



Continuous simulation of urban sanitary sewer network

PEDRO PAULO ALMEIDA SILVA

August 15th, 2019

BRANDENBURG UNIVERSITY OF TECHNOLOGY
COTTBUS-SENFTENBERG
Erasmus Mundus Joint Master Programme

This dissertation was accepted as the final thesis report of the degree of Master of Science in Hydroinformatics and Water Management.

Academic Supervisor: Frank Molkenthin, PhD
Head of EuroAquaee Programme at
Brandenburgische Technische Universität , Cottbus, Germany

Institutional Supervisor: Markus Sunela, PhD
Chief Technology Officer, Fluidit Oy,
Tampere, Finland

Thesis Presentation: August 21st, 2019

Funding:

Euroaquaee is a joint Master Programme funded by Erasmus+. This dissertation also received support from Fluidit Oy company where the student performed a professional practice with the title of Master Thesis Worker.



Funded by the
Erasmus+ Programme
of the European Union



Declaration:

I hereby declare that this work was written independently without any unauthorized assistance or sources not given credit within the work. All words, phrases, passages, and data taken from other sources have been properly cited. No parts of this work in the same form or a similar form have ever been previously handed in to fulfill an examination.

Pedro Paulo Almeida Silva



Brandenburgische
Technische Universität
Cottbus - Senftenberg



POLYTECH[®]
NICE SOPHIA



Contents

Contents	ii
List of Tables	iv
List of Figures	v
Abstract	vii
Dedication	vii
Acknowledgements	viii
Acronyms	ix
1 Introduction	11
1.1 Motivation	11
2 Methodology	14
2.1 Offline Model	14
2.2 Online Model	15
2.3 Parameter Optimization Algorithm	17
3 Hydrological Model	20
3.1 Methods to Quantify Rainfall Dependent Infiltration and Inflow	21
3.2 Physically-Based: SWMM Modules	23
3.2.1 Rainfall-Runoff	23
3.2.2 Snowpack & Snowmelt	25
3.2.3 Infiltration	27
3.2.4 Aquifer & Groundwater Flow	29
3.3 Synthetic Unit Hydrograph: RTK	30
4 Hydraulic Model	32
5 Case Study	33
5.1 Jokela Town	33
5.2 Data and Parameter Estimation	34
5.2.1 Sanitary Sewer Flow Data	34
5.2.2 Meteorological Data & Parameters	37
5.2.3 Weather Forecast	41
5.2.4 Topographic Data & Parameters	42
5.2.5 Water Table & Groundwater Flow Parameters	49
5.2.6 Synthetic Unit Hydrograph: RTK	53
5.3 Calibration	55

5.3.1	Calibration of physically based model	55
5.3.2	Calibration of RTK Model	57
5.4	Validation	58
5.5	Validation with historical Meteorological Data	58
5.6	Validation with historical Weather Forecast Data	58
6	Results and Discussion	61
6.1	Water Balance	61
7	conclusion	62
7.1	Conclusion	62
7.2	Recommendations	62
Bibliography		63

List of Tables

3.1	Snowpack & Snowmelt parameters range[25]	27
3.2	Modified Horton infiltration parameters range[25]	29
5.1	Monthly estimated amount of Sanitary Sewer network (SSN) flow components in 2018 for Jokela Catchment in [m ³]	38
5.2	Snowpack & Snowmelt estimated parameters [25]	41
5.3	SWMM Runoff parameter estimation	46
5.4	Modified Horton infiltration parameter estimation	48
5.5	Groundwater Flow parameter estimation	53
5.6	Range of RTK parameters by Vallabhaneni and Burgess [32]	55
5.7	RTK rain 04 simulation results evaluation. Nash-Sutcliffe = 1 means a perfect model	57

List of Figures

1.1	Wet-weather Infiltration	12
2.1	Hydrological Model Development Flowchart	15
2.2	System Diagram	17
3.1	Wet-weather flow components. Modified from [32]	20
3.2	Precipitation Losses relative to a Sanitary Sewer Network	22
3.3	Nonlinear reservoir model [25]	23
3.4	Extra D_s and $\%_{routed}$ for Nonlinear reservoir model. Modified from Rossman and Huber [25]	24
3.5	Seasonal variation of melt coefficients [25]	25
3.6	Horton infiltration capacity decay and recovery curves. Modified from [25]	28
3.7	Representation of two-zone groundwater model in SWMM [25]	30
3.8	RTK Short, Medium, Long-term and Resulting Hydrographs [32]	31
5.1	Vantaa Basin	33
5.2	RDII x DWF estimate for 2018 of total waste water load	35
5.3	RDII x DWF estimate for 2018 of total wastewater load	35
5.4	Estimated Dry-weather Hydrograph from EPA SSOAP tool	36
5.5	RDII x DWF estimate for 2018 of total wastewater load of Jokela SSN	37
5.6	Snow depth and Temperature Measurements FMI [13]	39
5.7	Results of Initial parameter estimation x manual calibration of Snowpack & Snowmelt parameters	40
5.8	HARMONIE weather model forecast coverage	41
5.9	Topographic data used per each Modeled Process	42
5.10	Jokela Subcatchment division	43
5.11	Comparison of sewershed delineation using Filled and Burned DEM	44
5.12	Delineation methods D1 (left) and D3 (right)	45
5.13	Pervious and Impervious division (left) and level 3 from SYKE's corine land cover for Jokela catchment (right) [27]	46
5.14	Jokela catchment soil superficial deposits	47
5.15	Location of Observation Wells around Jokela's catchment	49

5.16	Groundwater table measurements. Two upper plots are measurements from 2018 only whereas the other two are monthly mean values of the period informed. (A = H12, B = 0110651, C = 0118651). Data from Tuusulan Vesihuolto and SYKE [27]	50
5.17	Dupuit-Forcheimer lateral seepage to adjacent channel. Modified from	51
5.18	Position of h2 for each subcatchment	52
5.19	Groundwater table first simulation from January to May of 2018	54
5.20	Representation of Jokela's Hydrological Model using RTK UH method	55
5.21	Long-term simulation of winter period evaluated at Jokela Pumping station using Dupuit-Forcheimer and unidirectional flow condition. Simulated = Orange, Observed = Blue	56
5.22	Long-term simulation of winter period evaluated at Jokela Pumping station using different GWI equation and unidirectional flow condition. Simulated = Orange, Observed = Blue	56
5.23	Rain 04	57
5.24	Calibration results of RTK parameters using Rain04 and 100 iterations of DDS optimization algorithm	58
5.25	Calibration results of RTK parameters using Rain04 and 300 iterations of DDS optimization algorithm	58
5.26	Calibration results of RTK parameters using Rain04 and 600 iterations of DDS optimization algorithm	59
5.27	Calibration results of RTK parameters using Rain04 and 1000 iterations of DDS optimization algorithm	60

ABSTRACT

[To Do]

will be completed when final results are obtained.

DEDICATION

Dedication goes for my family that provided me with full support from my home country during the course of this work. Especially to my mother Maria Almeida da Silva who had to be far from her son.

ACKNOWLEDGEMENTS

The author acknowledges EuroAqua+ Joint Master Programme and the European Union for the organisation and funding which made this study possible for me and provided the same conditions to many other students.

Professors, students, assistants, and staff of EuroAqua+ programme from all five participant universities played important role by teaching and organising this challenging programme and master thesis which involved six different countries. Special gratitude to Dr. Frank Molkenthin from Brandenburg Technical University for his tutoring during this thesis.

Fluidit Oy company in Finland and all its employees who contributed in different parts of this study. Acknowledgments to Dr. Markus Sunela for his patience and to gently guidance. Kalervo Kylätie and Timo Ranta-Pere for their proactive help and availability during the project. Janne Väyrynen, Mika Kuronen, Panu Kuitunen, Johan Ånäs, and Lauri Rantala helped either explaining the case study hydraulic model, development of scripts or valuable opinions.

Author would like to thank Tuusula Water Utility (Tuusulan Vesihuolto) and Hanna Riihinen for providing the data used and their sewer network.

ACRONYMS

BWF	Base Wastewater Flow
CSN	Combined Sewer network
CSO	Combined Sewer Overflow
DDS	Dynamically Dimensioned Search
DEM	Digital Elevation Model
DWF	Dry-Weather Flow
EPA	U.S. Environmental Protection Agency
FMI	Finnish Meteorological Institute
GTK	Geological Survey of Finland
GWI	Groundwater Infiltration
HD-DDS	Hybrid Discrete Dynamically Dimensioned Search
I/I	Infiltration and Inflow
NLS	Finnish National Land Survey
NWS	U.S. National Weather Service
RDII	Rainfall-derived Infiltration and Inflow Real
SCADA	Supervisory Control and Data Access
SSN	Sanitary Sewer network
SWSN	Stormwater Sewer network
SSOAP	Sanitary Sewer Overflow Analysis and Planning
SSO	Sanitary Sewer Overflow
SYKE	Finnish Environment Institute
SWMM	Stormwater Management Model
TMF	Traffic Management Finland
UH	Unit Hydrograph
WWF	Wet-Weather Flow
WWTP	Wastewater Treatment Plant
XML	Extensible Markup Language

1 INTRODUCTION

Modelling sanitary sewer network flows allow cities to understand and solve issues that impact its society, environment and economy. To model sanitary sewer network flows, it is first necessary to identify the behavior of the system. As described by Vallabhaneni and Burgess [32] the flows in the network usually have two different behaviors that are classified as: 1. dry-weather flow (DWF); 2. wet-weather flow (WWF). To apply a continuous simulation and successfully forecast the flows in the sanitary sewer system, the model should be able to account for both DWF and WWF.

Typically, DWF pattern can be estimated by analyzing historical data from flow measurements available along the network or at the downstream end during dry periods [7]. Intuitively, usually more water is discharged into the sewer in day time than during the night. More complexity is added when trying to estimate WWF in the sanitary sewer network. Flow increases in the network due to inflow and infiltration often triggered by rainfall or snowmelt. This incremental quantity of flow which finds its way into the sanitary sewer network during a rainfall event is known by Rainfall Dependent Inflow and Infiltration (RDII).

The challenge on RDII estimations lies upon the different ways stormwater enters a sanitary or combined sewer [19]. Stormwater can inflow to the sanitary or combined sewer network directly through foundation and roof drains connections, leaky manhole covers, or stormwater drains. Infiltration occurs due to defects in the network components such as: damaged pipes, joints, manholes, etc. [25] Quantity of inflow and infiltration (I&I) increases proportionally with intensity of the rainfall and snowmelt. Urban areas located in cold climate can have a significant inflow due to snowmelt and was, therefore, considered during this study. For that, the incremental flow caused by snowmelt will be referred here as Snowmelt Dependent Inflow and Infiltration (SDII). Event-based Inflow and Infiltration (EBII) will be used as a general term for both snowmelt and rainfall inflow and infiltration.

1.1 Motivation

Event-base inflow and infiltration (EBII) causes increase of the quantity of flow into sanitary sewer systems. This increase may cause sanitary sewer overflow (SSO) due to exceedance of the system capacity [25]. In some cities, sanitary sewer is combined with stormwater sewer network. The capacity of this coupled systems may also be exceeded during an event causing combined sewer overflow (CSO) [32]. When the capacity is exceeded, untreated water rejected by the wastewater treatment plants (WWTPs) is released into surface waters [9] such rivers and streams. Upstream capacity-related issues may cause wastewater to find its way into basements or streets [24]. Untreated water released on the surface water bodies or urban area increases the risk of human contamination by infectious

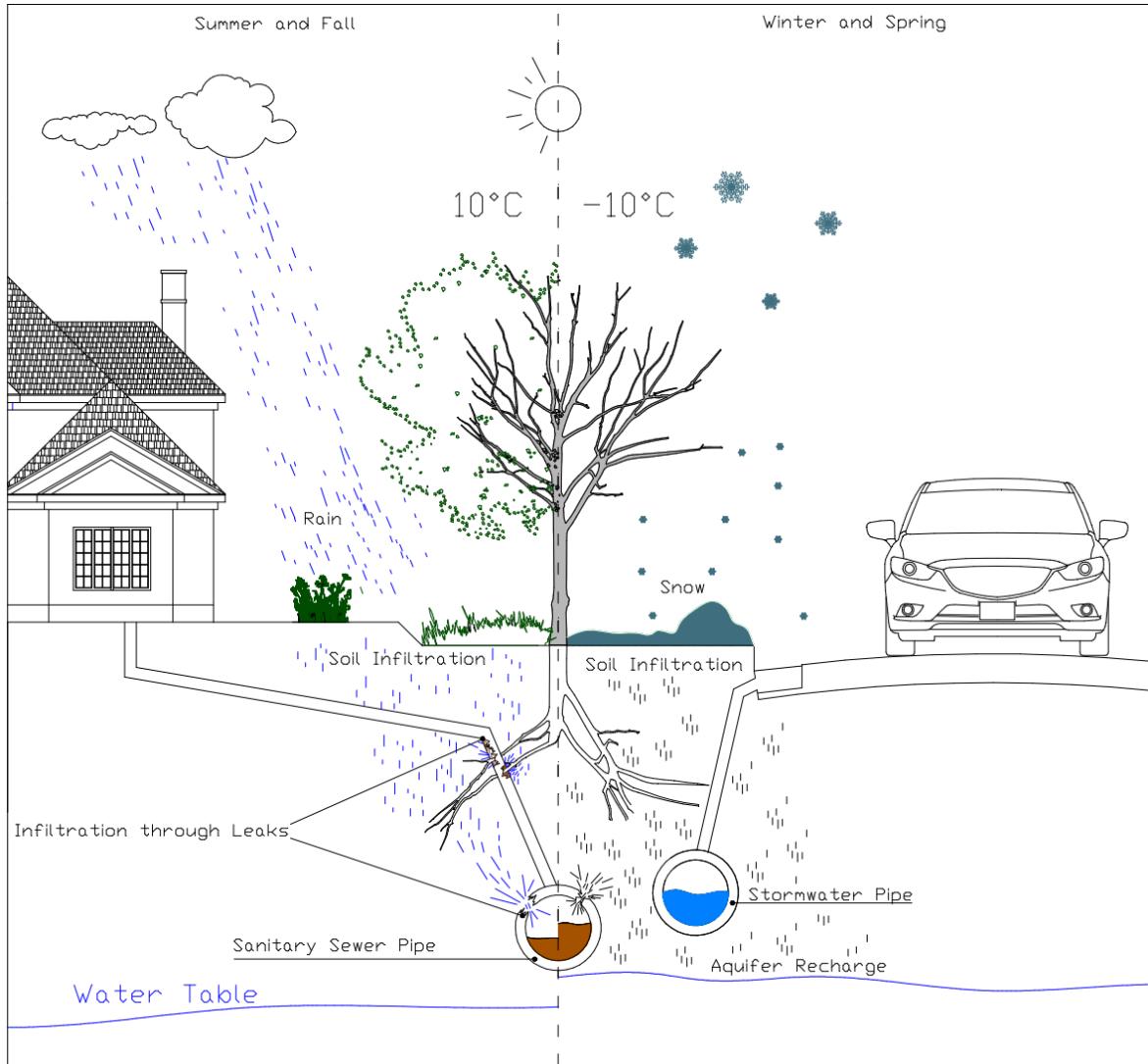


Figure 1.1: Wet-weather Infiltration

diseases. Although these are well known problems, they are still present in urban centers around the globe. Frequency of CSOs are seasonal dependent. The current prediction of more frequent intensive rainfall events also increases the frequency of CSO [9] and enlarge the damage caused by these overflows. Moreover, wastewater overflows can cause conflicts in the society when streams or rivers are used both as an option to dispose wastewater and recreational area as described by Heikkinen et al. [16]. To give cities, municipalities and water utilities the ability to predict SSOs and CSOs is one of the motivations of this study. However, other benefits are also aimed. A continuous simulation considering urban hydrology and hydraulics can also be used to improve the service and reduce operational costs for water utilities. EBII can increase the inflow of wastewater to WWTPs for weeks due to possible long response times [19]. Obviously, the operational cost of the plant will increase since more wastewater needs to be treated. A continuous simulation might be able to identify increases

over time on the flow pattern in specific pipes or sub-divisions of the network. This would be an indicator for the water utility to carry further inspections and evaluate whether the infrastructure is damaged allowing infiltration. Furthermore, a digital model of the network gives to the water utilities the ability to analyze the impacts on the whole network caused by changes in the network such as: decommissioning of a water tower, changing pumping schedule, or analyzing impacts of a future new neighborhood.

