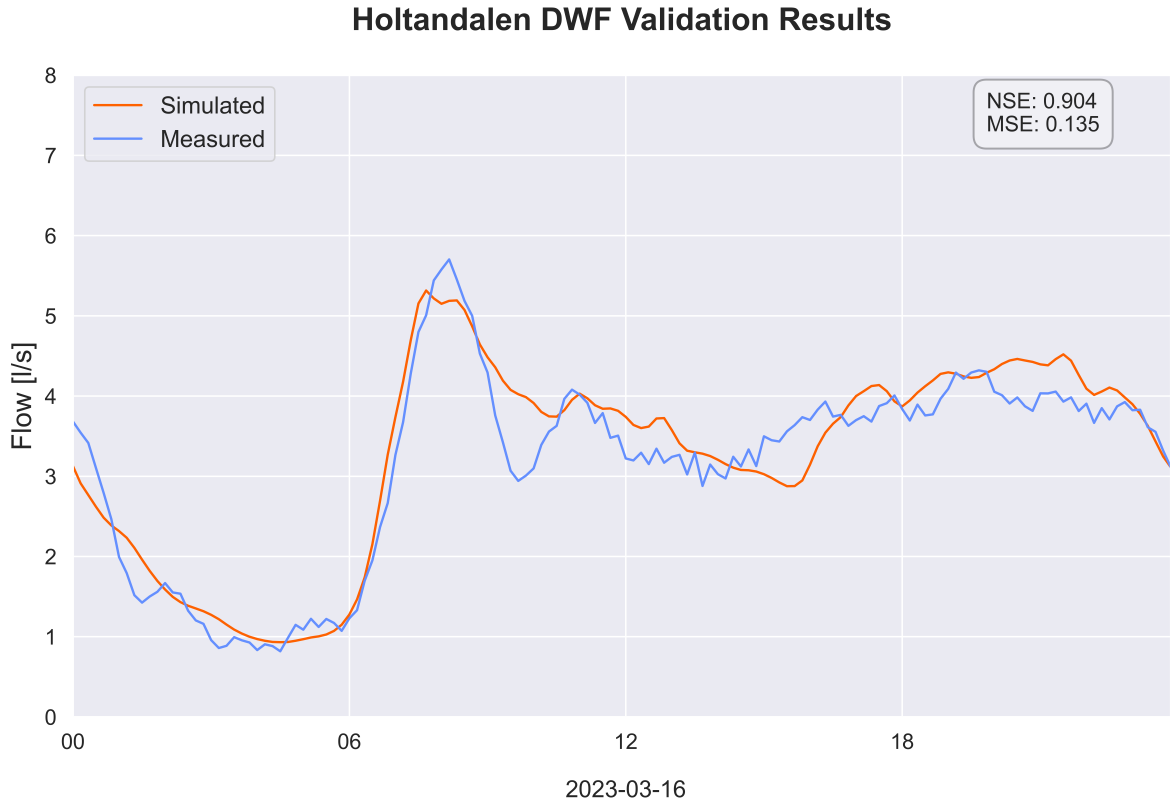


Figure 5.12: Holtandalen Validation Results

As mentioned in subsection 4.4.2, the time periods utilized for the calibration process were chosen due to their flow rates occurring in periods with minimal influence from rainfall and snowmelt, as depicted in Figure 5.3. Nevertheless, an interesting observation in Figure 5.11 is the absence of a large flow peak recorded approximately 12 o'clock on the 8th of March. Upon reviewing the climate data depicted in Figure 5.3, a slight positive temperature of 2.7°C as a results of a significant temperature spike was evident during this period, suggesting the possibility of RDII resulting from snowmelt.

The roughness parameter was calibrated by applying a multiplicative factor, effectively adjusting the initial values. The calibration process produced satisfactory results, with a minor adjustment of 0.9 implemented as the multiplicative value. While the exact cause of the reduction is uncertain, one possible explanation for this decrease is the implementation of the alternative solution for lateral connections through the upstream nodes, which may require a slight increase in velocity to compensate for the extended length. Hence, a reduction in the roughness.

Regarding the daily production rate, it was calibrated to a slightly lower value than the

initially set value. In addition, the measured wastewater flow rates were adjusted by subtracting the estimated daily values calculated using the MNF-method, as this was necessary to establish comparable flow domains for analysis. As a result, the uncertainty associated with the MNF-method is expected to be compensated for by a reduction or increase in the estimated BSF. Consequently, any overestimation of the GWI would induce a reduction to the daily production rate. The observed slightly lower daily production rate may be attributed to a minor overestimation of the GWI. However, it still falls within the acceptable range of 100 to 200 liters per day, which aligns with typical expectations (The exact value is subjected to confidentiality). Notably, this lower flow rate is also consistent with the winter and spring seasons when water consumption tends to be lower compared to the summer.

Overall, the calibration of the SWMM model was a crucial step in achieving accurate flow rate predictions for the specified time period, enabling for the subsequent utilization of the model in quantifying I/I. The application of only two calibration variables proved to be effective, resulting in promising outcomes for the model. This streamlined approach could provide the municipality with an easy access to the calibration process, as it only required manual adjustments on two parameters, both leading to distinct changes in the resulting flow rate pattern.

5.5 Quantifying I/I in Holtandalen Sewer Zone

To accurately address the main research objective of this study, regarding the quantification of I/I in Holtandalen sewer zone, the calibrated model was executed throughout the measurement period, providing estimates of I/I levels for each day. The separation and estimation of BSF, GWI and RDII quantities were performed using the method described in section 4.5. The results of this quantification are presented in Table 5.3, where days with faulty measurements are marked with (-).

Table 5.3: I/I Quantification Results, Including Total Measured Volume and Respective Shares of BSF, GWI, RDII (from 03.03.2023 to 12.04.2024)

Date	Total Volume [m^3]	BSF [%]	GWI [%]	RDII [%]
03.03.2023	-	-	-	-
04.03.2023	-	-	-	-
05.03.2023	797	35	65	0
06.03.2023	763	36	64	0
07.03.2023	743	41	59	0
08.03.2023	728	40	60	0
09.03.2023	695	41	59	0
10.03.2023	672	42	58	0
11.03.2023	651	43	57	0
12.03.2023	672	42	58	0
13.03.2023	645	42	58	0
14.03.2023	677	53	47	0
15.03.2023	715	50	50	0
16.03.2023	675	41	59	0
17.03.2023	1848	15	22	63
18.03.2023	2386	12	17	71
19.03.2023	1908	14	21	64
20.03.2023	1791	16	22	62
21.03.2023	1613	18	25	57
22.03.2023	-	-	-	-
23.03.2023	2024	14	20	66
24.03.2023	2088	13	19	68
25.03.2023	2253	13	18	69
26.03.2023	2081	13	19	68
27.03.2023	-	-	-	-
28.03.2023	1420	21	28	51
29.03.2023	1323	21	30	49
30.03.2023	1215	23	33	44
31.03.2023	1153	25	35	40
01.04.2023	1087	26	37	37
02.04.2023	1037	27	39	34
03.04.2023	969	28	41	31
04.04.2023	932	30	43	27
05.04.2023	916	28	44	28
06.04.2023	929	28	43	29
07.04.2023	924	28	44	28
08.04.2023	942	26	43	31
09.04.2023	921	27	44	29
10.04.2023	930	27	43	30
11.04.2023	-	-	-	-
12.04.2023	-	-	-	-

The results demonstrate a significantly high level of GWI (approximately 60%) during the initial period with DWF, indicating that the zone is generating a considerably greater amount of flow than what is produced by households alone. Furthermore, in days described with WWF, there are drastic changes in flow patterns, with the BSF contributing to a minimum of 12% of the total daily volume. Hence, suggesting that the zone is highly influenced by I/I in this period.

In general, as the total share of I/I over this evaluation period was 76 % (RDII and GWI), it indicates that this zone can be considered worse than the national average for I/I, as described by Jørgensen and Rostad (2021) and Scherling et al. (2020).

Upon visual examination of the results from the 21st of March 2023, as shown in Figure 5.13, the analysis further reveals substantial levels of I/I resulting from the combined effects of GWI and RDII. The initial levels of RDII can be attributed to the snowmelt and rainfall that occurred the previous day. Additionally, the flow rates demonstrate a gradual decay, followed by significant increases shortly after precipitation events. These observations provide verification of the anticipated impact of precipitation on the sewer system, as presented by the flow-based method in subsection 5.1.1.

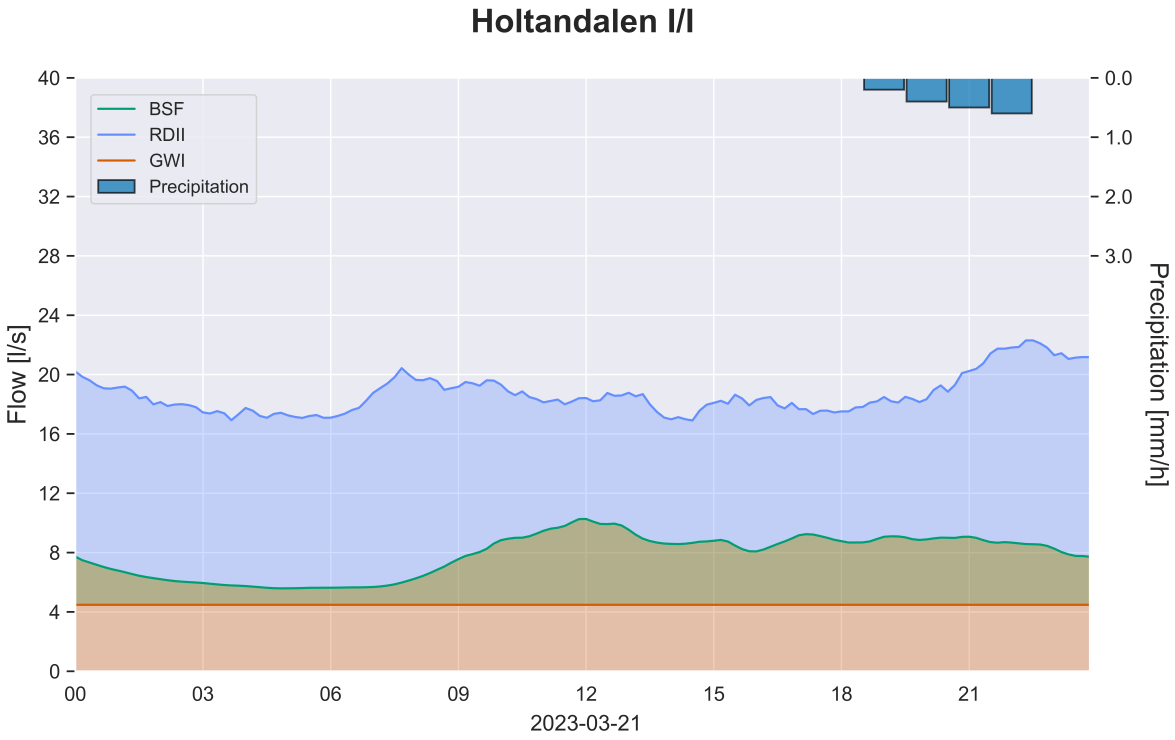


Figure 5.13: Comparison of Simulated DWF and Measured WWF: Separated Into RDII, BSF and GWI Components

Moreover, the model was executed on one of the days characterized by the highest flow rates, as depicted in Figure 5.14. It was evident that the flow rate reached a plateau, as indicated in the figure. As mentioned in section 5.2, this phenomenon occurs because the flow rate reaches the capacity of the pumping station, particularly considering the continuous operation of the pumps. Consequently, the method employed for flow measurements demonstrates inefficiency whenever the flow rate consistently exceeds the pumping station's capacity. Nevertheless, it should be noted that the procedure manages to handle higher flow rates, but for prolonged periods the flow rate flattens out as a mean over the emptying cycle. Hence, the daily total value is still expected to provide reasonable results.