2 METHODOLOGY

To tackle the issues presented in the previous section, this study will investigate the main aspects behind the development of two hydraulic models with the aim to approximate the behavior of the sanitary sewer network flows continuously in a cold climate influenced by annual snowmelt periods. In the future, the model should be able to cope with existing monitoring infrastructure of Supervisory Control and Data Access (SCADA) and weather forecast data as input to predict the future status of the network. Even though this study does not go as far as description of real implementation of an integrated operational and CSOs/SSOs early warning system, it focuses the development and data acquisition of the hydrological models in real-time operation. It is important to keep in mind the purpose of the model since it influenced decisions of methods to be used and data fetching routines. Thus, key aspects of this study involves:

- fetching of real-time monitoring data;
- automatic calibration & validation;
- State of network in different nodes;
- Transfer times among pumping stations;
- 24h forecast: possible overflows, capacities, etc.

There is an existent hydraulic model of the study site. Therefore, details about this model are briefly presented later in section ???. Focus is given here to the hydrological model and the methodology was divided in two parts named: 1. Offline Model and 2. Online Model. The first aimed to investigate the model's data necessities, the process of parameter estimation based on spatial data of the study site, and first simulation with previously estimated parameters. The Online Model uses not only historical meteorological events, but also weather forecast for comparing the two proposed hydraulic models results and a preliminary analysis of their performance for real-time applications.

2.1 Offline Model

First step was the creation of hydrological model and coupling to hydraulic model using Stormwater Management Model (SWMM) developed by the U.S. Environmental Protection Agency (EPA). The model was built, calibrated and validated using historical data. The goal of the offline model part was to identify the best set of parameters and how to handle challenging parts such as: initial soil moisture content, soil frost, percentage of rainfall lost to stormwater sewer network, aquifer interactions, etc. Figure 2.1 shows a simplified flowchart of the hydrological model development steps.

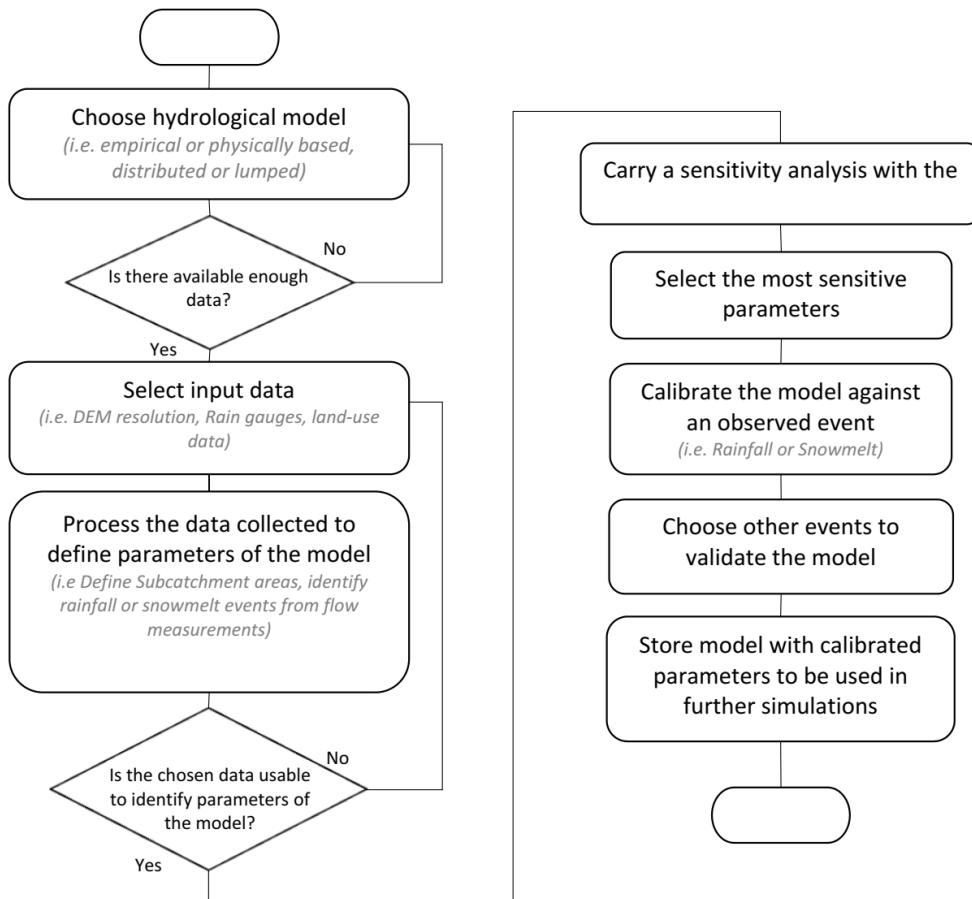


Figure 2.1: Hydrological Model Development Flowchart

2.2 Online Model

[To Do]

This section will aim to answer the questions posed below

This section presents questions related to the continuous simulation process. The answers express initial ideas to explain how the continuous simulation will work and what are the main concerns.

- How long will the simulations run? This depends on many factors such as the type of hydrological, hydraulic model, calibration process, routine, hardware used to process the information.

- How often will the simulation run? Routine of simulation is constraint by data acquisition routine. The routine of simulation can also vary during a storm event applying a shorter time of simulation. As an example, imagine if all necessary data to run a simulation is available every 15min and the model takes 30min to run a complete simulation, the iterations necessary for the optimization of parameters during the calibration process can be reduced as an attempt to decrease the time of simulation down

to 15min. This could allow a better decision-making process during intense rainfall events, even if some accuracy is compromised.

- What information is relevant before, during, and after a rainfall or snowmelt event? Before the rainfall and/or snowmelt: When and where peak flows will happen in the sanitary sewer system. Where and when can CSOs or SSOs can happen. What would be the transfer times between points of the network? During rainfall and/or snowmelt: Status of the system with focus in short-time forecast (+1-2h): Where are the peak flows, CSOs, SSOs, happening and how will then be within 1-2h. Are all the measurement instruments safe and providing good enough data? Maybe a quick check of the input data will be necessary to identify whether the instrument is broken or not. If we cannot rely on the real-time data, what would be the strategy? Maybe use the last reliable data collected to forecast the behavior of the system. After rainfall and/or snowmelt: Maybe the peak flows will still happen. How long will the flows in the system still be considered RDII and higher than average peak flows of DWF?

- Which parameters will be calibrated? According do Choi and Ball (2002) (Choi and Ball 2002) there are two types of parameters related to quantity part of runoff block in SWMM: 1. Measured Parameters; 2. Inferred Parameters. The first are parameters directly measured such length of channels/pipes, catchment land-use, or recorded rainfall depth. The second, parameters not directly measured, and coefficients used by empirical models that approximate complex physical processes. More about this is discussed in section 4.4. Inferred parameters approximate characteristics of the system (i.e. impermeability) and processes (i.e. flow coefficients such as hydraulic conductivity or manning's). Inferred parameters are the least known values since no measurements were carried. Therefore, they are the first candidates to be chosen when calibrating since the uncertainties tend to be higher than the measured parameters. A sensitivity analysis will be done during the offline modelling part to identify which of the inferred parameters have greater impact to the results. The inferred parameters with greater impacts will be then chosen for calibration.

- Why not all parameters will be calibrated? The number of parameters necessary to run SWMM and the range of possible values for each parameter can be a very large number. This increases the search space for the calibration algorithm where many iterations will be required to identify optimal set of parameters. This can rapidly increase the time of simulation limiting the model's forecast capability.

- How will the Automatic Calibration Process work? The automatic calibration will have a routine according to the necessity of the models. A calibration run might happen in scale of hours, days, weeks, months, etc. A new set of parameters will then be proposed and an evaluation to compare whether the new set of parameters is better than the previous one. The frequency of calibration and the possible values of the parameters must be further investigated since less frequent events might have very distinct behavior than more frequent events. It is also possible to define different set of parameters for different seasons and event's magnitude. An approximated system diagram is shown in figure 2.2.

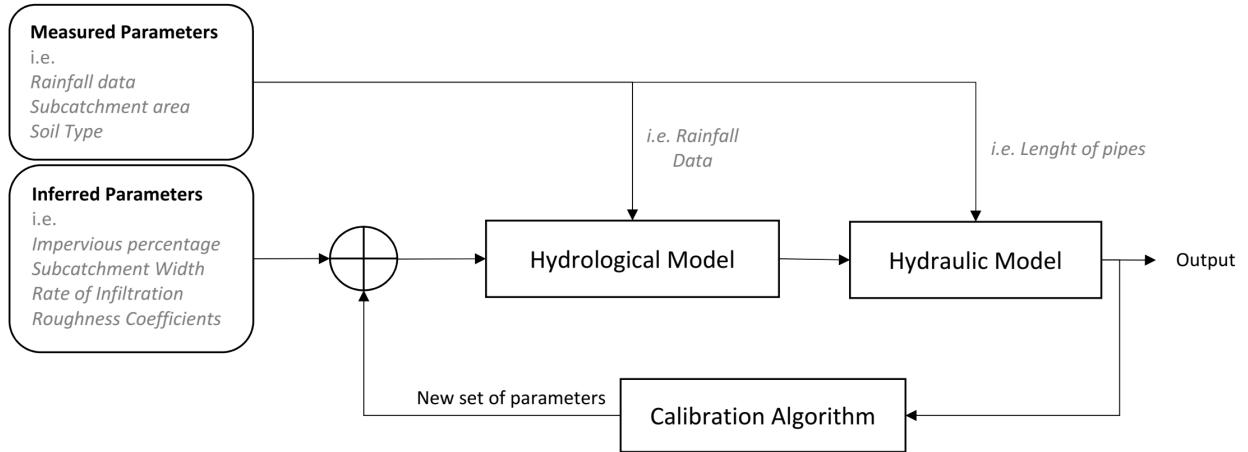


Figure 2.2: System Diagram

2.3 Parameter Optimization Algorithm

Dynamically Dimensioned Search (DDS) is a stochastic single-solution based heuristic global search algorithm that was created for the automatic calibration process of watershed simulation models. It was developed to find good global solution for a set of parameters faster than previously available search algorithms. DDS was proposed by Tolson and Shoemaker [2007] [30]. DDS mimics the process of manual calibration. According to Dent et al. [10] a calibration algorithm workflow can be defined as:

1. Change parameters of the model;
2. Run the model simulation;
3. Measure error of simulated vs observed;
4. Repeat previous steps and choose the most fit set of parameters.

For hydrologic models with multiple subcatchments and parameters, the number of possible combinations of parameters are of several orders of magnitude. This complexity motivates the adoption of optimization algorithms for automatic calibration. For a given hydrologic model with only one catchment and three different parameters: area, slope, and roughness. Assuming that area is constant but the other two can assume each five other possible values. The possible combination of parameters to represent the catchment would be 165. Now let's assume the catchments is divided in two subcatchments. In this case, the possible number of combinations would raise to 74613. The name Dynamically Dimensioned is given due to the ability of the algorithm to scale the search based on user-specified maximum number of iterations. Global search approach is used for the first iterations. DDS switches to a local search approach by selecting and reducing the search space when the number of iterations nears the maximum allowed. The algorithm reduces the search space by strategic reduction of the number of parameters to be calibrated when it approximates to the end of the search. It

also respects the constraints of each parameter given by the user. Therefore, it does not choose values for a parameter out of the specified range. Other relevant aspects of Automatic Calibration and DDS for SWMM:

- Automatic Calibration avoids modelers bias, accelerates the process of calibration, and handles multiple objectives such as peak flows, hydrograph shape, and total volume [10];
- DDS was created for computationally expensive calibrations [5]. Therefore, it is suitable for SWMM where a possible large number of parameters should be simultaneously optimized;
- DDS converges rapidly finding a good solution for a set of parameters and successfully avoid poor local solutions [30];
- For distributed model such as used in SWMM, comparisons available in the literature have proven that DDS is one of the fastest to converge and the best finding good solutions. In other words, DDS does outperform other algorithms for complex models[30, 34, 5];

Algorithm 2.1 Dynamically dimensioned search algorithm by Tolson and Shoemaker [30] and algorithm structure presentation by Sunela [26]

```

 $f_{best} \leftarrow f(\bar{x}_0)$ 
 $\bar{x}_{best} = \bar{x}_0$ 
for  $i \leftarrow 1, m$  do
    Randomly select the decision variables that will be perturbed.
     $p \leftarrow 1 - \frac{\ln i}{\ln m}$ 
     $N \leftarrow \emptyset$ 
    for  $d \leftarrow 1, n$  do
         $X \sim U([0, 1])$ 
        if  $X \leq p$  then  $N \leftarrow N \cup \{d\}$ 
    end for
    if  $N = \emptyset$  then ▷ Ensure variable change
         $X \sim U([1, n])$ 
         $N = \{X\}$ 
    end if
    Construct new solution by perturbing the current best
     $\bar{x} \leftarrow \bar{x}_{best}$ 
    for  $\forall j \in N$  do
         $x_j \leftarrow x_j^{best} + r \cdot (x_j^{max} - x_j^{min}) \cdot N([0, 1])$ 
        if  $x_j < x_j^{min}$  then
             $x_j \leftarrow x_j^{min} + (x_j^{min} - x_j)$ 
            if  $x_j > x_j^{max}$  then  $x_j \leftarrow x_j^{min}$ 
            else if  $x_j > x_j^{max}$  then
                 $x_j \leftarrow x_j^{max} - (x_j - x_j^{max})$ 
                if  $x_j < x_j^{min}$  then  $x_j \leftarrow x_j^{max}$ 
            end if
        end if
    end for
    Evaluate the objective function value for the new solution
     $f \leftarrow f(\bar{x})$ 
    if  $f \leq f_{best}$  then
         $f_{best} = f$ 
         $\bar{x}_{best} = \bar{x}$ 
    end if
end for

```

3 HYDROLOGICAL MODEL

Flow in the sanitary sewer network can be classified as Dry-Weather Flow (DWF) and Wet-Weather Flow (WWF). DWF can be further divided in two components: 1. Base Waste Flow (BWF): inflow of waste water coming from households, commercial and industrial sites; and 2. Groundwater Infiltration (GWI): Water from aquifers that infiltrates into the network thought defects such as pipe cracks and leaky joints ([32]). The choice of the hydrological model in this study aims the representation of RDII, which is the incremental flow into the sanitary sewer system caused by precipitation (rainfall or snowmelt). Figure 3 shows the typical characteristics of different components of sanitary sewer flow. RDII needs to be first separated from DWF when processing raw data coming from flow meters. More about the methods to separate the components are discussed on section 4.2.