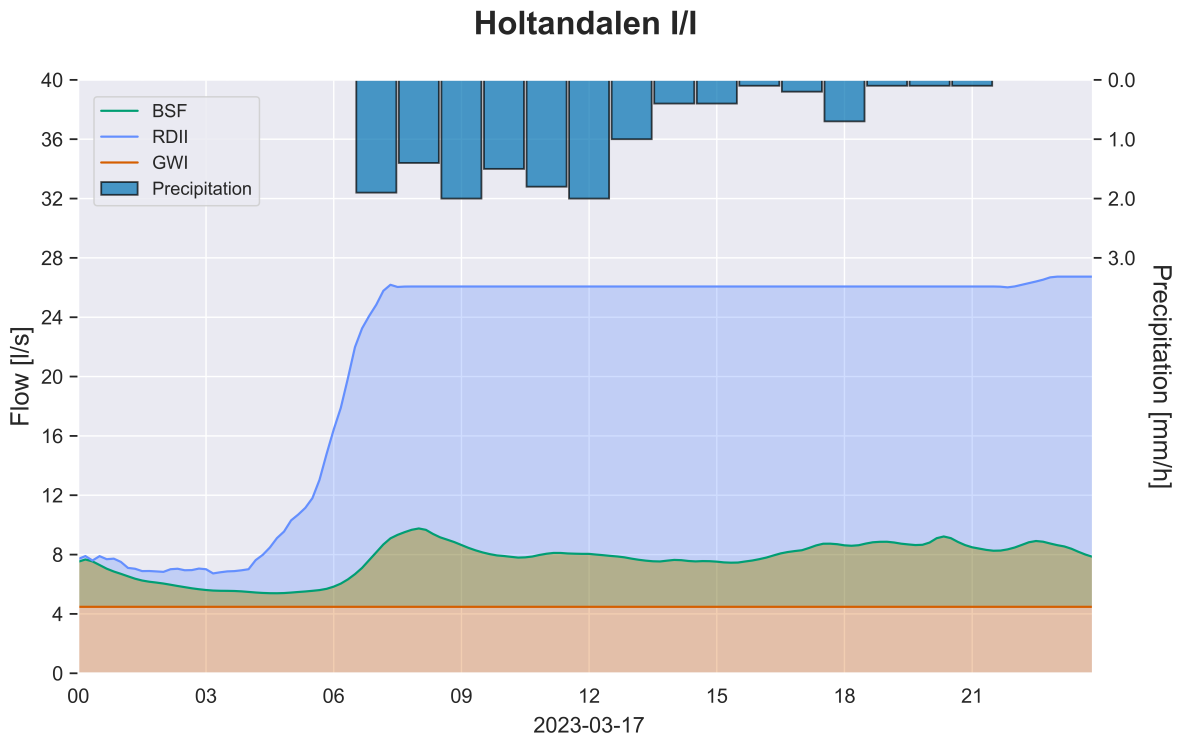


Figure 5.14: Flow Rates Reaching a Plateau due to Pumping Station Capacity

In connection with Figure 5.14, a closer examination of the climate data presented in Figure 5.3 reveals that the observed slight increase in flow rate preceding the precipitation event can be attributed to an anticipated volume resulting from snowmelt. This occurrence is anticipated as the reduction of the snowpack contributes to increased runoff and subsequent flow rates.

The GWI component was calculated based on the preceding DWF, representing the constant component and seasonal variation of the infiltration rate. In this approach, the infiltration resulting from snowmelt and rainfall is included in the RDII component. By differentiating between GWI and RDII, the model considers the distinct mechanisms contributing to the overall I/I, providing a more comprehensive understanding of the system dynamics. However, the accuracy of the GWI and RDII quantities is influenced by the extrapolation of GWI values, which can lead to uncertainty regarding the differentiation between GWI and RDII. Nevertheless, considering that GWI typically represents a consistent flow rate, its effect on the results is therefore expected to be minimal.

When simulating the BSF while considering the WWF, it is important to acknowledge that solely simulating the BSF component might have overlooked certain hydraulic aspects of the overall actual flow rate. While the simulated BSF takes into account the inflowing sewage, there is a possibility that the inflow rates of sewage may not precisely correspond to the conditions, especially considering the actually increased volumes resulting from RDII, which can significantly impact the hydraulic conditions of the system. However, despite this potential discrepancy, the overall volumes estimated through the daily simulation are still expected to be accurate. While including the WWF in the sim-

ulation, this could potentially provide a more comprehensive representation of the flow dynamics and enhance the reliability of the results, in this study, due to the reliance on estimating the discrepancy between measurements and simulations, it was not feasible to incorporate the WWF component.

Overall, the obtained results align with the anticipated I/I identified in the correlation investigation (subsection 5.1.1) and the potential inflow effects associated with the combined manholes (subsection 3.1.1). Moreover, the notably high infiltration rates observed during this period may be attributed to the pipe placement in the ditches, as discussed in section 3.2. Additionally, as a benefit of this study, these accurate estimations of the extent of I/I in Holtandalen provide valuable insights for the municipality and serve as a decision support tool for their upcoming system renovation, as mentioned in subsection 3.1.2.

6 Conclusion

The primary objective of this study was to quantify the extent of I/I in a municipal sewer zone using simple approaches applicable to small municipalities. The findings of this research successfully demonstrate the potential for this in a municipal sewer system, even with limited data availability. By leveraging the existing information stored in the municipal database and utilizing graph theory and other techniques, this study estimated expected BSF and compared it to measured wastewater flow, providing valuable insights into the response to precipitation events, and estimates of GWI and RDII.

The analysis of the existing information revealed challenges associated with data reliability and accuracy, but through reverse engineering the available data generated by a pumping station, the accuracy of information was significantly improved. This, in turn, enabled a more thorough analysis of the effects caused by various weather events, and presenting an accessible approach for small municipalities to acquire flow measurements. However, a challenge in this approach emerged when the pumping station operated at maximum capacity for extended periods, due to the constant pumping of the pumps, rendering the calculation method ineffective and unable to provide detailed information. Nevertheless, these findings highlighted the importance of considering data quality and the influence of different weather events when assessing I/I.

Moreover, the process of creating the sewer network graph from the municipal database proved to be efficient and straightforward, enabling quick analysis and data filling based on the graph representation of the sewer network. This approach, with minor adjustments, can be easily replicated and implemented in other municipalities, making it particularly advantageous for small municipalities with limited resources and expertise. The effectiveness, resource optimization, and improved system evaluations associated with this approach provide significant co-benefits for small municipalities.

While the inclusion of hydraulic models may pose challenges and require some level of modeling knowledge, the benefits of incorporating flow-based approaches with hydraulic models to validate and refine the results cannot be overlooked. A balance between practicality and accuracy is maintained when utilizing this approach, contributing to improved I/I management and maintenance decisions for small municipalities.

Overall, the results obtained from this study hold significant value for the municipality in this case, serving as decision support for their sewer network. By identifying areas of concern and quantifying the extent of I/I, the municipality can can make informed decisions to enhance the performance of their sewer systems, specifically by prioritizing necessary interventions and optimizing their resources.

In terms of future work, a valuable direction to explore would be the verification of the estimated flow rate composition by comparing it with other methods. This validation

process would further enhance the accuracy and reliability of the proposed model, increasing confidence in its applicability for similar studies or other real-world scenarios. Such validation steps would strengthen the robustness of the methodology and contribute to a broader understanding of I/I dynamics in sewer systems, particularly for municipalities with limited data availability.

Bibliography

Bäckman, H., Peterson, B.-L., Marklund, B., Olsson, R., & Wästlin, T. (1993). *Indirekt nederbördspåverkan i spillvattensystem* (tech. rep.). Retrieved March 6, 2023, from https://vattenbokhandeln.svensktvatten.se/produkt/indirekt-nederbordspaverkan-i-spillvattensystem/

Beheshti, M., & Sægrov, S. (2018). Quantification Assessment of Extraneous Water Infiltration and Inflow by Analysis of the Thermal Behavior of the Sewer Network. *Water*, *10*(8), 1070. https://doi.org/10.3390/w10081070

Beheshti, M., Sægrov, S., & Ugarelli, R. (2015). Infiltration / Inflow Assessment and Detection in Urban Sewer System. *Vannforeningen*. https://vannforeningen.no/wp-content/uploads/2015/01/Beheshti.pdf

Bennett, N. D., Croke, B. F. W., Guariso, G., Guillaume, J. H. A., Hamilton, S. H., Jakeman, A. J., Marsili-Libelli, S., Newham, L. T. H., Norton, J. P., Perrin, C., Pierce, S. A., Robson, B., Seppelt, R., Voinov, A. A., Fath, B. D., & Andreassian, V. (2013). Characterising performance of environmental models. *Environmental Modelling & Software*, *40*, 1–20. https://doi.org/10.1016/j.envsoft.2012.09.011

Bentes, I., Silva, D., Vieira, C., & Matos, C. (2022). Inflow Quantification in Urban Sewer Networks [Number: 4 Publisher: Multidisciplinary Digital Publishing Institute]. *Hydrology*, *9*(4), 52. https://doi.org/10.3390/hydrology9040052