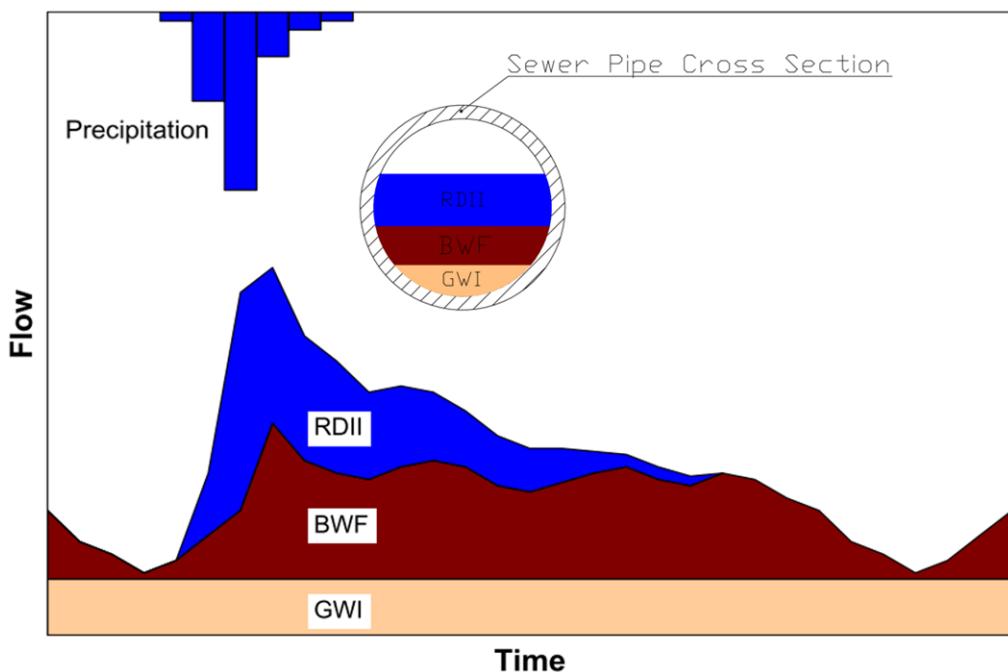


Figure 3.1: Wet-weather flow components. Modified from [32]

As mentioned on section 1. There are different ways stormwater or snowmelt finds its way into the sanitary sewer lines which ideally would have only wastewater from urban developments such as households, commercial centers, factories, etc. Sanitary sewer network flow increase can be triggered by an event such as a storm, snowmelt or increase of soil moisture content. From rainfall or snowmelt water flows over the soil surface and inflows to the sanitary sewer through manhole leaky covers or directly from roof-drain and foundation connections. This extra amount of wastewater inflow is

generally observed few hours after the beginning of the storm up to days after a period with intense snowmelt. [25]. The flow above the customary values observed as a long-term effect (days or weeks) can be explained by the features of subsurface flow. Once the water infiltrates, it moves through the soil porous with a much slower velocity due to the characteristics of the groundwater flow before entering the network system through its defects.

3.1 Methods to Quantify Rainfall Dependent Infiltration and Inflow

Rainfall dependent infiltration and inflow have been modeled with different methods. Bennet et al. [1999] [7] carried a literature review and case study of around 10 different methods for quantifying RDII. The study concluded that only the regression and unit hydrograph methods are suitable when applying continuous simulation for long-term modelling. However, a physically-based method was not assessed. The unit hydrograph (UH) method also provided the best consistent match to storm peaks among the benchmark methods. Vallabhaneni and Burguess [2007] [32] and U.S. EPA [2008] [12] also considered sewer network rehabilitation capabilities as a factor for evaluation of the methods and suggested that regression should be used when at least more than 2 years of recorded flow and rainfall data is available. When no flow is available, the Constant Unit Rate RDII Method seems to be useful since it accounts for spatial characteristics (topographical data) of the Sewershed, information of pipe characteristics and population. Moreover, U.S. EPA [2008] study concluded that Unit Hydrograph RTK method can be useful to identify which portion of the wet-weather flow is caused by direct inflow and which portion is caused by infiltration. Knowing whether RDII is more impacted by inflow or infiltration is relevant when evaluating the sanitary sewer network for rehabilitation. It is important to mention that the studies also concluded that there is no RDII quantification method that can be universally applied, since their use depend on available data and characteristics of the catchment. The goal of Vallabhaneni and Burguess [2007] and U.S. EPA [2008] reviews were to choose the most suitable method to be first implemented in a toolbox named as Sanitary Sewer Overflow Analysis and Planning (SSOAP) that is later discussed.

Examples of Empirical models such as multiple linear regression (Li et al. [18]), Artificial Neural Networks (ANN) (Djebbar and Kadota [11], Walker [33]) and RTK unit hydrographs (Muleta and Boulos [20], Gheith [14]) which is perhaps the most commonly used method, were published with different advantages and limitations.

Physically based models have also been used to model sanitary sewer flows. Robinson [23] modelled groundwater infiltration into SSN using SWMM5 and two aquifer approach to investigate rehabilitation method of urban catchment in Seattle, USA. [28] used SWMM5 rainfall-runoff module to study the impacts of climate change to combined sewer overflows (CSOs) in Toledo, Ohio, USA. [8] PhD dissertation thesis compared the use of three physically-based models: 1. roof downspout; 2. Sump pump, and 3. leaky lateral with RTK unit hydrograph method. Even though there are studies available in the literature using phisycally based models applied to sanitary sewer flows, it seems to be easier to find cases where process are analyzed separated and few include also snowmelt models

to the processes.

Therefore, a comparison of a physically based model and empirical model as hydrological inputs to the sanitary sewer hydraulic model was proposed in this study. SWMM 5 was chosen and the two available routines (SWMM physically based modules and RTK Unit Hydrographs) used to perform comparisons.

SWMM 5 was chosen for this study for three main reasons:

1. It includes snowpack and snowmelt model, infiltration, runoff and aquifer & groundwater flow, and RTK Unit Hydrograph;
2. Free open source package;
3. Well documented user manuals and several published case studies.
4. Existential hydraulic model.

The approach chosen focus on the flows in the Sanitary Sewer network (SSN) without further information of the Stormwater Sewer network (SWSN) or any Stormwater Harvesting System. These artificial units are treated as losses as depicted in figure 3.2.

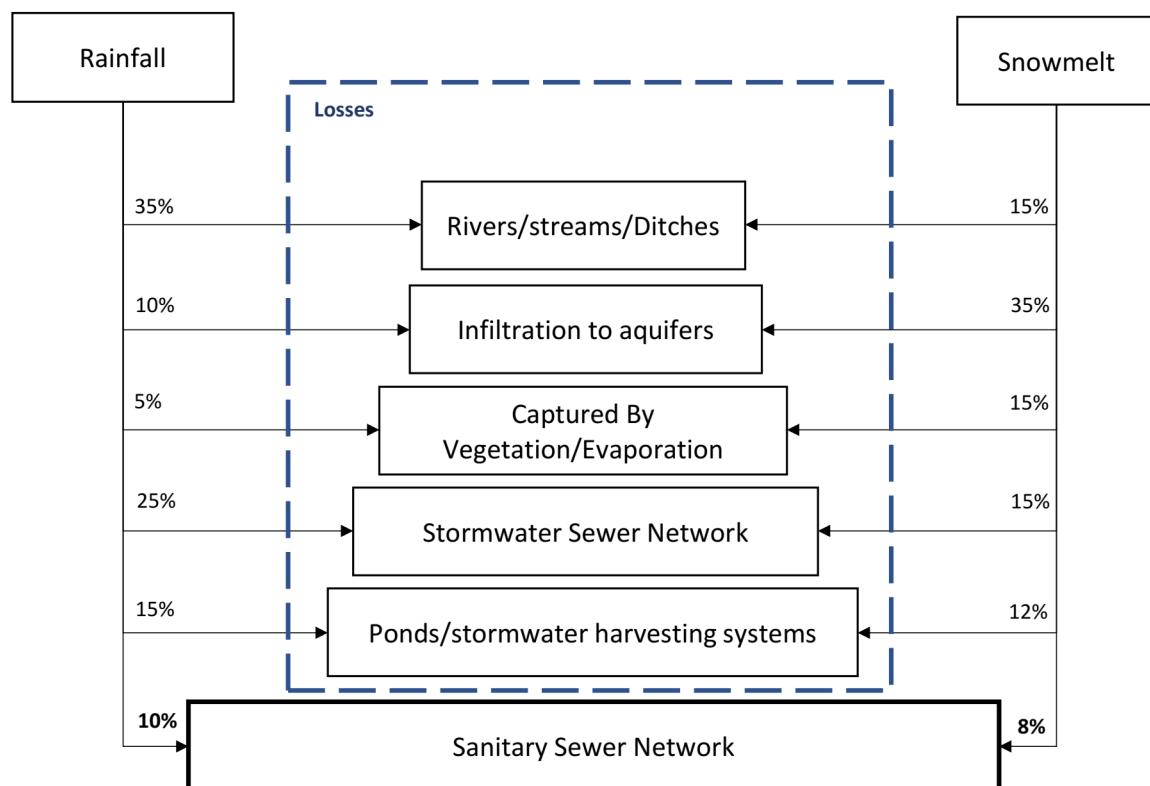


Figure 3.2: Precipitation Losses relative to a Sanitary Sewer Network

//

3.2 Physically-Based: SWMM Modules

The use of SWMM packages in this study aimed to model four processes happening simultaneously in the watershed to simulate fast, medium and long term response observed in SSN wet-weather flows. The four processes/SWMM modules are described here as: 1. Runoff; 2. Snowpack & Snowmelt; 3. Infiltration; 4. Groundwater. A summary of the four modules is presented on the following sections based on Rossman and Huber [25].

3.2.1 Rainfall-Runoff

The area in SWMM is discretized by subcatchments. The size of each subcatchment is dependent on the purpose of the model. See section ?? for further discussion on subcatchment delineation for this study. The rainfall-runoff is computed in SWMM for each one of the subcatchments using a nonlinear reservoir model as depicted in figure 3.3 and expressed in 3.1 [25].

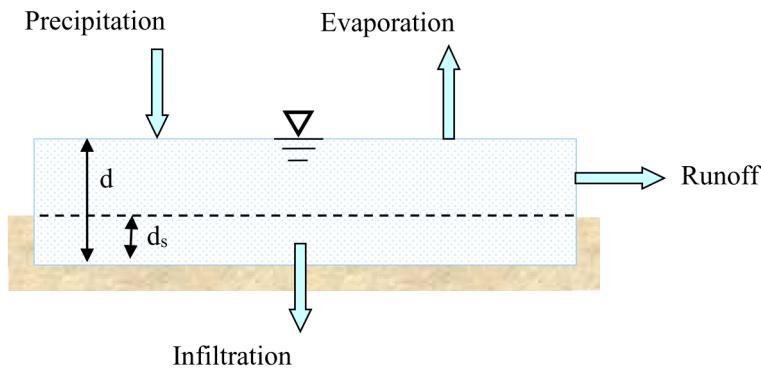


Figure 3.3: Nonlinear reservoir model [25]

$$\frac{\partial d}{\partial t} = i - e - f - q \quad (3.1)$$

Where:

i = precipitation rate [m/s]

e = surface evaporation rate [m/s]

f = infiltration rate [m/s]

q = runoff rate [m/s]

Runoff happens when water exceeds the depression storage (D_s) and the overland flow is assumed as uniform in a rectangular channel expressed by Gauckler-Manning-Strickler formula (3.2). Each subcatchment in SWMM can be divided in three portions: 1. Pervious; 2. Impervious; 3. Impervious without D_s .

$$q = \frac{1.49 \cdot W \cdot S^{1/2}}{A \cdot n} \cdot (d - d_s)^{5/3} \quad (3.2)$$

Where:

A = area [m^2]

W = Flow width [m]

S = Slope [1]

n = Manning's roughness coefficient [$\text{s}/\text{m}^{1/3}$]

Runoff can be divided and routed to three different areas: 1. Outlet (node within the pipe network); 2. Pervious or impervious portion of the subcatchment; 3. Other subcatchment. The modeller can input a percent of runoff routed ($\%_{routed}$) as a parameter for SWMM model.

No information of Stormwater Sewer network (SWSN) was assessed in this study. Therefore, the amount of stormwater that finds its way into the SWSN is treated as a loss (see figure 3.2). It was also assumed that the fast response on sanitary sewer wet-weather hydrograph follows the same pattern as surface runoff. However, with a reduction in its volume. One can imagine that the hydrograph coming from a subcatchment entering the SWSN will have the same shape as the "short term hydrograph" entering the Sanitary Sewer network (SSN), but with greater volume.

The volume of water from precipitation lost to the SWSN can be represented by two existent parameters in SWMM model: 1. depression storage (D_s); and/or 2. percent routed ($\%_{routed}$). Therefore, the values chosen for these two parameters in this study may differ greatly from other SWMM models focused more on modeling SWSN. Figure 3.4 represents the conceptual difference considered in this study in comparison with the original nonlinear reservoir model.