Bogusławski, B., Sobczak, P., & Głowacka, A. (2022). Assessment of extraneous water inflow in separate sewerage system by different quantitative methods. *Applied Water Science*, *12*(12), 278. https://doi.org/10.1007/s13201-022-01793-2

Brin, S., & Page, L. (1998). The anatomy of a large-scale hypertextual Web search engine. *Computer Networks and ISDN Systems*, *30*(1), 107–117. https://doi.org/10.1016/S0169-7552(98)00110-X

Bruaset, S. (2019). *Long-term sustainable management of the urban water and wastewater pipe networks* (Doctoral thesis) [Accepted: 2019-12-03T14:00:52Z ISBN: 9788232638017 ISSN: 1503-8181]. NTNU. Retrieved May 30, 2023, from https://ntnuopen.ntnu.no/ntnu-xmlui/handle/11250/2631569

Bruaset, S., Becker, M. A., Reksten, H., & Baade-Mathiese, T. (2021). *A 259 Kommunalt investeringsbehov for vann og avløp 2021-2040* (tech. rep.). Norsk Vann. Retrieved November 23, 2022, from https://va-kompetanse.no/butikk/a-259-kommunalt-investeringsbehov-for-vann-og-avlop-2021-2040/

Butler, D., Digman, C. J., Makropoulos, C., & Davies, J. W. (2018). *Urban Drainage* (4th ed.). CRC Press. https://doi.org/10.1201/9781351174305

Catalfio, C., Hughes, C., Zorza, G., Mullett, N., & Toumari, D. (2007). Alternatives for Finding Illicit Discharges to Stormwater Systems [ISSN: 1938-6478], 2079–2104. Retrieved May 30, 2023, from https://www.accesswater.org/publications/

proceedings/-294072/alternatives-for-finding-illicit-discharges-to-stormwater-systems

Choi, N., & Schmidt, A. R. (2023). Rainfall-Derived Infiltration and Inflow Estimate in a Sanitary Sewer System Using Three Impulse Response Functions Derived from Physics-Based Models. *Water Resources Management*, *37*(1), 305–319. https://doi.org/10.1007/s11269-022-03370-3

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2002). *Applied Multiple Regression/-Correlation Analysis for the Behavioral Sciences* (3rd ed.). https://doi.org/10.4324/9780203774441

Comeau, A., El-Sayegh, H., Kauffman, J., Swarner, B., & Vallabhaneni, S. (2019). *Sanitary Sewer Flow Monitoring and Data Analytics* (Fact sheet). Retrieved December 4, 2023, from https://www.wef.org/globalassets/assets-wef/direct-download-library/public/03---resources/wsec-2019-fs-011---csc---flow-monitoring-and-data-analytics---final.pdf

Coughanowr, D. R., & LeBlanc, S. E. (2009). *Process Systems Analysis and Control, 3rd Edition* (3rd ed., Vol. 2). McGraw-Hill New York. https://udghoshna.files.wordpress.com/2013/06/136649035-process-systems-analysis-and-control-d-coughanowr-3rd-ed.pdf

Council Directive 91/271/EEC. (1991). Retrieved May 28, 2023, from http://data.europa.eu/eli/dir/1991/271/oj/eng

Davalos, P., Humphrey, G., Eagle, S., Roque, R., Bedoya, J., Edwards, D., Torrealba, F., Perez, J., & Dvorak, A. (2018). Cost Effective Infiltration and Inflow Analysis and Remediation Efforts in Miami-Dade County [ISSN: 1938-6478], 2994–3023. Retrieved May 30, 2023, from https://www.accesswater.org/publications/proceedings/-299308/cost-effective-infiltration-and-inflow-analysis-and-remediation-efforts-in-miami-dade-county

Ellis, B., & Bertrand-Krajewski, J.-L. (2010). *Assessing Infiltration and Exfiltration on the Performance of Urban Sewer Systems*. https://www.iwapublishing.com/books/9781843391494/assessing-infiltration-and-exfiltration-performance-urban-sewer-systems

Engineering ToolBox. (2004). Manning's Roughness Coefficients. Retrieved April 28, 2023, from https://www.engineeringtoolbox.com/mannings-roughness-d_799.html

Esri. (2022). Intersect (Analysis)—ArcGIS Pro. Retrieved November 26, 2022, from https://pro.arcgis.com/en/pro-app/latest/tool-reference/analysis/intersect.htm

European Commission. (2022). The Urban Waste Water Treatment Directive (proposed revision). Retrieved February 27, 2023, from https://environment.ec.europa.eu/topics/water/urban-wastewater_en

Fugledalen, T., Rokstad, M. M., & Tscheikner-Gratl, F. (2021). On the influence of input data uncertainty on sewer deterioration models – a case study in Norway [Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/15732479.2021.1998142]. *Structure and Infrastructure Engineering*, *19*(8), 1064–1075. https://doi.org/10.1080/15732479.2021.1998142

Furqan, M., Sitompul, O., Mawengkang, H., Siahaan, A. P. U., Donni Lesmana Siahaan, M., Wanayumini, & Nasution, N. (2018). A Review of Prim and Genetic Algorithms in Finding and Determining Routes on Connected Weighted Graphs. https://www.researchgate.net/publication/328146864_A_Review_of_Prim_

and _ Genetic _ Algorithms _ in _ Finding _ and _ Determining _ Routes _ on _ Connected_Weighted_Graphs

Gillies, S. (2023). Rasterio: Fast and direct raster I/O for use with Numpy and SciPy. Retrieved May 21, 2023, from https://github.com/rasterio/rasterio

Gillies, S., van der Wel, C., Van den Bossche, J., Taves, M. W., Arnott, J., Ward, B. C., et al. (2023). Shapely. https://doi.org/10.5281/zenodo.5597138

Google. (2023). Google Maps. Retrieved March 7, 2023, from https://www.google.no/maps/@59.4002982,10.4199843,12z?hl=no

Goulding, R., Jayasuriya, N., & Horan, E. (2012). A Bayesian network model to assess the public health risk associated with wet weather sewer overflows discharging into waterways. *Water Research*, *46*(16), 4933–4940. https://doi.org/10.1016/j.watres.2012.03.044

Haas, D., Ng, S., Dahl, N., & Baulch, D. (2021). Estimating ADWF at Sewage Treatment Plants. *Water e-Journal*, *6*(1). https://doi.org/10.21139/wej.2021.005

Hagberg, A. A., Schult, D. A., & Swart, P. J. (2008). Exploring Network Structure, Dynamics, and Function using NetworkX. In G. Varoquaux, T. Vaught, & J. Millman (Eds.), *Proceedings of the 7th Python in Science Conference*. https://conference.scipy.org/proceedings/SciPy2008/paper_2/full_text.pdf

Hansen, I. L. (2022). Byliv, idyll og historie - Horten kommune. Retrieved October 16, 2022, from https://www.horten.kommune.no/om-horten/

Harju, T. (2014). *Lecture notes on Graph theory*. Retrieved March 28, 2023, from https://docplayer.net/20751931-Lecture-notes-on-graph-theory-tero-harju.html

Heiderscheidt, E., Tesfamariam, A., Marttila, H., Postila, H., Zilio, S., & Rossi, P. M. (2022). Stable water isotopes as a tool for assessing groundwater infiltration in sewage networks in cold climate conditions. *Journal of Environmental Management*, *302*, 114107. https://doi.org/10.1016/j.jenvman.2021.114107

Hey, G., Jönsson, K., & Mattsson, A. (2016). The impact of infiltration and inflow on wastewater treatment plants: A case study in Sweden. *VA-Teknik Södra*. https://va-tekniksodra.se/wp-content/uploads/2016/12/06-2016-Report-on-Infiltration-Inflow_Hey-et-al-2016.pdf

Hoes, O. A. C., Schilperoort, R. P. S., Luxemburg, W. M. J., Clemens, F. H. L. R., & van de Giesen, N. C. (2009). Locating illicit connections in storm water sewers using fiber-optic distributed temperature sensing. *Water Research*, *43*(20), 5187–5197. https://doi.org/10.1016/j.watres.2009.08.020

Jordahl, K., Bossche, J. V. d., Fleischmann, M., Wasserman, J., McBride, J., Gerard, J., Tratner, J., Perry, M., Badaracco, A. G., Farmer, C., Hjelle, G. A., Snow, A. D., Cochran, M., Gillies, S., Culbertson, L., Bartos, M., Eubank, N., maxalbert, Bilogur, A., ... Leblanc, F. (2020). Geopandas/geopandas: V0.8.1. https://doi.org/10.5281/zenodo.3946761

Jørgensen, T. L., & Rostad, M. (2021). *Tilstandsvurdering av kommunale vann– og avløpstjenester* (tech. rep.). bedreVANN. Retrieved November 23, 2022, from https://bedrevann.no/pdf/bedreVANN2021.pdf

Kabir, G., Tesfamariam, S., Hemsing, J., & Sadiq, R. (2020). Handling incomplete and missing data in water network database using imputation methods [Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/23789689.2019.1600960]. *Sustainable and Resilient Infrastructure*, *5*(6), 365–377. https://doi.org/10.1080/23789689.2019.1600960

Karimzadeh, A., Kaykhosravi, S., & Shoghli, O. (2017). Developing a Diurnal Pattern of Sewage Flow at Inlet Nodes Using Limited Measured Nodes–A Case Study of Greater Tehran Sewer System. Retrieved February 2, 2023, from https://www.researchgate.net/publication/325454918_Developing_a_Diurnal_Pattern_of_Sewage_Flow_at_Inlet_Nodes_Using_Limited_Measured_Nodes-A_Case_Study_of_Greater_Tehran_Sewer_System

Kartverket. (2023). Kartverket. Retrieved May 10, 2023, from https://hoydedata.no/LaserInnsyn2/

King County. (2021). What is infiltration and inflow? Retrieved February 1, 2023, from https://kingcounty.gov/services/environment/wastewater/ii/what.aspx

Klimaservicesenter. (2022). Klimaservicesenter. Retrieved November 25, 2022, from https://klimaservicesenter.no/kss/klimaprofiler/vestfold

Kracht, O., Gresch, M., & Gujer, W. (2008). Innovative tracer methods for sewer infiltration monitoring. *Urban Water Journal*, *5*(3), 173–185. https://doi.org/10.1080/15730620802180802

Lawande, S. R., Jasmine, G., Anbarasi, J., & Izhar, L. I. (2022). A Systematic Review and Analysis of Intelligence-Based Pathfinding Algorithms in the Field of Video Games [Publisher: Multidisciplinary Digital]. *Applied Sciences*, *12*(11), 5499. https://doi.org/10.3390/app12115499