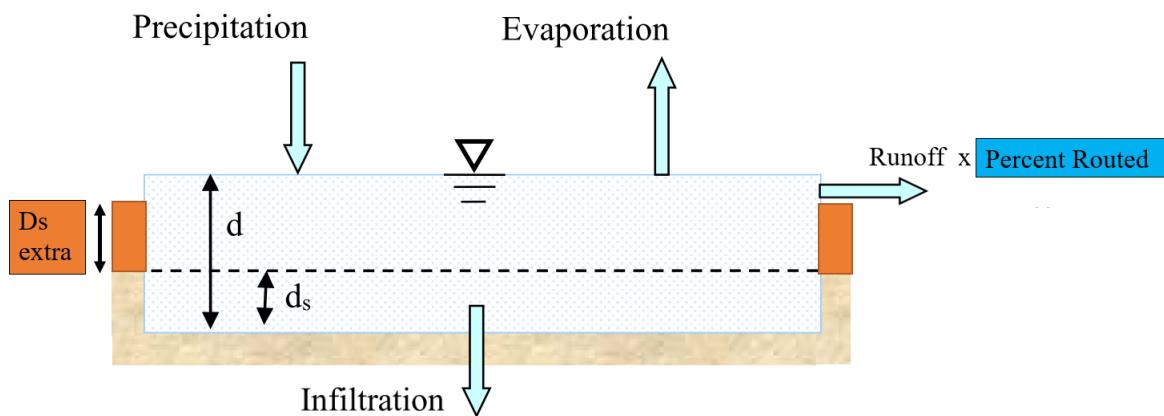


Figure 3.4: Extra D_s and $\%_{routed}$ for Nonlinear reservoir model. Modified from Rossman and Huber [25]

D_s accounts for the amount of water "absorbed" by the watershed from precipitation before runoff occurs. Wetting and ponding of the surface, and interception are usually the losses modeled by this parameter. The D_s Extra and $\%_{routed}$ represented in figure 3.4 are the increment value for D_s and $\%_{routed}$ to represent losses to SWSN system. The parameter estimation is discussed later in section 5.2.4.

3.2.2 Snowpack & Snowmelt

Snowpack & snowmelt module was used to simulate the variations of flows in the Sanitary Sewer network (SSN) occurring during winter conditions since a considerable incremental quantity of infiltration occurs during snowmelt periods as showed on the available data of section 5.2.1.

Snowpack & snowmelt calculation routines available in SWMM were based on models developed by U.S. National Weather Service (NWS) [4, 3]. SWMM models the depth of water equivalent as the snowpack. The depth is increased during snow accumulation periods and decreased when snowmelt occurs. The amount of water released from the snowpack during snowmelt is transformed in precipitation rate [mm/h] and summed to the rainfall as "net precipitation" that is used as input to compute surface runoff. Therefore, snowmelt calculations are part of runoff module [25].

Three of the key parameters for snowmelt routine are:

1. T_a : air temperature of the current time step [C°]
2. T_{base} : The base temperature of which snowmelt starts to occur [C°]
3. DHM : melt coefficient [mm/h/ C°]

These three parameters are used in the linear type equation 3.3 to compute the snowmelt [mm/h] during dry periods. Calculations of snowmelt during wet periods (greater than 0.51 mm/h) take also in consideration the wind speed and local atmospheric pressure. Refer to Anderson [4, 3] or Rossman and Huber [25] for detailed description of snowmelt calculations during rainfall.

$$SMELT = DHM \cdot (T_a - T_{base}) \quad (3.3)$$

Melt coefficient (DHM) varies seasonally and is calculated based on a sinusoidal equation and two user-supplied constants: 1. Minimum melt coefficient (DHM_{min}) which happens on December 21th; 2. Maximum melt (DHM_{max}) happening on June 21th as depicted in figure 3.5.

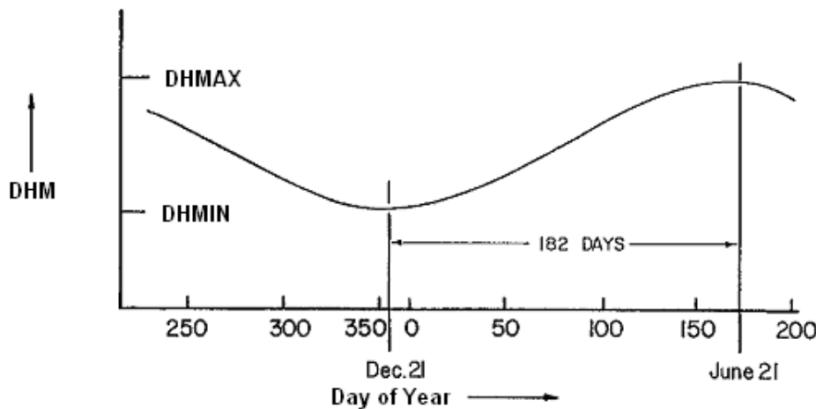


Figure 3.5: Seasonal variation of melt coefficients [25]

Before snowmelt occurs, the snowpack status has to be assessed. For this, there are two condition:

1. The snowpack has to be heated with air temperatures higher than T_{base} .
2. Snowmelt has to fill the voids within the snowpack. Meaning that there is a quantity of water contained in the snowpack and it is considered to be a fraction of the "depth of water equivalent" and named as fraction of free water capacity (FWC).

The heat content in the pack is calculated and FWC a user-supplied value. Therefore, liquid melt will only become a component of "net precipitation" after the two conditions, above mentioned, are satisfied.

The difference between heat content of the snowpack and T_{base} is named as "cold content" ($COLD C$). This variable is used to compute how much heat is necessary to be transferred to the snowpack before snowmelt can occur, as the first condition mentioned above. The $COLD C$ value is updated every time step based on the heat transfer between the pack and the atmosphere. The variation of the cold content (ΔCC) is calculated every time step assuming a negative value during melting periods. The following two user-supplied constant fractions are necessary to compute ΔCC :

1. RNM : Negative melt ratio [fraction]
2. $TIPM$: ATI weight ratio [fraction]

The rate of which heat transfer occurs is calculated based on SWMM's internal parameter of antecedent temperature index (ATI) which is function of T_a and $TIPM$. Values of $TIPM$ towards tending to zero represent a thicker pack which warms and cools slowly as a greater weight is given to more antecedent temperatures. Equation 3.4 is used when $T_a < T_{base}$ and equation 3.5 when $T_a > T_{base}$ [25].

$$\Delta CC = RNM \cdot DHM \cdot (ATI - T_a) \cdot TimeStep \quad (3.4)$$

$$\Delta CC = -SNOWMELT \cdot RNM \cdot TimeStep \quad (3.5)$$

The Negative Melt Ratio (RNM) in equations 3.4 and 3.5 is used to account for a reduced heat transfer during periods without "actual liquid melt". Snow plowing and areal depletion were not used in this study for lack of data and simplicity. Other three parameters used were:

1. U : Monthly average wind speed [m/s]
2. Z_{el} : elevation above mean sea level [m]
3. $SNOTMP$: Dividing temperature between snowfall and rainfall [C°]
4. SCF : Snow catch factor [ratio]

Where 1. and 2. were used to compute the influence of wind speed on the melting of snow during rainfall periods and 3. and 4. used to define the amount of snowfall from raw precipitation input data.

Table 3.1 depict all parameters used for the snowpack & snowmelt module in this study and their proposed range based on other study cases available in the literature.

Table 3.1: Snowpack & Snowmelt parameters range[25]

Parameter	Proposed Range	Units
SNOTMP	0 - 2	[°C]
SCF	1 - 2	[1]
T _{base}	-4 - 0	[°C]
DHM _{min - max}	0.019 - 0.11	[mm/°C-h]
RNM	0 - 1	[1]
FWFRAC	0.01 - 0.25	[1]
TIPM	0 - 1	[1]
T _a ; Z _{el} ; U	Location Based (see section 5.2.2)	

RNM and TIPM bare the full possible range. However, suggestions available in the literature were used as initial values in this study. All other ranges of parameters in 3.1, except by $DHM_{min-max}$, were proposed as suggested by Rossman and Huber [2016] in the SWMM hydrology reference manual [25].

Tikkanen [2013] suggested values for the degree-hour melt coefficients ($DHM_{min-max}$) when modelling a catchment in Finland based on values calibrated by Valeo and Ho [2004] for a catchment in Calgary, Canada. Tikkanen used reduced values in comparison to Valeo and Ho to account for fewer solar radiation due to difference in latitude. Valeo and Ho calibrated different values of $DHM_{min-max}$ for snow covered pervious and impervious areas varying from 0.02 for DHM_{min} to 0.150 DHM_{max} . Therefore, the proposed range in this study was based on Tikkanen and Valeo and Ho findings.

3.2.3 Infiltration

An infiltration model was used in this study to assess long-term simulations (up to 6 months) of the winter periods using snowmelt routine. As Groundwater Infiltration (GWI) is one of the components of Sanitary Sewer network (SSN) flows, an aquifer and groundwater inflow models were included. The gradient of groundwater infiltration to the SSN is dependent on the water table elevation (see section 3.2.4). Therefore, the infiltration routine was included as a way to recharge the modeled aquifer varying the saturated zone elevation (water table) providing a connection between effective precipitation and the GWI component.

SWMM version 5.1 package offers the modeller five different infiltration models. The Modified Horton method [1, 2] was chosen among the options for three main reasons: 1. it is simply one of the default methods available in SWMM; 2. it has the same parameters as the well-known Horton method which parameter estimates are suggested in Rossman and Huber [25]; 3. Appears to be more accurate low intensity rainfall events than the original Horton method [25].

The two governing equations of the method describes the infiltration capacity decay during wet

periods 3.6 and its recovery curve during dry periods 3.8 and an example of these two curves and how the infiltration capacity would change over time is plotted in figure 3.6.

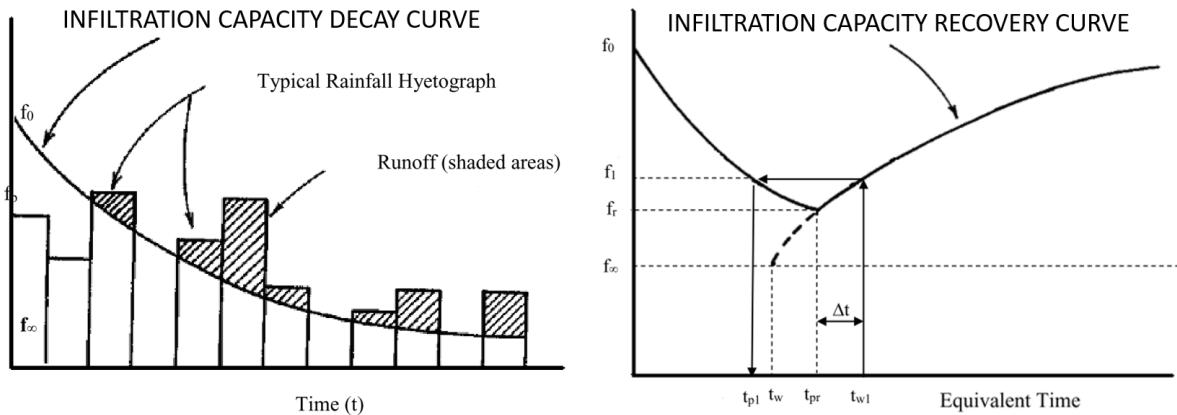


Figure 3.6: Horton infiltration capacity decay and recovery curves. Modified from [25]

$$F = f_\infty \cdot t + \frac{(f_0 - f_\infty)}{k_d} \cdot (1 - e^{-k_d \cdot t}) \quad (3.6)$$

Where:

F = cumulative infiltration capacity [ft]

f_∞ = minimum or equilibrium value of infiltration capacity at $t = \infty$ [ft/sec]

f_0 = maximum or initial value of infiltration capacity at $t = 0$ [ft/sec]

t = equivalent time [sec]

k_d = decay coefficient [sec^{-1}]

It is worth to mention that 3.6 is an integrated form of Horton's original equation. SWMM uses integrated form to consider the intensity of the rainfall event also as a function of the infiltration capacity reduction [25].

$$\frac{df_r}{dt} = kr \cdot (f_{r0} - f_r) \quad (3.7)$$

Where:

f_r = infiltration capacity during recovery [ft]

f_{r0} = maximum or initial value of infiltration capacity at $t = 0$ [ft/sec]

k_r = regeneration coefficient [1/sec]

t = time [sec]

the infiltration capacity at time t after integrating 3.7 when infiltration capacity is f_{r0} is:

$$f_r = f_0 - (f_0 - f_{r0}) \cdot e^{-k_d \cdot t} \quad (3.8)$$

SWMM computation scheme first checks for wet-period (rainfall/snowmelt) or dry period to apply either of the equations 3.6 or 3.8 and compute the current infiltration capacity and the amount of water infiltrating the soil. More details of the equations and computational scheme are available in Rossman and Huber [25].

Table 3.2 presents a rough estimate of the range of four input parameters for Horton infiltration model. The range was extracted from EPA SWMM user help.

Table 3.2: Modified Horton infiltration parameters range[25]

Parameter	Typical Range	Units
Maximum infiltration rate	8.5 - 254	mm/h
Minimum infiltration rate	0.254 - 120.4	mm/h
Decay coefficient	2 - 7	h ⁻¹
Drying Time	2 - 14	days

3.2.4 Aquifer & Groundwater Flow

There are medium and long term wet-weather infiltration observed in Sanitary Sewer network (SSN) measured flow data. This infiltration raises the flow above its average for days or even weeks as discussed in the previous sections. It is challenging or not possible to model the short, medium and long term hydrographs using only SWMM's runoff module since its parameters such as roughness and slope would be distorted as an attempt to reproduce the delayed flows. Moreover, a short term response can occur at the same period as the medium and long term. One can imagine a high intensity rainfall happening right after a snowmelt period. A delayed portion of snowmelt infiltrates and slowly recharges the aquifer and discharges into the SSN while rainfall causes runoff being fully discharged in minutes or hours. The Aquifer & Groundwater Flow module of SWMM was implemented together with the aforementioned modules as an attempt to represent the most important hydrological processes happening in the sewershed. Aquifer in SWMM is represented as a two zones model containing an unsaturated and a saturated zone as depicted in figure 3.7.

The first zone is intermediary located between the soil surface and the groundwater table. Six fluxes among soil surface and the two aquifer zones can be computed every time step varying the elevation of the groundwater table and, therefore, changing the size of the aquifer zones. The variation of the saturated zone elevation (d_L) affects the flow from the aquifer to the receiving node in the pipe network system (f_G).