Lee, H., Calvin, K., Dasgupta, D., Krinner, G., Mukherji, A., Thorne, P., Trisos, C., Romero, J., Aldunce, P., Barrett, K., Blanco, G., Cheung, W. W. L., Connors, S. L., Denton, F., Diongue-Niang, A., Dodman, D., Garschagen, M., Geden, O., Hayward, B., ... Zommers, Z. (2023). *Synthesis Report of the IPCC Sixth Assessment Report (AR6): Summary for Policymakers* [Publisher: Intergovernmental Panel on Climate Change]. https://www.ipcc.ch/report/ar6/syr/

Lepot, M., Makris, K. F., & Clemens, F. H. L. R. (2017). Detection and quantification of lateral, illicit connections and infiltration in sewers with Infra-Red camera: Conclusions after a wide experimental plan. *Water Research*, *122*, 678–691. https://doi.org/10.1016/j.watres.2017.06.030

Lindholm, O. (2017). Fremmedvann i avløpsledninger : VA-Miljø. Retrieved November 28, 2022, from https://www.va-blad.no/fremmedvann-i-avlopsledninger/

Lindholm, O. G., Jarle T. Bjerkholt, & Ole Lien. (2012). Fremmedvann i nordiske avløpsledningsnett. Retrieved November 27, 2022, from https://vannforeningen.no/dokumentarkiv/fremmedvann-i-nordiske-avlopsledningsnett/

Masasi, B., Frazier, R., & Taghvaeian, S. (2017). Review and Operational Guidelines for Portable Ultrasonic Flowmeters. https://extension.okstate.edu/fact-sheets/review-and-operational-guidelines-for-portable-ultrasonic-flowmeters.html

Meier, R., Tscheikner-Gratl, F., Steffelbauer, D. B., & Makropoulos, C. (2022). Flow Measurements Derived from Camera Footage Using an Open-Source Ecosystem [Number: 3 Publisher: Multidisciplinary Digital Publishing Institute]. *Water*, *14*(3), 424. https://doi.org/10.3390/w14030424

Meijer, D., Van Bijnen, M., Langeveld, J., Korving, H., Post, J., & Clemens, F. (2018). Identifying Critical Elements in Sewer Networks Using Graph-Theory [Number: 2 Publisher: Multidisciplinary Digital Publishing Institute]. *Water*, *10*(2), 136. https://doi.org/10.3390/w10020136

Miljødirektoratet. (2023). Norske utslipp, Avløpsannlegg. Retrieved November 22, 2022, from https://www.norskeutslipp.no/no/Avlopsannlegg/?SectorID=100

Mitchell, P. S., Stevens, P. L., & Nazaroff, A. (2007). QUANTIFYING BASE INFILTRATION IN SEWERS: A Comparison of Methods and a Simple Empirical Solution [ISSN: 1938-6478], 219–238. Retrieved May 30, 2023, from https://www.accesswater.org/?id=-294502&fromsearch=true#iosfirsthighlight

NGU. (2022). NGU - Kart min kommune. Retrieved November 24, 2022, from https://geo.ngu.no/kart/minkommune/

Nie, L., Lindholm, O., Lindholm, G., & Syversen, E. (2009). Impacts of climate change on urban drainage systems – a case study in Fredrikstad, Norway [Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/15730620802600924]. *Urban Water Journal*, *6*(4), 323–332. https://doi.org/https://doi.org/10.1080/15730620802600924

Pereira, A., Pinho, J. L. S., Faria, R., Vieira, J. M. P., & Costa, C. (2019). Improving operational management of wastewater systems. A case study. *Water Science and Technology*, *80*(1), 173–183. https://doi.org/10.2166/wst.2019.264

Pichler, M. (2022). Swmm-api: API for reading, manipulating and running SWMM-Projects with python. https://doi.org/10.5281/zenodo.7054804

Pulsar. (2023). FlowPulse | Clamp-on Flow Sensor | Pulsar. Retrieved April 12, 2023, from https://pulsarmeasurement.com/en/flowpulse#downloads

Ranck, C. (2017). Sanitary Sewer Systems: Rainfall Derived Infiltration and Inflow (RDII) Modeling [Publisher: Water Environment Federation]. Retrieved May 30, 2023, from https://www.accesswater.org/publications/-324221/sanitary-sewer-systems-rainfall-derived-infiltration-and-inflow-rdii-modeling

Ratnaweera, H., Zhang, D., Lindholm, G., & Martinez Almario, N. (2018). Exploiting Capacity of Sewer System Using Unsupervised Learning Algorithms Combined with Dimensionality Reduction [Accepted: 2019-02-22T13:36:30Z]. *1-29*. Retrieved May 30, 2023, from https://nmbu.brage.unit.no/nmbu-xmlui/handle/11250/2587071

Rokstad, M. M., & van Laarhoven, K. (2022). Technical note: Graph-theory-based heuristics to aid in the implementation of optimized drinking water network sectorization [Publisher: Copernicus GmbH]. *Drinking Water Engineering and Science*, *15*(1), 1–12. https://doi.org/10.5194/dwes-15-1-2022

Rossman, L. A., & Huber, W. C. (2016). Storm Water Management Model Reference Manual. https://nepis.epa.gov/Exe/ZyPDF.cgi?Dockey=P100NYRA.txt

Rossman, L. A., & Simon, M. A. (2022). Storm Water Management Model User's Manual Version 5. Retrieved November 8, 2022, from https://www.epa.gov/system/files/documents/2022-04/swmm-users-manual-version-5.2.pdf

Saletti, A. O. (2021). Infiltration and inflow to wastewater sewer systems - A literature review on risk management and decision support [Publisher: Chalmers University of Technology]. Retrieved May 30, 2023, from https://research.chalmers.se/en/publication/522159

Saletti, A. O., Lindhe, A., Söderqvist, T., & Rosén, L. (2023). Cost to society from infiltration and inflow to wastewater systems. *Water Research*, *229*, 119505. https://doi.org/10.1016/j.watres.2022.119505

Scherling, M. v., Svensson, G., Malm, A., & Røstum, J. (2020). *A 255 Bærekraftig fremmedvannsandel – modell for vurdering av riktig nivå | Norsk Vanns Kompetanseweb* (tech. rep.). Norsk Vann. Retrieved November 23, 2022, from https://va-kompetanse.no/butikk/a-255-baerekraftig-fremmedvannsandel-modell-for-vurdering-av-riktig-niva/

Schilling, J., & Tränckner, J. (2022). Generate_swmm_inp: An Open-Source QGIS Plugin to Import and Export Model Input Files for SWMM [Publisher: Multidisciplinary Digital Publishing Institute]. *Water*, *14*(14), 2262. https://doi.org/10.3390/w14142262

Schober, P., Boer, C., & Schwarte, L. A. (2018). Correlation Coefficients: Appropriate Use and Interpretation. *Anesthesia & Analgesia*, *126*(5), 1763–1768. https://doi.org/10.1213/ANE.0000000000002864

Skagsoset, C. (2022). *Feasibility study for quantifying infiltration and inflow* (Project Thesis).

Sola, K. J., Kvaal, K., Bjerkholt, J. T., Lindholm, O. G., & Ratnaweera, H. (2019). Identifying factors influencing infiltration and inflow (I/I-water) in wastewater systems using multivariate data analysis [Accepted: 2020-06-17T11:23:02Z], 263–278. Retrieved May 30, 2023, from https://nmbu.brage.unit.no/nmbu-xmlui/handle/11250/2658455

SSB. (2023). Kommunefakta. Retrieved February 8, 2023, from https://www.ssb.no/kommunefakta/horten

Sun, B., Chen, S., Liu, Q., Lu, Y., Zhang, C., & Fang, H. (2020). Review of sewage flow measuring instruments. https://doi.org/10.1016/j.asej.2020.08.031

Tekeli, S., & Belkaya, H. (1986). Computerized Layout Generation for Sanitary Sewers [Publisher: American Society of Civil Engineers]. *Journal of Water Resources Planning and Management*, *112*(4), 500–515. https://doi.org/10.1061/(ASCE)0733-9496(1986)112:4(500)

Turan, M. E., Bacak-Turan, G., Cetin, T., & Aslan, E. (2019). Feasible Sanitary Sewer Network Generation Using Graph Theory. https://doi.org/10.1155/2019/8527180

Van Assel, J., Kroll, S., & Delgado, R. (2023). Calculation of Dry Weather Flows in Pumping Stations to Identify Inflow and Infiltration in Urban Drainage Systems [Number: 5 Publisher: Multidisciplinary Digital Publishing Institute]. *Water*, *15*(5), 864. https://doi.org/10.3390/w15050864

Vannportalen. (2023). Retrieved January 26, 2023, from https://www.vannportalen.no/vannregioner/vestfold-og-telemark/vannomrader/horten---larvik/plandokumenter/planperioden-2022---2027/

Weiss, G., Brombach, H., & Haller, B. (2002). Infiltration and inflow in combined sewer systems: Long-term analysis. https://iwaponline.com/wst/article/45/7/11/9869/Infiltration-and-inflow-in-combined-sewer-systems

Zaidan, K. (2018). *Analyse av overvannsnettet i Holtandalen i Horten kommune ved bruk av simuleringsprogrammet SWMM* (Master's thesis). Retrieved April 18, 2023, from https://nmbu.brage.unit.no/nmbu-xmlui/handle/11250/2569478

Appendix A

Extended Results: Supplementary Visualizations

This document presents additional plots that provide a more comprehensive visualizations of the data obtained during the study. These plots include supplementary information that to what is presented in the main Results chapter.

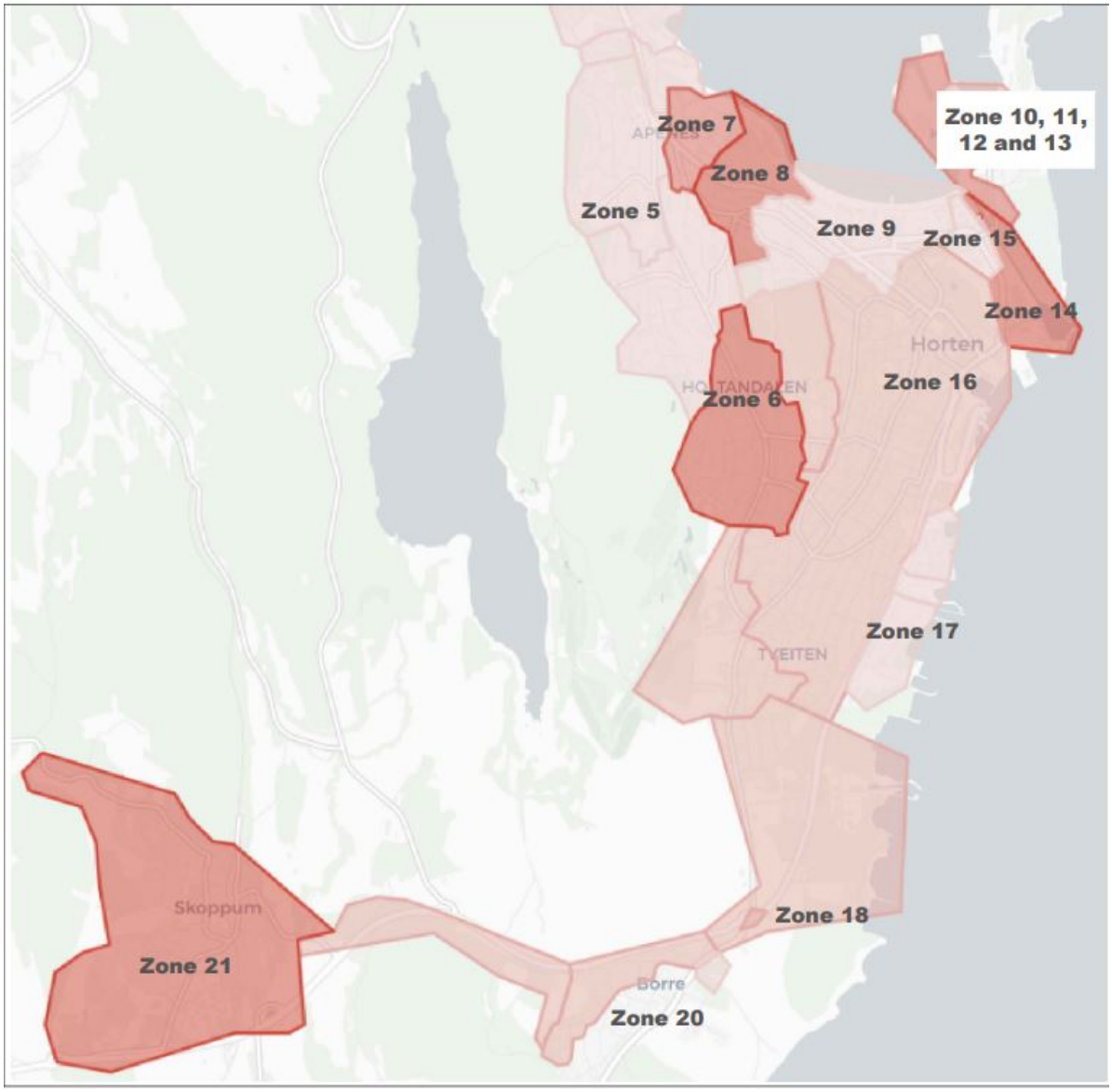


Figure A.1: Correlation to Precipitation in Major Zones: Color Strength Indicates Correlation Strength