Modellers can use a customized equation to describe f_G flux using available parameters from the two-zone model. SWMM's hydrology reference manual [25] mentions options such as Linear Reservoir, Dupuit-Forcheimer Lateral Seepage, and Hooghoudt's Tile Drainage. However, it is important

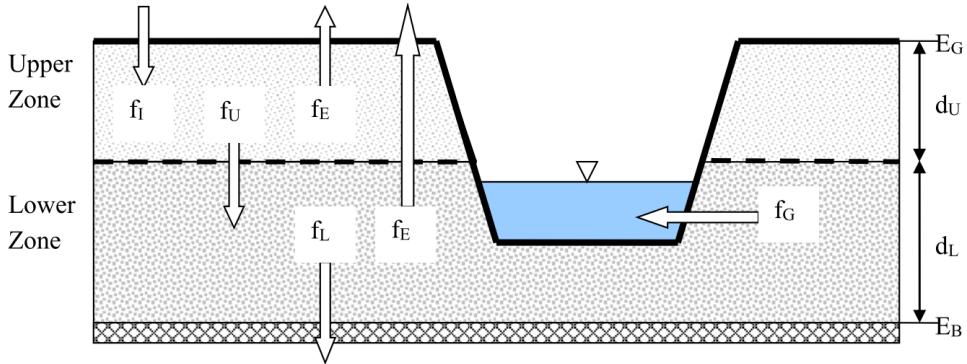


Figure 3.7: Representation of two-zone groundwater model in SWMM [25]

to remember that groundwater table elevation is constant along the subcatchment limiting the representation of the pressure gradient between the saturated zone and the receiving node to the difference between d_L and the elevation of water surface in the receiving node. Note that these two elevations can vary every time step and govern changes in f_G .

3.3 Synthetic Unit Hydrograph: RTK

Flows are higher than average into the sanitary or combined sewer network during and after a storm. This incremental quantity of stormwater inflows into the network from roof drain connections, foundation connections, leaky manhole covers, etc. [25] The inflow causes a relatively high peak discharges in the network in a short term, usually while the storm is happening. A bit less intuitively, flows and depths remain higher than average in the network after the rainfall event ranging from hours to even weeks [19]. This long term is caused by Infiltration of stormwater that enters the network system through defects such as damage pipes and joints [25]. Stormwater infiltrates the network after percolating through the soil. Groundwater also infiltrates into the network due to water table elevation caused by wet periods.

To simulate the inflow & infiltration caused by a rainfall event, SWMM incorporates the synthetic unit hydrograph RTK method. This unit hydrograph was first created to simulate RDII, therefore it accounts for the short, medium- and long-term effects. This model creates three triangular unit hydrographs, one for each term. Unit Hydrographs are then summed to create a final unit hydrograph that simulates the overall response of the system as shown in figure 3.8. The letters R-T-K refers to the parameters used to create the triangular hydrographs and are, respectively, fraction of precipitation that enters the network (area below hydrograph), time to peak, and recession time.

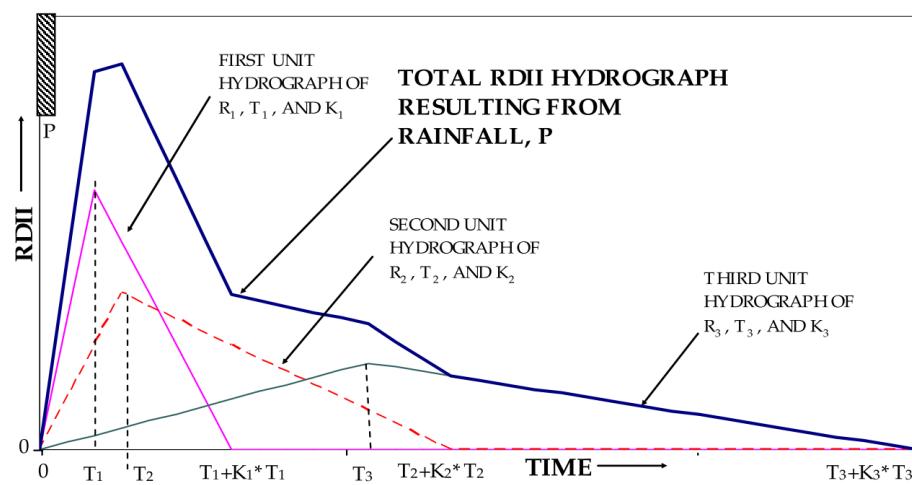


Figure 3.8: RTK Short, Medium, Long-term and Resulting Hydrographs [32]

4 HYDRAULIC MODEL

[TO DO]

- include relevant characteristics of the network (length, diameter, location, average elevation, network components, etc)
 - include map of network
 - characteristics of demand estimation
 - literature background of flow routing method: governing equations
 - Saint venant eq.
 - Darcy-Weisbach headloss formula
 - surcharge model
 - numerical scheme: dynamic wave and why its choice.

5 CASE STUDY

The town of Jokela in Finland was chosen as the study area. Pipe Network and flow meter data was provided by Tuusula Water Utility *Tuusulan Vesihuolto*. The following sections provide an summary of characteristics of the catchment, data providers and data treatment, and parameter estimation for physically based and unit hydrograph methods.

5.1 Jokela Town

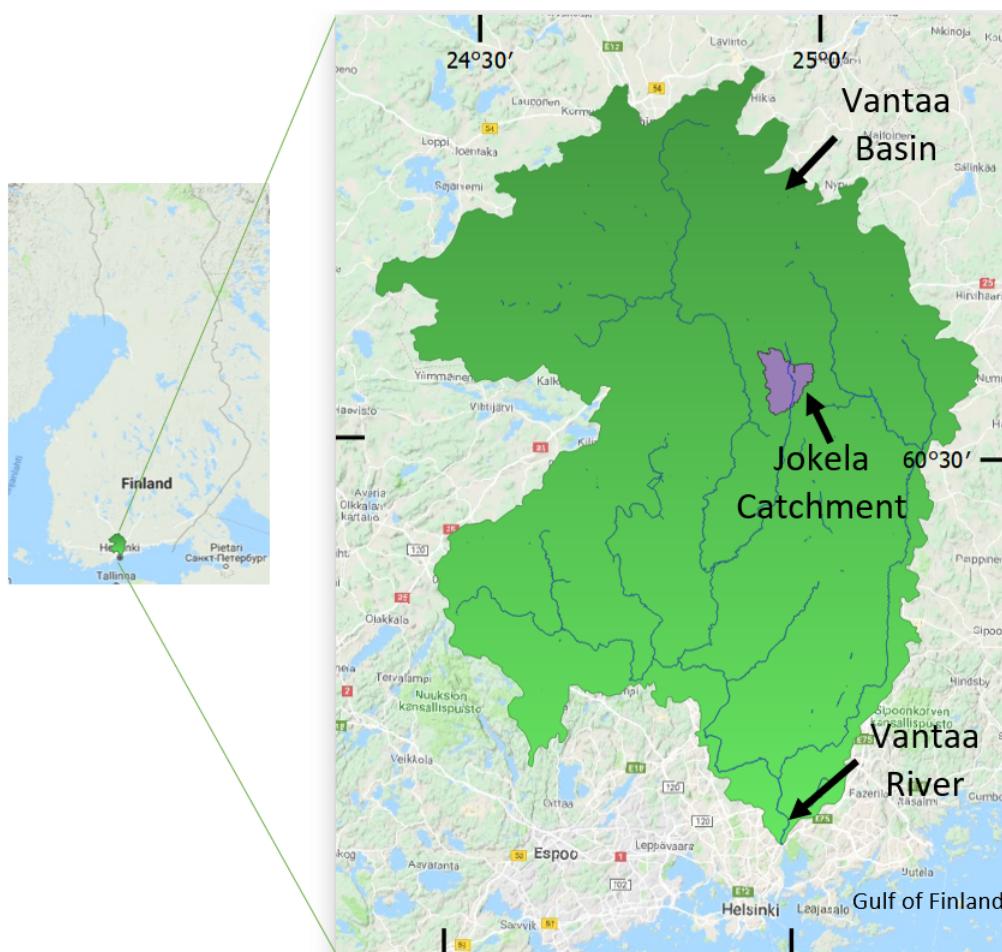


Figure 5.1: Vantaa Basin

Jokela is a town located in south region of Finland approximated 40 [km] from the capital city of Helsinki. It is part of Tuusula municipality. Jokela's urban area is mostly residential with an area

estimated of 315 [ha]. There are approximated 6500 residents and 2975 buildings (households, commercial, and factories). The area is rather flat with an average slope of 3%. Its land use can be considered semi-urban and roughly divided as approximated 43% of its catchment area with forests and semi-natural areas, 32% with urban artificial surfaces, 13% of transitional woodland or shrub and 12% or pastures. Its soil superficial deposit is mostly clay (66%) and sand (20%). Its main drainage stream has an rough average of 6.5 [m] width and cuts the catchment flowing from north to south draining to Vataanjoki (Vantaa River) and traveling approximated 50km before reaching the Gulf of Finland.

5.2 Data and Parameter Estimation

This section presents the details of parameter estimation for the physically-based model (SWMM modules) and RTK-Unit Hydrographs and the data sets used for model construction, calibration and validation.

5.2.1 Sanitary Sewer Flow Data

Flow meter data was provided by Tuusula Water Utility which is responsible for urban drainage system management of Jokela town and all Tuusula municipality. Data was measured at SSN pumping stations of which two are used for calibration and validation purposes in this study. One of the meters is located at the last pumping station (Jokela Pumping Station - subcatchment D4S4) of the delineated area and it is considered the system outlet. hourly measurements were recorded during almost whole period of 2018. The flow data was pre-processed to eliminate missing data, outliers and normalize it to reduce noisy measurements as an attempt to capture the overall evolution of the flow. Figure x depicts the distribution of measured data before and after the elimination of outliers and missing data. The outliers, flows greater than 200 [l/s], appears as spikes measured only at one data interval unit (1h). This suggests that these measurements are in fact unreal values and were filtered out. Missing measurements were completed using simple linear interpolation. Zero flow measurements were assumed also as missing data. When longer periods of missing data was observed (approximated >12h) were left out of the future analysis: RDII quantity estimation and calibration.

Outliers presenting extreme flow were relatively easy to be identified and filtered by defining a threshold and filtering out. However, unreal measurements presenting customary flow rates were harder to be filtered using the threshold method. An example of such measurement is depicted in figure 5.3 on February 22nd at 20:00. The flow data was then smoothed using Savitzky-Golay filter with window length of 11 and 5th order polynomial. It is important to mention that care should be taken when smoothing the data since it can erroneously reduce the peak flow measured. Periods with flows higher than average and peaks are most important when simulating wet-weather flow conditions into SSN. Higher order polynomials better reproduce the peak flows, but also keep unrealistic low flows. The pre-treatment procedure is most used for auto-regressive models that use historical flow data not only for comparing the results and calibrating, but also to build the model itself. Li et al. [18] performed similar pre-treatment on a sanitary sewer flow data derived from pump ON/OFF operation to

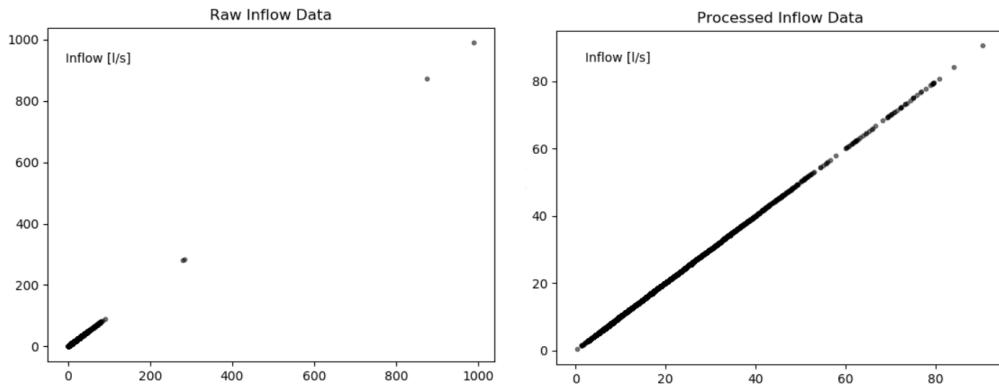


Figure 5.2: RDII x DWF estimate for 2018 of total waste water load

transform non-stationary data into stationary before applying autoregressive-moving-average model (ARMAX).

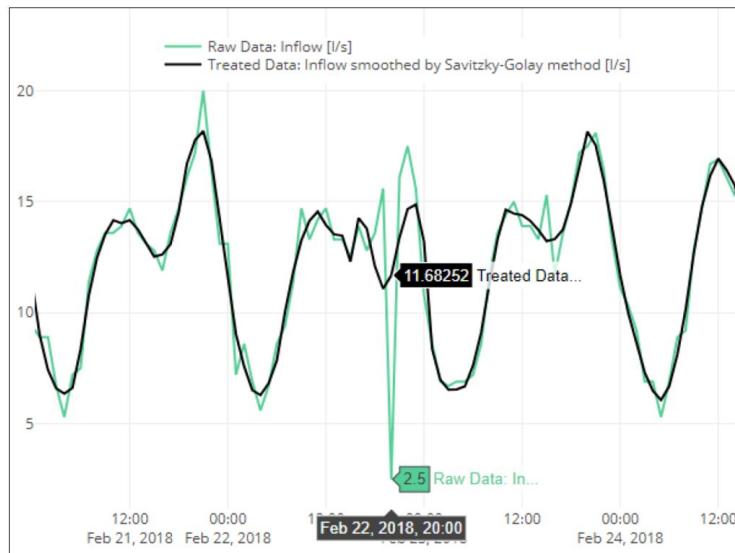


Figure 5.3: RDII x DWF estimate for 2018 of total wastewater load

An external script written in python 3 was developed to perform the flow data treatment. Other relevant tool used was the EPA SSOAP toolbox [32]. It helped to identify RDII flows from the raw flow data as described in the following section.

Separation of Sanitary Sewer Flow Components

It was relevant to identify the different components of the Sanitary Sewer network (SSN) to estimate the amount of RDII flows and compare measured with modelled when analysing the water balance for short or long-term simulations. The estimated "measured" RDII flows were identified using EPA SSOAP toolbox.

EPA SSOAP toolbox has many functionaries. As its core, it allows the users to visually build RTK Unit Hydrographs through an iterative process by adjusting RTK parameters and comparing the output hydrograph with the hydrograph measured [32]. However, the SSOAP toolbox was mostly used in this study to identify precipitation events and identify the flow components in Jokela's SSN. RTK parameters were estimated in this study using an optimization algorithm as described in previous sections.

Two input data were used in SSOAP: 1. Sanitary Sewer Flow Data of Jokela Pumping Station. 2. Precipitation data (better described on section 5.2.2). SSOAP utilizes precipitation data to identify wet-periods and define precipitation events. Days without Wet-Weather Flow (WWF) are then filtered to define a mean Dry-Weather Flow (DWF) hydrograph for weekdays and weekends. Current day and previous two days with any precipitation record were filtered out of the data set for Weekdays and Weekends. Past three to seven days with precipitation amount record higher than 10 [mm] and 15 holidays in Finland were also filtered out of the set. A total of 66 days weekdays and 32 weekends remained from DWF identification process. The output DWF hydrographs are depicted in figure 5.4. Weekdays average daily DWF of 11.4 [l/s] with standard deviation of 1.75 and 12.5 [l/s] and standard deviation of 2.35 for weekends. Refer to Vallabhaneni and Burgess [32] for detailed information about the SSOAP toolbox.