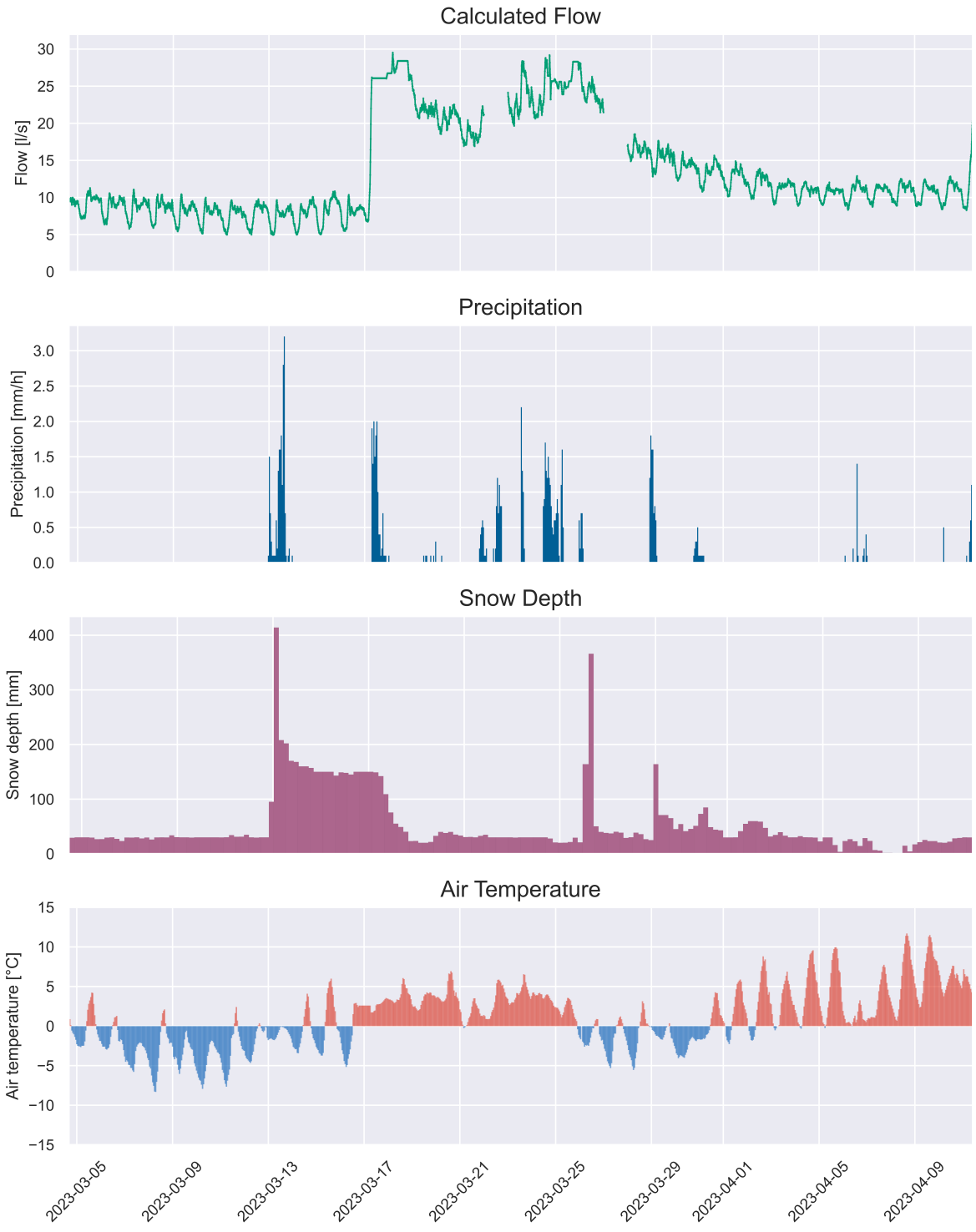


Figure A.2: Flow Measurements: Compared to Data from Horten Weather Station

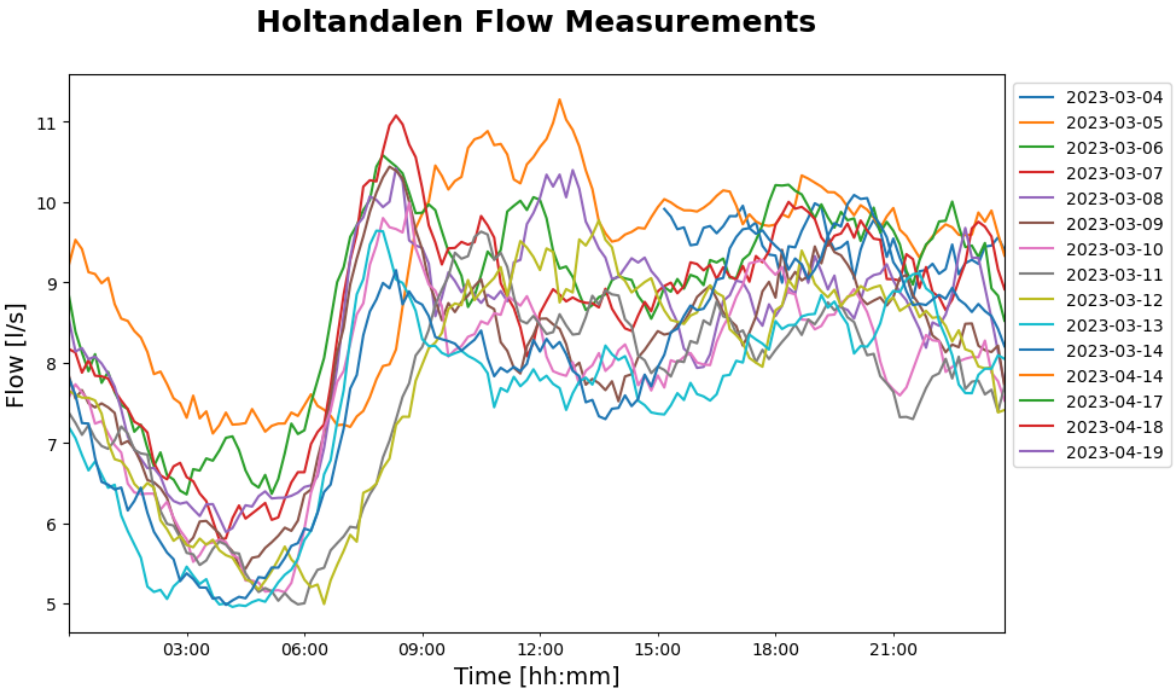


Figure A.3: Flow Measurements: Dry weather Flow

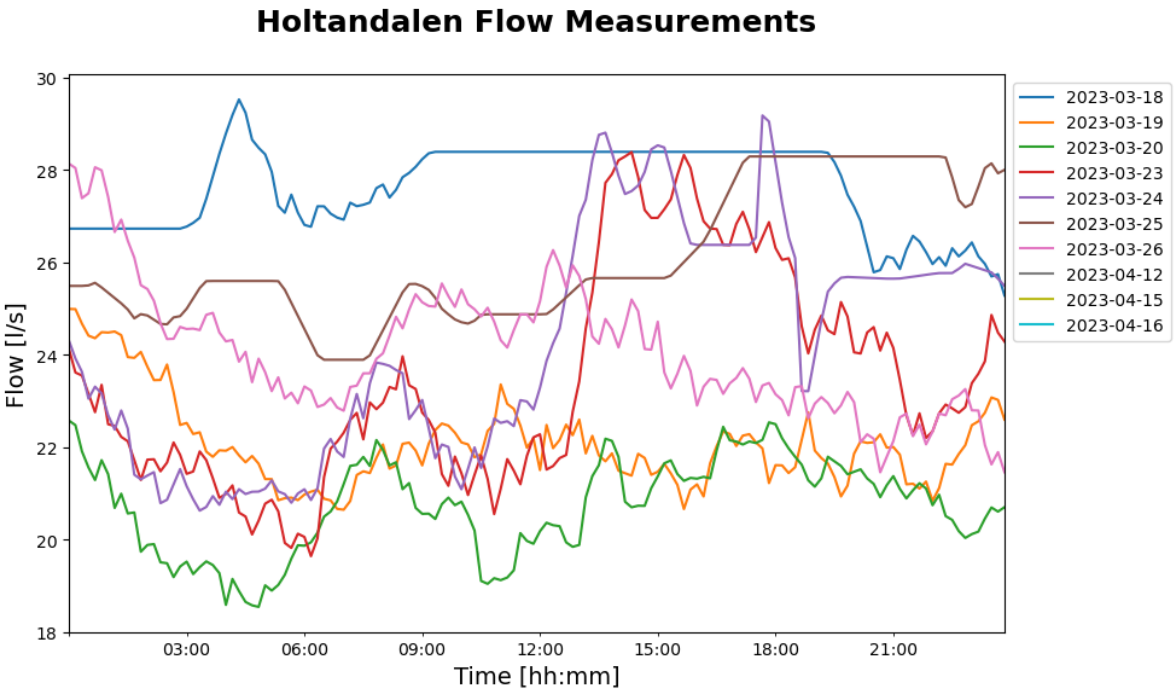


Figure A.4: Flow Measurements: Wet Weather Flow

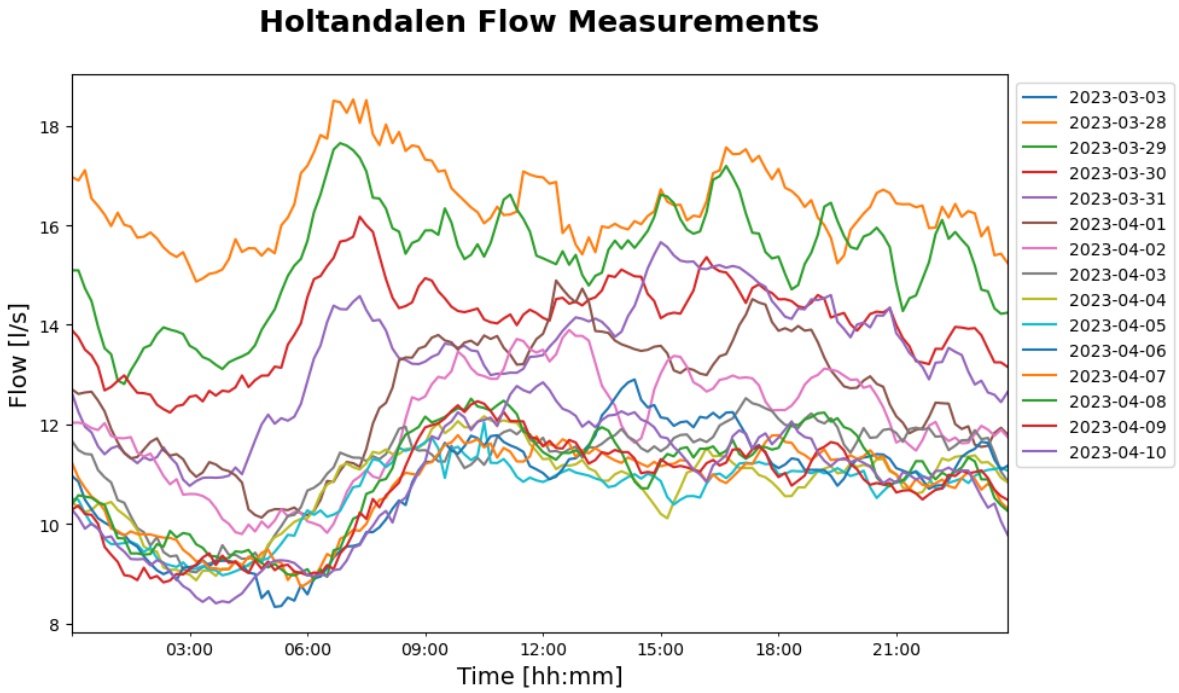


Figure A.5: Flow Measurements: Days with Snow Melt

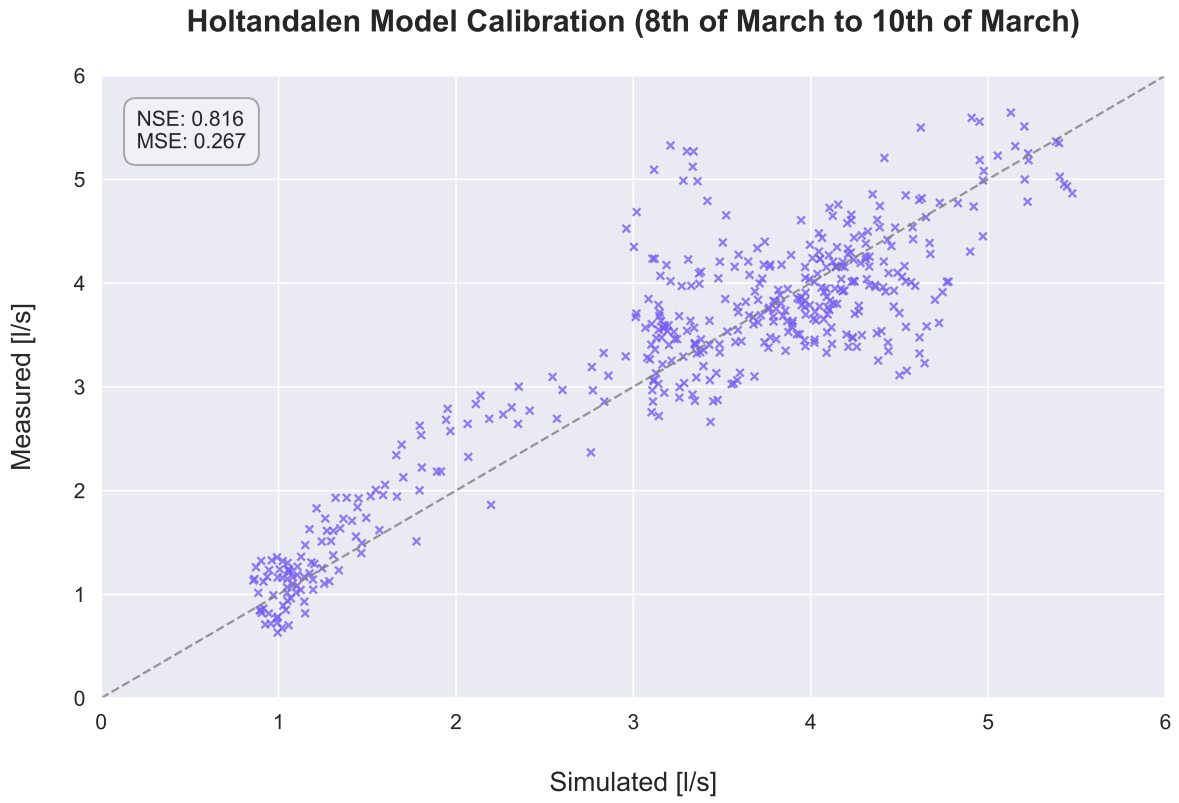


Figure A.6: Holtandalen Calibration Scatter Plot: Relationship between measurements and simulation results (08.03.2023 - 10.03.2023)

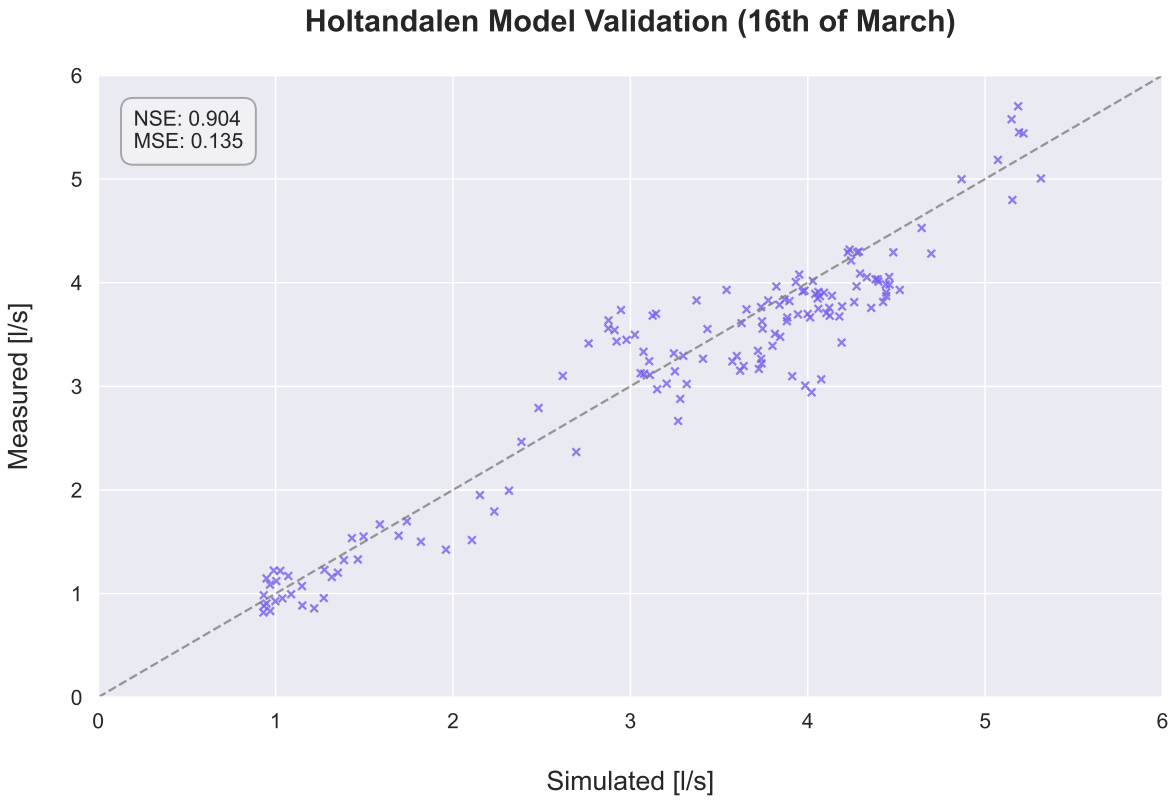


Figure A.7: Holtandalen Validation Scatter Plot: Relationship between measurements and simulation results (16.03.2023)

Appendix B

Calculation Procedure for the Inflow Rate

This document provides an explanation of the calculations used to determine the inflow rate in the pumping station for this study. The focus of this document is to shed light on the process of obtaining measurements. Specifically, it describes the calculations performed to determine the inflow rates during different cycles, including the filling cycle, emptying cycle, and the estimation of values in-between cycles. These steps are adaptations of the liquid level calculations described in Coughanowr and LeBlanc (2009) and the procedure of Van Assel et al. (2023).

Filling cycle

The filling cycle refers to the period during which the sump tank accumulates sewage primarily due to inflow from the connected sewer system. During this cycle, the sump storage gradually increases until it reaches a certain level, triggering the pumps and initiating the emptying cycle. This is how the inflow rate in the filling cycle is calculated:

The change in height is given by:

$$\Delta h = h_{t_1} - h_{t_0} \tag{B.1}$$

where h_{t_1} is the water level at time t_1, and h_{t_0} is the water level at time t_0. The water level in the pumping station sump is often reported with a one minute resolution.

The time difference is given by:

$$\Delta t = t_1 - t_0 \tag{B.2}$$

The change in volume is given by:

$$\Delta V = \Delta h \cdot A \cdot 1000 \tag{B.3}$$

Here, A represents the cross-sectional area of the container. In this particular case, the

sump had a pre-fabricated cylindrical design, and the water levels did not fall below the threshold where the cross-section changes. The factor of 1000 is used to convert cubic meters to liters.

The inflow rate is given by:

$$Q_{\text{inflow}} = \frac{\Delta V}{\Delta t} \tag{B.4}$$

Emptying cycle

When the pump is triggered in response to a rising sump level, the inflow rate to the sump typically decreases. The pump is designed to remove the accumulated sewage from the sump and transfer it to a downstream pipe. As the measured flow rate from the pumps and sump level is not reported simultaneously, they are not necessarily comparable. Therefore, the volume change throughout the cycle is used to calculate the flow rate for this period. Hence, constantly running pumps does provide flat curves, as the flow rate is set to the same throughout this period. The measured change in volume is given by:

$$\Delta V_{\text{measured}} = \frac{\sum \text{measured flow}}{1000} \tag{B.5}$$

where the measured flow is the sum of flow measurements over the cycle period.

The measured change in height is given by:

$$\Delta h_{\text{measured}} = \frac{\Delta V_{\text{measured}}}{A} \tag{B.6}$$

where A is the cross-sectional area of the sewage sump.

The total change in height is given by:

$$\Delta h = h_{\text{end}} - h_{\text{start}} + \Delta h_{\text{measured}} \tag{B.7}$$

where h_{end} is the water level at the end of the cycle period, and h_{start} is the water level at the start of the cycle period.

The change in volume is given by:

$$\Delta V = \Delta h \cdot A \cdot 1000 \tag{B.8}$$

where the factor 1000 is used to convert cubic meters to liters.

The inflow rate is given by:

$$Q_{\text{inflow}} = \frac{\Delta V}{\Delta t} \tag{B.9}$$

where Δt is the cycle period.

In between cycles

Interpolated values are used to fill in the gaps between flow calculations taken at specific intervals. By assuming a linear relationship between flow rates from different cycles, flow rates for these intermediate periods were estimated.

Appendix C

A Guide to Understanding Specialized Terminology

This document offers explanations of the concepts discussed in the main text, aiming to help readers understand specialized terminology effortlessly, without the need for external sources. Furthermore, it serves as a centralized reference for important terms, enhancing the overall comprehension of the reader. It encompasses a wide range of topics, including diurnal wastewater flow (C.1), dry-weather flow (DWF) (C.1.1), Groundwater infiltration (GWI), Minimum night-time flow (MNF) (C.1.2), Pearson correlation coefficient (C.2), Doppler ultrasonic flowmeters (C.3), graph theory (C.4), and evaluation metrics such as Nash–Sutcliffe Efficiency (NSE) and Mean Square Error (MSE) (C.5).

C.1 Diurnal Wastewater Flow

The diurnal wastewater flow pattern represents the daily fluctuations in the volume of wastewater transported through the sewer system (Butler et al., 2018). For residential areas, the pattern usually exhibits two peaks during the day and a minimum flow rate during night, as depicted in Figure C.1. The sewer catchment area is comprised of several branches in a network, each exhibiting distinct base sanitary flow (BSF) characteristics. Furthermore, due to distance between the measurement point and the sources, the actual flow patterns may be distorted and the measurement presents a mixture of multiple flow patterns. Hence, it could be difficult to distinguish the different sources based on flow measurements from a single point.

Nevertheless, the diurnal flow pattern of BSF can be used to estimate the quantity of infiltration and inflow (I/I) within an existing sewer system (Butler et al., 2018; Karimzadeh et al., 2017). This requires the analysis of the flow patterns in periods of DWF, which can be used to calculate the BSF by subtracting the estimated GWI (Mitchell et al., 2007).

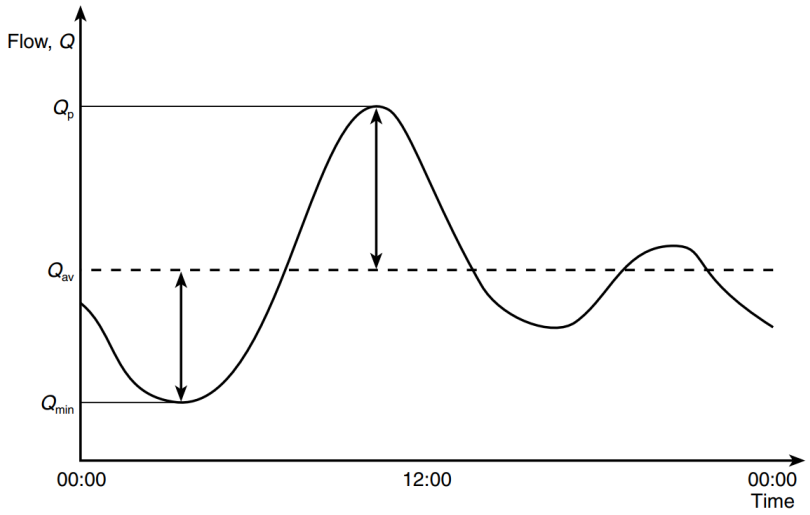


Figure C.1: Example of a residential diurnal wastewater flow pattern (Butler et al., 2018)

C.1.1 Dry Weather Flow

DWF is a combination of wastewater generated by households and industries and a component of GWI. To achieve accurate estimation of DWF, longer monitoring periods are recommended, since seasonal variations of the GWI component are to be expected (Hey et al., 2016). In conventional flow methods, calculation of DWF can be used to find variations and determine trends in the flow rate during dry weather conditions (Beheshti et al., 2015; Hey et al., 2016). Which could further be used to estimate the impact of I/I in wet weather conditions. Hence, wet weather flow is DWF with the inclusion of rainfall-derived infiltration and inflow.

DWF refers to the average flow rate of wastewater during periods with little to no impact from precipitation (Mitchell et al., 2007; Weiss et al., 2002). This time period is commonly based on a delay of a few days after a precipitation event (Bogusławski et al., 2022; Lindholm, 2017). However, Haas et al. (2021) describes alternative methods for estimating DWF, e.g. basic statistical methods or K-means clustering.

C.1.2 Groundwater Infiltration & Minimum Night-Time Flow

GWI is often assumed to be constant throughout a day with dry weather and can serve as a baseline flow for the DWF calculation. By assuming that a substantial portion of the GWI originates from groundwater and that there is minimal BSF during night-time, the GWI can be quantified through the calculation of the MNF (Hey et al., 2016).

The MNF is determined by analyzing daily sewer flow data and identifying the lowest point in the diurnal pattern. This method is effective in areas primarily comprised of residential connections (Mitchell et al., 2007). More accurate estimates of GWI can be obtained through the use of empirical formulas, rather than just using the MNF as the GWI.

C.2 Pearson Correlation Coefficient

The Pearson correlation is commonly used to find correlations between sets of variables and within sets (Cohen et al., 2002; Schober et al., 2018). Specifically, the Pearson correlation coefficient can be used to identify correlations between the flow rates and relevant variables. The mathematical calculations for the Pearson correlation can be seen in Equation C.1.

$$r_{xy} = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n\sum x_i^2 - (\sum x_i)^2} \cdot \sqrt{n\sum y_i^2 - (\sum y_i)^2}} \tag{C.1}$$

r_{xy} = Pearson correlation coefficient between x and y
n = Number of observations
x_i = Values for x
y_i = Values for y

As Table C.1 shows, a good linearity between parameters can be in both directions. Meaning that both high positive and negative values indicates that there are strong relationships between the variables.

Table C.1: Explanation of Pearson correlation results (Skagsoset, 2022)

Pearson's r Value	**Correlation Between x and y**
Equal to 1	Perfect positive linear relationship
Greater than 0	Positive correlation
Equal to 0	No linear relationship
Less than 0	Negative correlation
Equal to -1	Perfect negative linear relationship

C.3 Doppler Ultrasonic Flowmeters

Measuring sewer flow are often a difficult task due to sewage pipes are mostly buried and the flow is usually gravity driven (Sun et al., 2020). Hence, its mainly open channels with a variable flow and the optimal measurement location is difficult to reach. To measure the flows in these cases, intrusive measurement devices are often used, which can get affected by the flow. Alternatively, there are occasionally filled pressurized pipes related to sewer pumps, which offer additional possibilities for flow measurement since these pipes are filled and pressurized.

Ultrasonic flowmeters are commonly used to measure pressurized flow without making physical contact with the flow medium, making the method non-intrusive (Sun et al.,

2020). A transducer sends a signal through the pipe, and the meters measure the frequency change of the signal to determine the flow velocity (Masasi et al., 2017). This information is then multiplied by the pipe's cross-section to calculate the flowrate. Hence, accurate flow rates require a precise measurement of the pipe's internal diameter. The measured medium must contain sufficient particles to bounce the signal off, as visualized in Figure C.2. On the other hand, air pockets and sediments can cause inaccuracies, and the transducer's positioning is therefore vital as it should not be located at the top or bottom of the pipe where there could be disturbances to the signal. Additionally, turbulence in the flow can significantly impact the quality of measurements, so it's often recommended to place the device at least 10 times the diameter of the pipe away from valves, bends, or pumps. Because of practical reasons this is often not possible, and a correction factor is calibrated to reduce this effect.