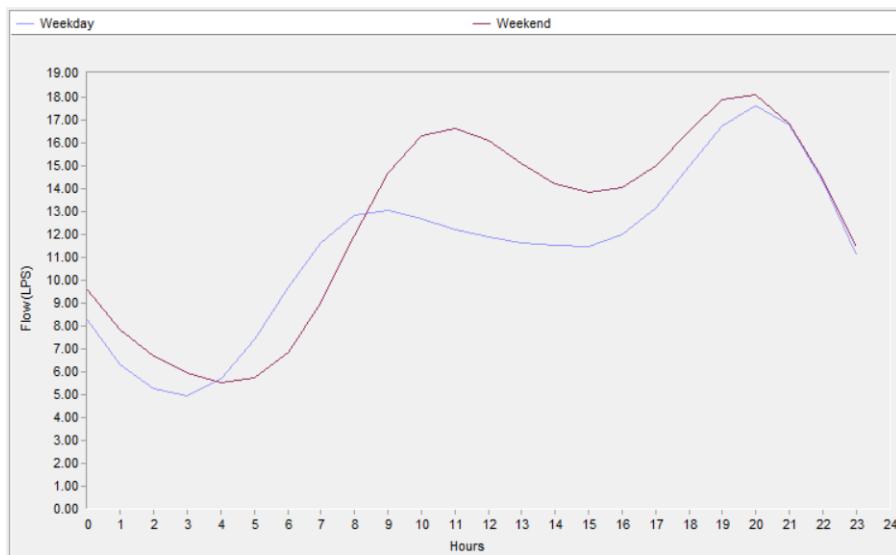


Figure 5.4: Estimated Dry-weather Hydrograph from EPA SSOAP tool

The next steps using SSOAP tool were towards subtracting the DWF estimated from metered flow to quantify RDII flow ($RDII = MeteredFlow - DWF$). The first step is to define the constant groundwater infiltration (GWI). This is important to further divide the DWF into two parts: Basewaste flow (BWF) and GWI. The GWI is estimated as a percentage of the minimum nighttime flow. SSOAP help file suggest that in residential areas typically 90% of the minimum nighttime flow represents GWI. However, a different approach was chosen based on available information from the water utility.

It was known that an estimate of 900[m³/day] is pumped to Jokela area from the water distribution system. Then, It was assumed that almost all water supplied eventually finds its way to the sanitary sewer network, meaning that the daily average BWF is roughly 900[m³/day]. To achieve this value for BWF with 2018 recorded flow an estimate of 25% of daily minimum nighttime flows was attributed to GWI with the rest (75%) representing wastewater flow. Table 5.1 depicts the estimated daily average flow components of Jokela's SSN with approximated 37% of the total wastewater amount transported to the Wastewater Treatment Plant (WWTP) being non revenue water (NRW) from RDII and constant GWI. Figure 5.5 has the monthly amount of the amount of RDII versus DWF with approximated 87% of total annual RDII amount happening from January to May of 2018.

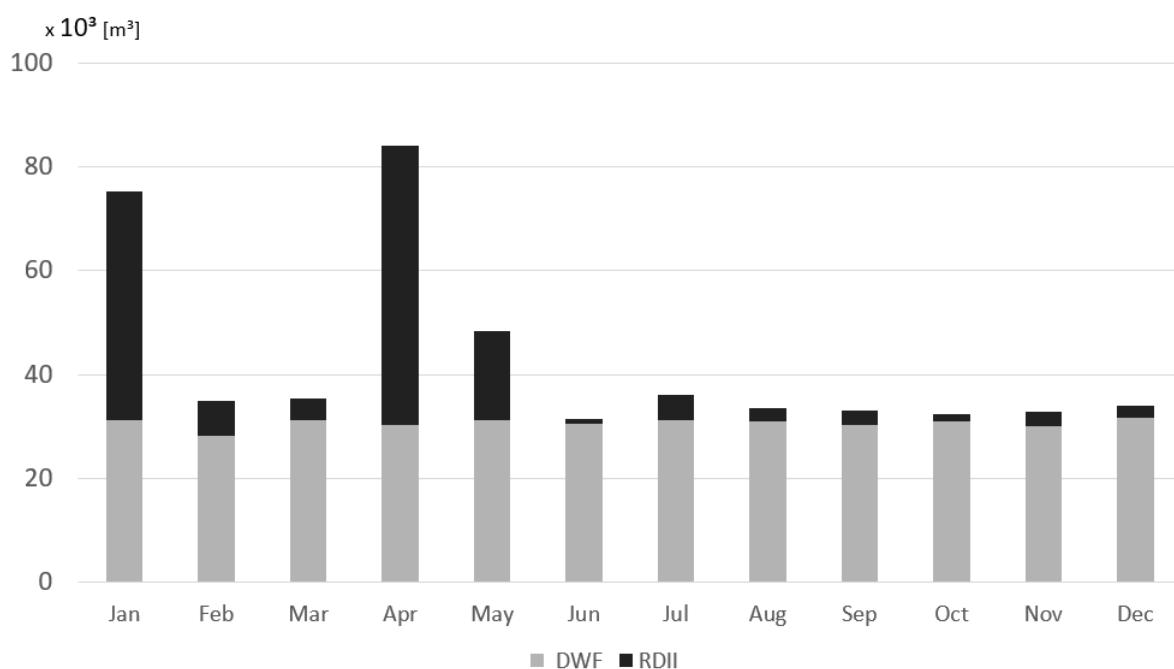


Figure 5.5: RDII x DWF estimate for 2018 of total wastewater load of Jokela SSN

It is import to remember that the RDII flow can also be further divided (i.e. short term hydrograph from runoff and long term hydrograph from subsurface infiltration) and the varying groundwater infiltration happening after rainfall or snowmelt is considered here different than the constant GWI used in SSOAP toolbox. Different GWI in SSN were discussed in section 3.1.

5.2.2 Meteorological Data & Parameters

Meteorological data is presented for both historical and weather forecast. The first is used in this study as input for hydraulic models during calibration and validation process whereas the forecast is utilized to compare the model performances using HARMONIE model forecast. Precipitation data is used for both physically-based and RTK unit hydrograph method. Snow depth measurements were used as input for the snowpack & snowmelt SWMM module process and to calibrate its parameters.

Table 5.1: Monthly estimated amount of Sanitary Sewer network (SSN) flow components in 2018 for Jokela Catchment in [m³]

Month	Obs Flow	DWF	BWF	RDII	RDII + GWI	% RDII	% RDII + GWI
Jan	2429	1007	897	1423	1533	58.57	63.09
Feb	1241	1006	896	237	347	19.13	28.00
Mar	1136	1010	899	136	246	11.93	21.68
Apr	2801	1011	901	1790	1900	63.90	67.85
May	1556	1010	900	547	657	35.13	42.22
Jun	929	1015	903	32	143	3.42	15.40
Jul	1041	1007	897	159	269	15.33	25.89
Aug	932	1003	894	82	192	8.80	20.58
Sep	1028	1011	901	94	204	9.10	19.85
Oct	986	1003	894	39	149	3.96	15.07
Nov	1066	1004	894	89	199	8.31	18.63
Dec	1025	1022	910	77	189	7.51	18.45
Total	16169	12110	10785	4704	6028	29.1	37.3

Precipitation

[TO DO]

- Relevant characteristics of precipitation data (monthly amount, intense precipitations, snowfall and rainfall, etc)
- maybe further comparison of 2018 precip. data and past 10 years.
- possible data treatment (i.e. missing values)
- rain gauge used
- how precip. data was input to the models
- Events chosen for calibration.

Temperature

[TO DO]

- Relevant characteristics of Temp. data
- maybe further comparison of 2018 Temp. data and past 10 years.
- possible data treatment (i.e. missing values)
- how temp. data was input to the models

Note: One missing data that was filled using linear interpolation between the previous and next

hour.

Wind

[TO DO]

- Relevant characteristics of wind data (average speed and direction)
- maybe further comparison of wind data from different weather stations
- possible data treatment (i.e. missing values)
- how wind data was input to the models

Snow Depth & Snowpack Parameters

[TO DO]

- brief comments on the snowdepth data (measure method, location of the measurements and the limitation when extrapolating for the entire catchment area

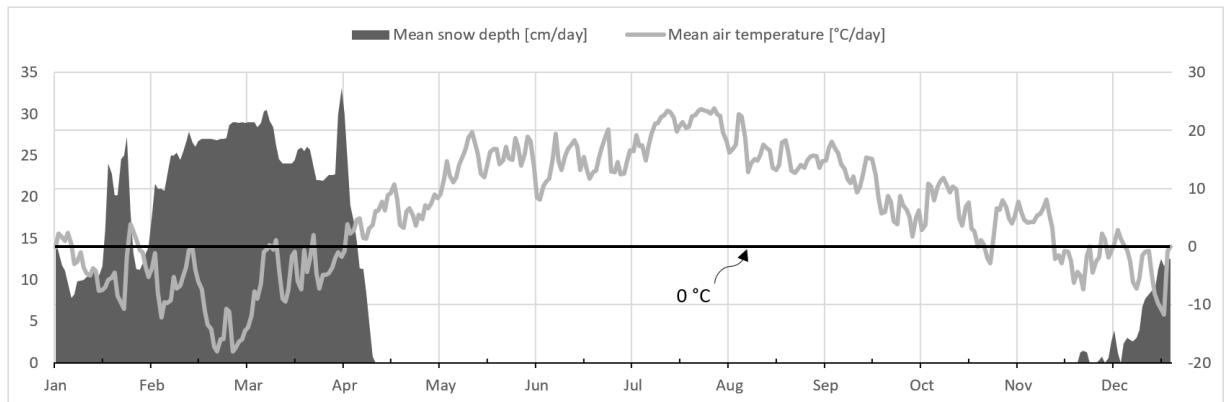


Figure 5.6: Snow depth and Temperature Measurements FMI [13]

Estimation of parameters was carried based on the range proposed in section 3.2.2. Sensitivity analysis was carried by manually varying individual parameters while others were static. It was found that the degree-hour melt coefficients (DHM_{min} , DHM_{max}), dividing temperature ($SNOTMP$), base temperature (T_{base}) and snow catch factor (SCF).

Base temperature and dividing temperature were initially estimated by observing relations among hourly temperature and snow depth. Whenever the snow depth measurement varied, values of temperature were collected. Temperatures during the increase of the snow depth were stored to estimate

the dividing temperature(*SNOTMP*) whereas temperatures during decrease of snow depth were used to estimate the base temperature(*T_{base}*). Periods when the temperatures were out of the range limited (see table 3.1) were discarded. The mean of the stored temperature were then used as the final estimation of *SNOTMP* and *T_{base}* and were respectively 0.1°C and -1.9°C. It is important to emphasize that this was a rough estimate used only to input initial values for the simulations. *DHM_{min}* and *DHM_{max}* were set initially as the limits of table 3.1. *TIPM* and *RNM* were set as used by Rossman and Huber [25] and Anderson [4]. No information on the rain gauge deficiency to record snowfall and fraction of free water capacity values were assessed. Therefore, *SCF* and *FWFRAC* values were set initially to their middle range. Two parameters for initial condition of the simulation are also required: 1. Initial depth of water equivalent (*SD₀*); 2. initial free water (*FW₀*). The first was estimated as 10% of the first measured value of snow depth assuming a ratio of 10:1 for the snow pack depth and water equivalent depth as rule of thumb suggested by Rossman and Huber [25].

Figure 5.7 show the results of the simulation using the initially estimated parameters and the results of a simulation with parameters manually calibrated. Four first months of 2018 between 1st of January and 15th of April when all the snow depth measured was already melted.

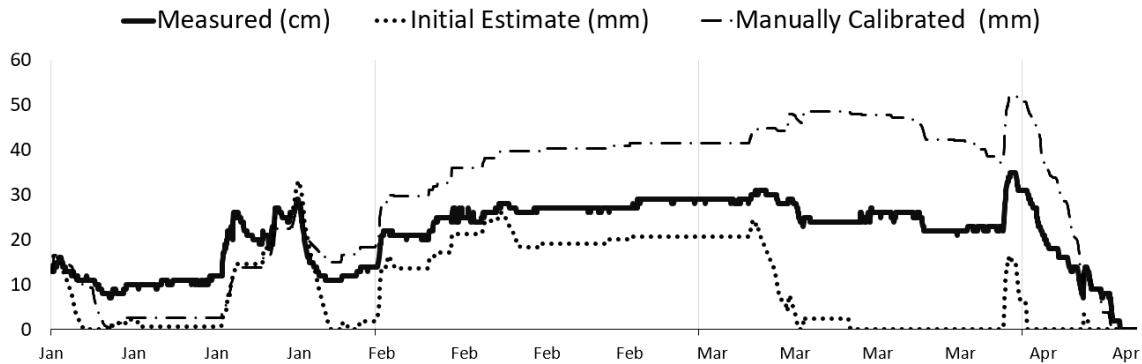


Figure 5.7: Results of Initial parameter estimation x manual calibration of Snowpack & Snowmelt parameters

By a visual analysis of the results of the first simulation and comparison to the measured snow depth values, it is possible to identify a higher rate of melt and snow accumulation than the values measured. This was corrected after a manual calibration of parameters. A lower rate of snow accumulation was achieved by assuming that the rain gauge was able better measure the snowfall. Therefore, *SCF* was reduced to 1.2. To reduce melting rate, the temperature which snow melt starts (*T_{base}*) was increased to its upper limit of 0°C. Although the results were better, a higher melt rate than expected was still with the pack being completely melted around three weeks before the expected time. Therefore, the minimum melt coefficient (*DHM_{min}*) was lowered further than its lower range limit from 0.019[mm/°C-h] to 0.001[mm/°C-h].

Table 5.2 depicts the initially estimated parameters, parameter values after manual calibration against measured data, and changes in the previously proposed range based on literature review (see

table 3.1) after results of the manual calibration. The only changed was a reduction of DHM_{min} parameter of approximated 47%.

Table 5.2: Snowpack & Snowmelt estimated parameters [25]

Parameter	Initially Estimated	Manually Calibrated	Changes in Previous Range	Calibrated
SNOTMP	0.1	0.1	-	[°C]
SCF	1.5	1.2	-	[1]
T _b	- 1.9	0	-	[°C]
DHM _{min} - DHM _{max}	0.019 - 0.10	0.009-0.03	0.009 - 0.15	[mm/°C-h]
RNM	0.6	0.6	-	[1]
FWFRAC	0.13	0.13	-	[1]
TIPM	0.5	0.5	-	[1]
SD ₀	13	13	-	[mm]
FW ₀	0	0	-	[mm]

5.2.3 Weather Forecast

[To Do]

- Description of Harmonie weather forecast data provided from FMI
- grid resolution of 2.5km, Weather forecast variables and rainfall up to h+ 66h, forecast Updated every 6 hours, lambert's projection
- Description of the routine created for automatic data acquisition from FMI APIs



Figure 5.8: HARMONIE weather model forecast coverage

5.2.4 Topographic Data & Parameters

Topographic data was assessed to estimate parameters used to model different flows happening in Jokela's catchment. The chart presented in figure 5.9 shows the relation among topographic data sets and the process being modeled.