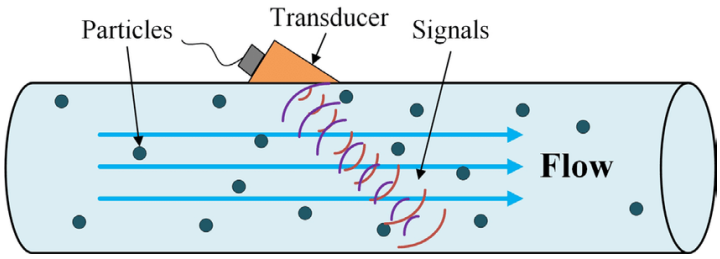


Figure C.2: Doppler ultrasonic flowmeters (Masasi et al., 2017)

C.4 Graph Theory

Graph theory, which is a mathematical concept, has broad applications in fields related to connectivity and routing problems, as demonstrated by the famous Königsberg bridge problem (Harju, 2014; Meijer et al., 2018). Furthermore, the concept is also used in modern search engines such as Google, which uses a PageRank algorithm to retrieve the most relevant web pages (Brin & Page, 1998).

A graph is a structure that represents a network and its connectivity trough nodes and links, called vertices and edges, respectively (Meijer et al., 2018). A sewer network can modeled as a graph, where the manholes usually represents the vertices and the pipes represents the edges between the vertices (Turan et al., 2019). A graph can either be directional or undirectional, with directional graphs having edges that only work in one direction and undirectional graphs allowing edges to go both ways, creating cycles between vertices. Figure C.3 illustrates this difference.

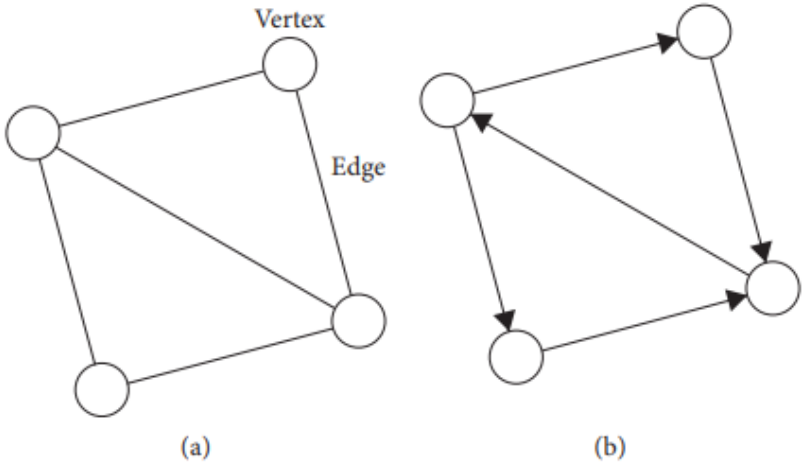


Figure C.3: Undirected graph (a), Directed graph (b) (Turan et al., 2019)

A base graph is an undirected graph that includes all possible vertices and links. On the other hand, directed graphs without cycles are known as trees or spanning trees (Tekeli & Belkaya, 1986; Turan et al., 2019). For modeling sewer networks, directed graphs are particularly useful due to their ability to represent flow paths to a given root vertex. Adding weights to edges allows for the definition of a travel "cost" between vertices, enabling algorithms like Prim's Algorithm to find the minimum spanning tree with the lowest "cost" (Furqan et al., 2018). Other algorithms, such as Breadth First Search, Depth First Search or Dijkstra's Algorithm, can also be used to find the shortest paths through the graph when weights are present (Lawande et al., 2022).

C.5 Nash–Sutcliffe Efficiency & Mean Square Error

NSE is a statistical metric used to assess the model's ability to explain the variance in the observations relative to employing the mean as a prediction (Bennett et al., 2013). A value of one indicates a perfect fit between the model and the simulation, while a value of zero implies that the simulation performs no better than using the mean of the measurements for predictions. NSE compares the squared differences between the observed values (y_i) and the predicted values ($\hat{y}$) with the squared differences between the observed values and their mean ($\bar{y}$). It ranges from negative infinity to one, with higher NSE values indicating superior model performance.

$$\text{NSE} = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{C.2}$$

MSE is a statistical metric used to quantitatively assess how well a model's predictions align with the actual observations (Bennett et al., 2013). Spesifically, it measures the average squared difference between the observed values (y_i) and the related predicted values ($\hat{y}$). Mathematically, MSE is calculated by taking the squared differences between each observed value and its corresponding predicted value, summing these squared differences, and then dividing by the total number of observations (n):

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{C.3}$$

MSE considers both positive and negative errors, squaring them to make them positive, and then averaging them (Bennett et al., 2013). This weighting penalizes larger errors more heavily, as the squared term amplifies their contribution to the overall error. In practical terms, a lower MSE value indicates that the model's predictions are closer to the actual observations, suggesting a better fit or performance of the model. On the other hand, a higher MSE value indicates larger prediction errors and worse model performance. MSE is widely used in various fields, including statistics, machine learning, and data analysis, as goodness-of-fit criteria, similar to NSE.

Appendix D

Technical Specifications

This document provides the technical specifications for the FlowPulse Ultrasonic Clamp-on Flow Sensor and the FlowPulse Handheld Controller, as referenced from Pulsar (2023).

Table D.1: Technical Specifications for FlowPulse Ultrasonic Clamp-on Flow Sensor (Pulsar, 2023)

Specifications	**Performance**
Accuracy	± 5% typical subject to installation and pipe conditions
Resolution	3 mm/s
Velocity Range	300 mm/s to 4 m/s
Response Time	Fully adjustable (1-second minimum)
Minimum Particle Size	>100 µm
Minimum Particle Concentration	>200 ppm
Pipe Diameter	30 mm to 350 mm
Min. range	0.3 m/s
Max. range	4.0 m/s
Pipe Wall Thickness	Metal or rigid pipe up to 20 mm thick
Signal Processing	RSSA (Refracted Spread Spectrum Analysis)
Power Supply	18 - 24 V DC
Power Consumption	2.4 W at 24 V

Table D.2: Technical Specifications for FlowPulse Handheld Controller (Pulsar, 2023)

Specifications	Performance
Accuracy	± 0.25% of the measured range or 6 mm, whichever is greater.
Resolution	± 0.1% of the measured range or 2 mm, whichever is greater
Max Range	Dependent on application and transducer
Min Range	Dependent on application and transducer
Rate Response	Fully Adjustable
Storage Media	Internal flash memory
Storage Capacity	3.8 GB
Storage Access	File transfer to PC via USB — no driver required

Appendix E

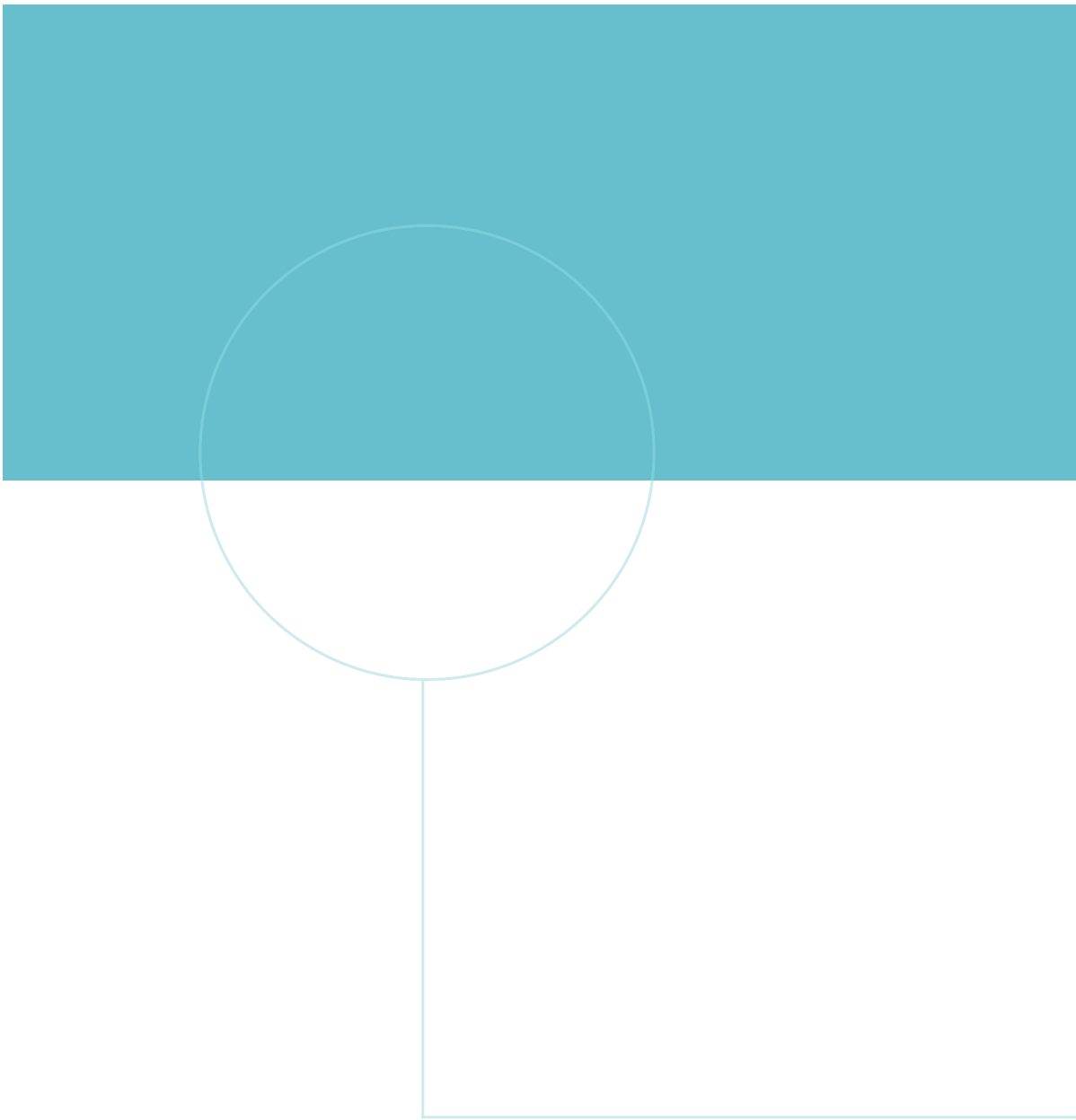
Feasibility study for quantifying infiltration and inflow

The Project Thesis, conducted as part of the course *TBM4500 - Civil and Environmental Engineering*, served as the foundational work for this Master's Thesis. The Project Thesis document is included as a separate file.

Appendix F

Python Scripts

Python scripts are included as a separate zip file.



NTNU

Norwegian University of
Science and Technology