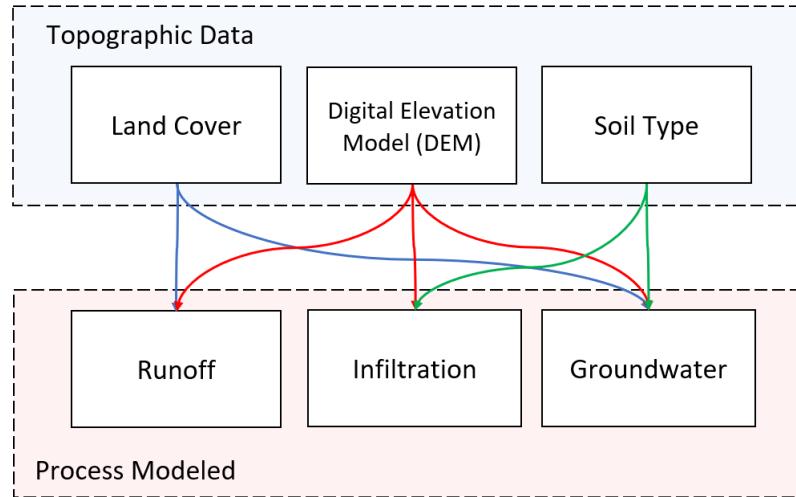


Figure 5.9: Topographic data used per each Modeled Process

Data description and initial parameter estimates are provided in the following sections.

Terrain Data & Runoff Parameters

The Digital Elevation Model (DEM) was used in many parts of this study. It is used in the Runoff model to delineate the subcatchments that provided area information for all the SWMM modules used. The DEM used in this study was provided by the National Land Survey of Finland (NLS) through its open data platform as 2 x 2 [m] resolution based on laser scanning data collected in the summer of 2015. The vertical resolution varies from 0.3 to 1 meter [21]. The DEM provided had already bridges cut off and a "filling" operation was carried using Qgis to exclude possible holes existent in the data set.

The definition of the *Area* parameter for Sanitary Sewer network (SSN) is rather conceptually challenging since the pipe network is located under the surface. As mentioned previously, surface and groundwater flow influence SSN and their delimited area (watershed or aquifer) as distinct. This can be imagined with the classical example of a rain drop falling over a specific point on the soil surface. The drop can flow over the surface towards the direction of the steeper slope or infiltrate and then flow through the porous in the soil to very different direction than it would flow if remained on the surface. Boundaries of a catch basin can be defined using the highest elevations such as topographic crest and road center line or lowest points of the area such as rivers and streams [17]. Cadastral parcels, which are property developed urban areas, are also often used to delineate the sewer catch basin. However, parcels are rather divided for administrative reasons than soil or subsoil characteristics. Thus, it can be

challenging to estimate some topographical parameters such as soil surface slope and roughness using administrative areas. As a hypothetical exercise one can imagine the possible issue when using parcels when modelling a urban area surrounded by mountains. In case the delineated area is limited only to the properties, no information relative to the mountains' is included in the parameter estimation and simulated hydrographs can mismatch the observed by volume and shape.

The sewershed delineation is, therefore, a discretization of the space domain. In this study, it is assumed as the area division that supplies water to a specific point within the SSN. The DEM is usually the data assessed to delineate the area of influence when modeling natural rivers. This method, however, is not conceptually valid for SSN once the network's slopes are not always follows the soil surface slopes estimated with the DEM. Therefore, SSN has a different slope gradient than the soil surface in some parts and pumping stations are required to transport waste water towards the Wastewater Treatment Plant (WWTP). This is one of the reason why the space domain was discretized per pumping station in this study. The second reason is that flow meter observations was available for two pumping stations within the catchment and the remaining stations are probably the best candidates to receive flow meter devices in the future.

DEM, pumping station locations and pipe network were the data used for the subcatchments delineation. Qgis application was chosen to perform GIS operations and delineate sewersheds due to its free open services and embedded automatic delineation tools such as *GRASS – r.watershed*. Proposed subcatchment partition for Jokela town is depicted in figure 5.10

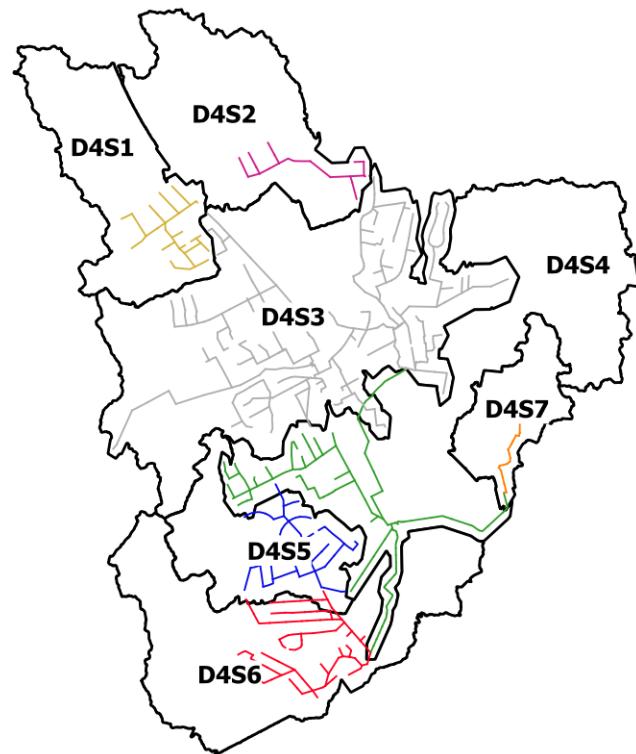


Figure 5.10: Jokela Subcatchment division

Pumping stations are usually located in areas with lower surface elevation when compared to its upstream pipe network to direct the flow utilizing gravitational force and reduce energy consumption. However, differences between soil surface and network still exists in some parts of the pumping station upstream service area. This happens because no information about soil type and infiltration rates is used when delineating the areas. Therefore, subcatchment of pumping station one (PS 1) may overlay the pipe network upstream pumping station (PS 2). Only automatic sewershed delineation procedure and DEM is not enough to avoid overlaying. To overcome this, a pipe network buffer area can be delineated and summed with its subcatchment or the pipe network vector data can be "burned" into the DEM raster data. Figure 5.11 depicts the differences on watershed delineation using *GRASS – r.watershed* and a "Filled DEM" vs. "Burned DEM".

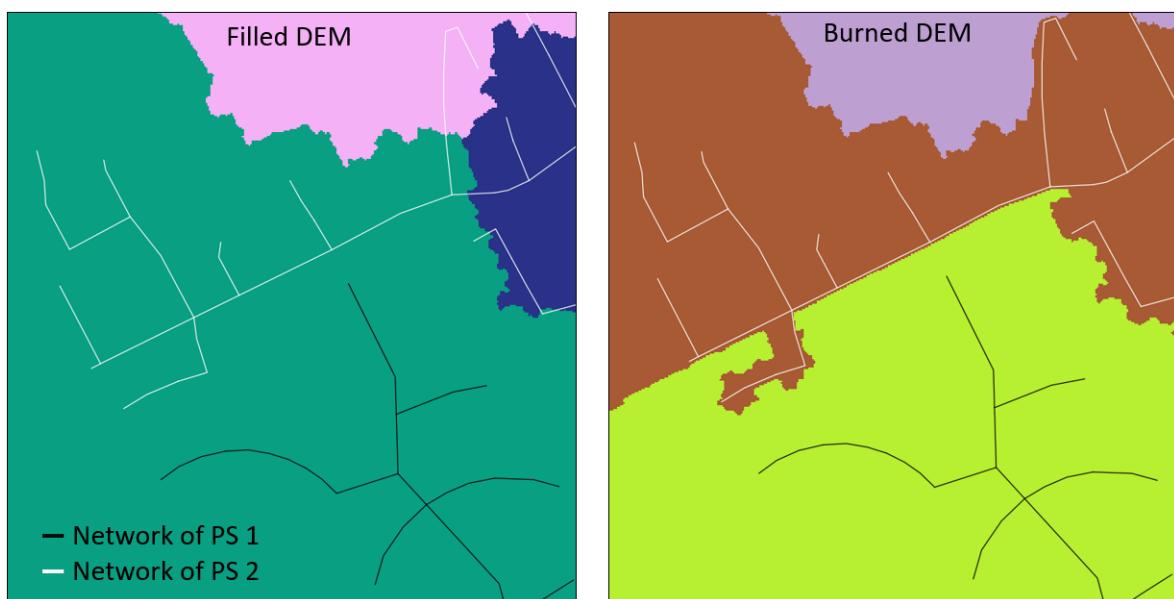


Figure 5.11: Comparison of sewershed delineation using Filled and Burned DEM

The pipe network included in the DEM acts as an artificial stream and successfully limits the subcatchment to its pumping station when pipes from different pumping station do not cross. Some level of manual adjustment is still required when pipe crossing exists. Other possible way to delineate the area influencing RDII flows into SSN is to assume that the area of influence is proportional to the size of the network components (i.e. pipe length and diameter). This assumes that the amount of defect (i.e. pipe cracks) increases when the size of the pipe also increases. In this study a combination of the buffer area proportional to the pipe length and diameter and the topographic crest was used to delineate the subcatchments depicted in figure 5.10. Further adjustments could still be done considering roads, railways, artificial barrers and streams present in the study area. Figure 5.12 depicts the two different delineation methods (D_1 = buffer over pipe size, D_3 = topographic crest delineation) used for the final delineation ($D_4 = D_1 + D_3$).

For organizational purposes, subcatchments where named after delineation method (i.e. D_4) and

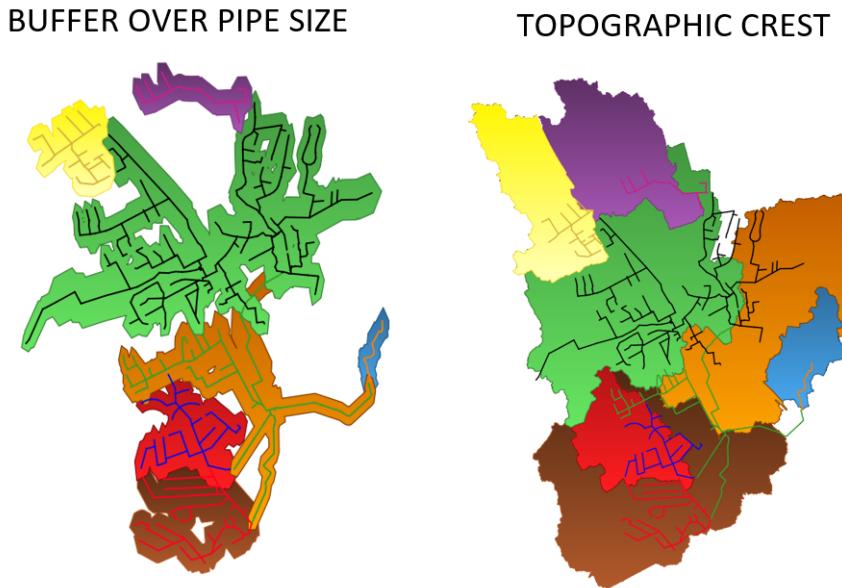


Figure 5.12: Delineation methods D1 (left) and D3 (right)

subcatchment id number (i.e. S2). For simplicity the letter were droped and they are referred hereafter simply by the numbers (i.e. 42). call the table and talk about the outlet (D4S4) INCLUDE Table with the area of each subc, network size, subcatchment area/total area, Subcatchment Area/Pipe Length, and totals.

To define other parameters used in SWMM's runoff module a Land Cover data set was assessed. The *FinnishCorineLandCover2018* (CLC2018) distributed by SYKE's open data platform. This data is an ensemble of different datasets such as Topographic database, digital road database of finland, building, land parcels and also data interpreted data from satellite images. The source data are from 2016-2017 period and the final raster data has 20m resolution. This dataset was produced by SYKE as part of EU Copernicus Land project and follows its standard nomenclature for land cover class. The raster dataset has four hierarchy levels of land cover class and several sub-classes [27]. If all four hierarchy levels and their sub-classes are used, a better spatial representation of land cover is obtained. However, the choice of which level to use depends on the size of the delineated sewershed. For this study three level of hierarchy were assumed to suffice. The vectorized and clipped dataset for the study area is depicted in figure 5.13.

Land cover of artificial surfaces were assumed as impervious areas whereas agricultural, forest, seminatural and wetlands assumed as pervious areas. This division was then used to calculate the percentage of impervious area, an input parameter for SWMM runoff module. The roughness coefficient (n manning's) was estimated also from the land cover. For this, a relation between standard copernicus land cover class description and manning's roughness coefficient table available in the literature ([25]) was created. An area weighted average calculation was carried as different types of land cover and, therefore, roughness coefficients are present within a subcatchment. In other words,

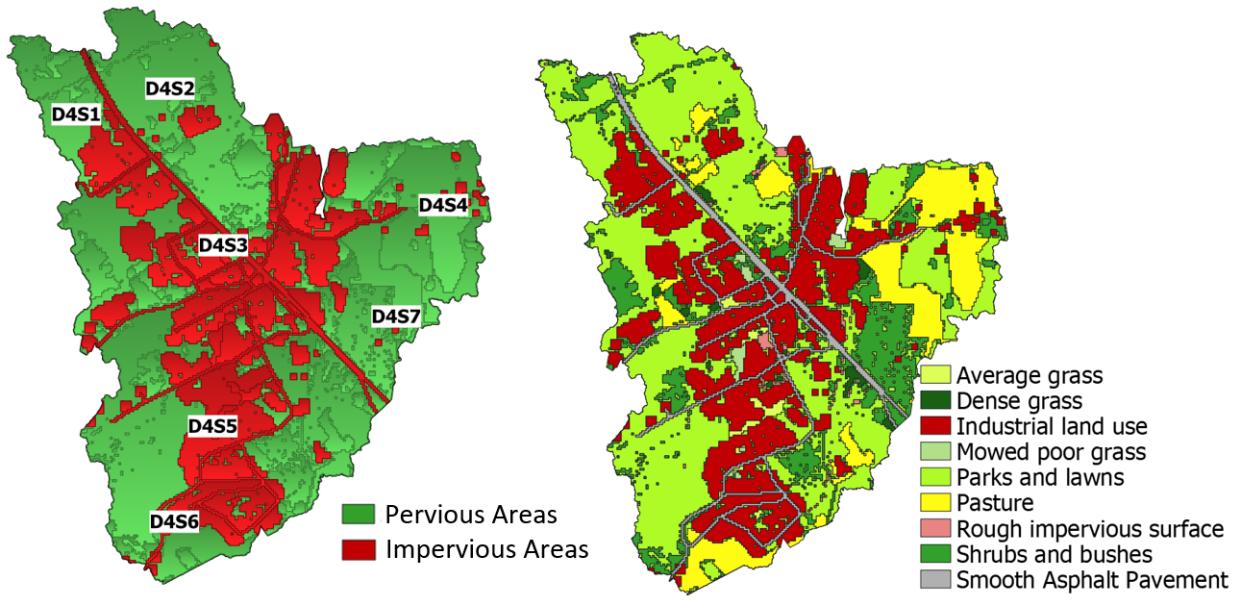


Figure 5.13: Pervious and Impervious division (left) and level 3 from SYKE's corine land cover for Jokela catchment (right) [27]

roughness coefficients of land cover with higher percentage over the total subcatchment's area have higher influence in the final value assigned to the subcatchment. This was done for both impervious and pervious area roughness estimate. The depression storage parameter were roughly estimated based on recommendations from Denver Urban Drainage and Flood Control District (UDFCD, 2007). The percentage of impervious land covers areas within a subcatchment were multiplied by 2.54[mm] and pervious land cover areas by 10.16[mm] before the area weighted average was taken. Table 5.3 depicts the estimated values for runoff parameters for each subcatchment as well as some relations to their respective SSN.

Table 5.3: SWMM Runoff parameter estimation

sub	area [ha]	area/ total [%]	Mean Slope [%]	n_imp [s/m ^{1/3}]	n_perv [s/m ^{1/3}]	Imp. area [%]	Ds imp [mm]	Ds perv [mm]	Area/ net [m]	Mean elev[m]
41	133.5	9.7	6.33	0.0321	0.081	34.6	0.88	6.64	488	87.36
42	156.8	11.6	6.05	0.0323	0.075	10.6	0.27	9.08	889	80.20
43	382.9	28.3	5.90	0.0317	0.082	52.5	1.33	4.83	179	77.10
44	313.9	23.2	4.89	0.0310	0.077	25.1	0.64	7.61	315	71.97
45	97.7	7.2	6.05	0.0310	0.081	49.4	1.25	5.14	240	77.78
46	212.2	15.7	6.01	0.0309	0.076	29.1	0.74	7.20	363	73.02
47	58.1	4.3	2.61	0.0350	0.092	1.3	0.03	10.03	930	71.09
sum* / mean	1355*	100*	5.41	0.0323	0.0806	28.9	0.74	7.22	486.3	76.93

Soil Superficial Deposits & Infiltration Parameters

The soil type coverage data was fetched from Geological Survey of Finland (GTK) [15] through its open data online service and was used to estimate the three parameters of Horton infiltration (see, 3.2.3). The available information was obtained as vector data containing superficial deposits of Finland with material produced between 1972-2007. GIS operations were carried using Qgis to delimit the area concerned Jokela's catchment. Coverage of superficial deposit is depicted in figure 5.14. Mixed soil types were simplified to facilitate model's parameter estimations.

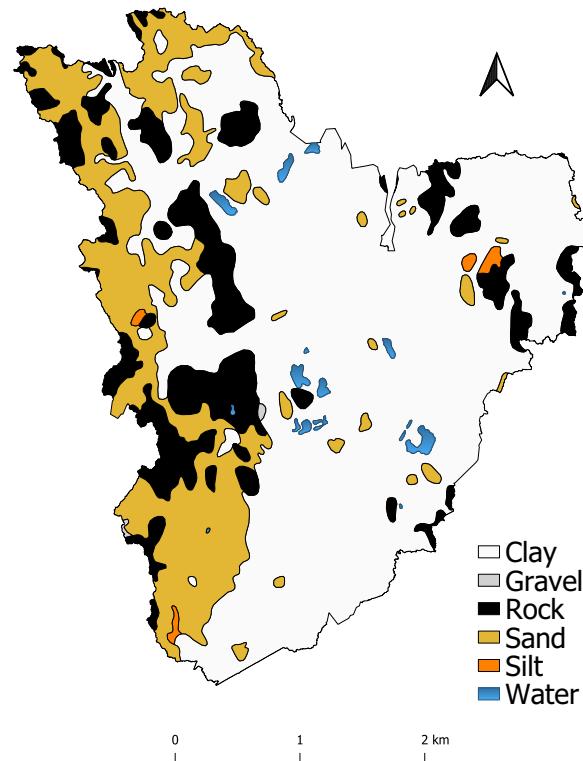


Figure 5.14: Jokela catchment soil superficial deposits

As informed in section 3.2.3, there are four parameters necessary to satisfy the Modified Horton infiltration model: Initial infiltration capacity (f_0); Minimum infiltration capacity (f_∞), decay coefficient (k_d), and recovery coefficient (k_r) that is calculated based on drying time. For simplicity, the parameters were averaged by the whole area of Jokela catchment without considering subcatchment's divisions per pumping station. Therefore, it is assumed that all subcatchments have the same parameters for the infiltration model. Estimation of parameters was carried as follows:

1. Initial infiltration capacity (f_0): values per soil type as estimated by Rossman and Huber [25] for DRY soils multiplied by 1.2 to account for vegetation present in the catchment.

2. Minimum infiltration capacity (f_∞): considered equal to saturated hydraulic conductivity with values estimated by Rawls et al. [22] and available in SWMM user help.
3. Decay coefficient (k_d): set to 4 [h^{-1}] as suggested by Rossman and Huber [25].
4. Recovery coefficient (k_r) in days: equation 5.1 as function of minimum infiltration capacity in inches/hour.

$$k_r = \frac{3.125}{\sqrt{f_\infty}} \quad (5.1)$$

The estimation of each parameter for each soil type is depicted in table 5.4. Area and percentage of coverage for each soil type was also calculated to spatially weight f_0 , f_∞ , and k_r using equation 5.2 to obtain a unique value of each parameter to represent the entire catchment.

Table 5.4: Modified Horton infiltration parameter estimation

soil type	area [ha]	coverage [%]	f_0	f_∞	k_r
Clay	887.5	66.0	30.5	0.3	28.8
Sand	267.9	19.9	152.4	117.8	1.5
Rock	181.6	13.5	3.0	0.03	90.9
Silt	5.9	0.4	91.4	6.5	6.2
Gravel	1.2	0.1	1524.0	1180.0	0.5
Water	11.5	0.9	-	-	-
Total	1355.7	100.0	-	-	-

$$\text{Parameter} = \sum_n^m (\text{Parameter}_n \cdot \text{Coverage}_n \cdot 0.1) \quad (5.2)$$

where:

$\text{parameter} = f_0, f_\infty \text{ or } k_r$

$n = \text{soil type (excluding water)}$

after applying equation 5.2 was applied to the estimated parameters of table 5.4. The result of k_r was set to its upper limit range since it was greater than 14 days because of the predominance of soil types with relatively low saturated hydraulic conductivity (*Clay and Rock*). The following unique parameters were obtained and included to SWMM model:

$$f_0 = 52.2 \text{ [mm/h]}, f_\infty = 24.6 \text{ [mm/h]}, k_d = 4 \text{ [1/h]}, k_r = 14 \text{ [days]}$$

5.2.5 Water Table & Groundwater Flow Parameters

As mentioned in the literature Bennett et al. [7], Vallabhaneni and Burgess [32], Barden et al. [6], infiltration into the sewer lines can be caused by the seasonal elevation of groundwater table or other condition that increased soil moisture content causing a temporary saturated zone. Elevation of the groundwater table around Jokela town was assessed in this section as an attempt to identify a possible correlation with seasonal variations of the water table and the flow measurements of the town's sanitary sewer network. Information of water table levels was collected from the Finnish Environmental Institute (SYKE) through its open data service [27] and provided by Tuusula Water Utility. Data of three observation wells from SYKE were available surrounding and Jokela town and one from Tuusula Water Utility database within the delineated catchment as depicted in figure 5.15. The recording period and routines among the four stations varies considerably - from one record per month to one record per year.

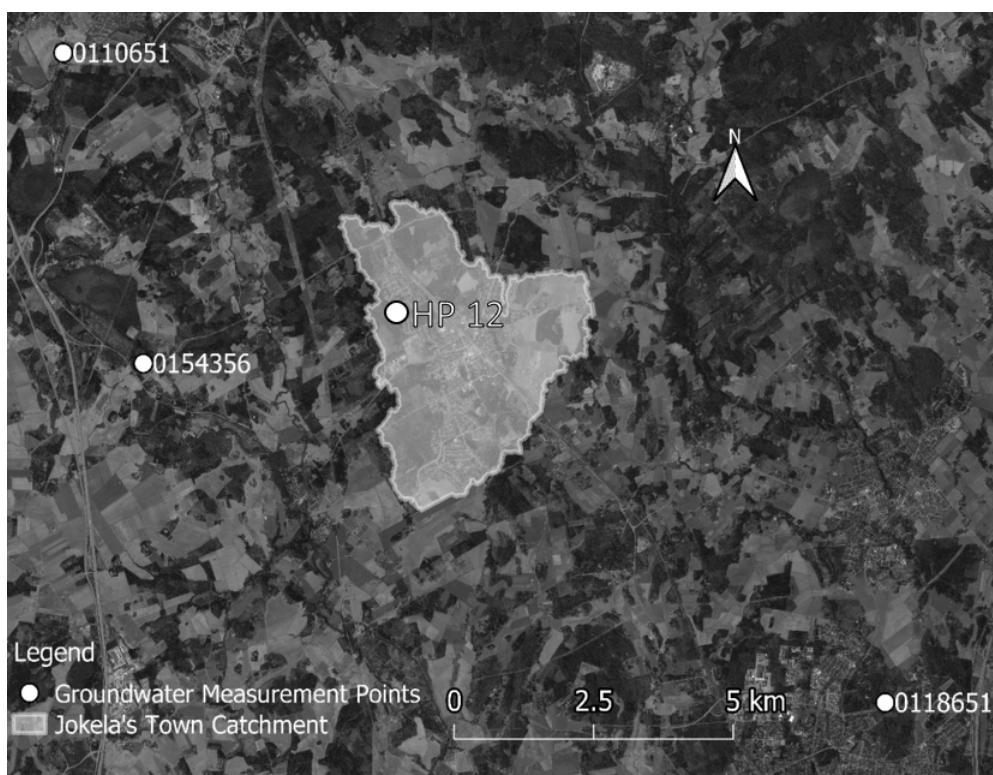


Figure 5.15: Location of Observation Wells around Jokela's catchment

Data of measurements from 2004 to 2016 of station 0118651 -located southeast from Jokela- were collected. Years with less than six months recorded were left out: 2007; 2008; and 2017. All the eleven years records showed an elevation on the groundwater table from March to May.

Only yearly measurements were available for the station 0154356 located around 5km west from Jokela. The records are from different months, mostly during spring and summer. Therefore, assessment of monthly variation for the same year was not possible. However, the available data suggests

slightly higher water table levels on average from January to June for the period of 1999-2017.

Only Station 0118651 and HP 12 had measurements from 2018. This year was relevant to compare with flow measurements available in the sewer network (see section 5.2.1). According to the measurements of these two stations, the water table elevation was higher in the first five months of the year (Jan-May) with increase period from February to May and decrease from May to December 2018. Monthly mean of measured groundwater table from all SYKE's station as well as monthly measurements of 2018 from stations 0118651 and HP 12 were plotted in figure 5.16.

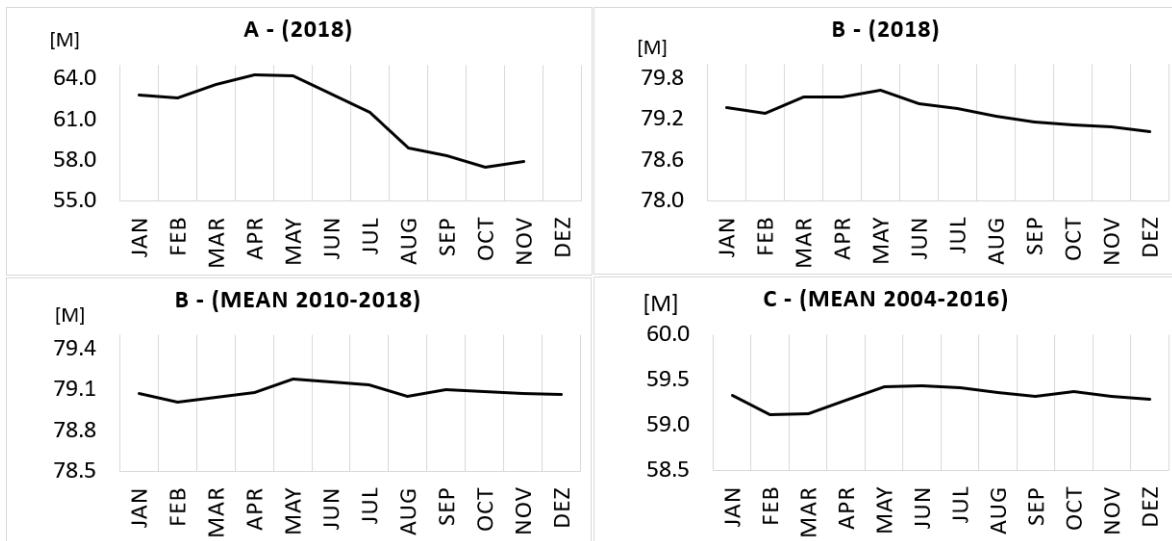


Figure 5.16: Groundwater table measurements. Two upper plots are measurements from 2018 only whereas the other two are monthly mean values of the period informed. (A = H12, B = 0110651, C = 0118651). Data from Tuusulan Vesihuolto and SYKE [27]

By comparing the two plots of figure 5.5 and figure 5.16, it is possible to observe that a correlation between higher groundwater table and infiltration amount for the first five months into the sanitary sewer of 2018 exists. The estimated amount of RDII peaks on April as the groundwater table elevation measured by the closest observation well to the Jokela's SSN (H 12).

As the other data assessed in this study, care should be taken when extrapolating the behavior of the measurements to different years. Groundwater elevation also varies on year by year basis and the same behavior is not always observed. However, the plots of figure 5.16 presenting the mean values for two different observation wells suggests that an increase of the groundwater table occurs from February to May on average from 2004-2018.

For simplicity, it was assumed that all subcatchments modeled in SWMM share an aquifer with the same characteristics, but with different groundwater infiltration capacity. Therefore, the estimation of parameters is divided here in two parts: 1. Aquifer Parameters and 2. Groundwater Flow Parameters. The choice for different groundwater infiltration parameters for each of the subcatchments was made as an attempt to investigate differences in groundwater infiltration quantity during calibration process. In other words, a higher rate of groundwater infiltration (f_G) in one of the subcatchments

could suggest that there are more defects (i.e. pipe cracks) to its network and, therefore, could be candidate for further rehabilitation analysis.

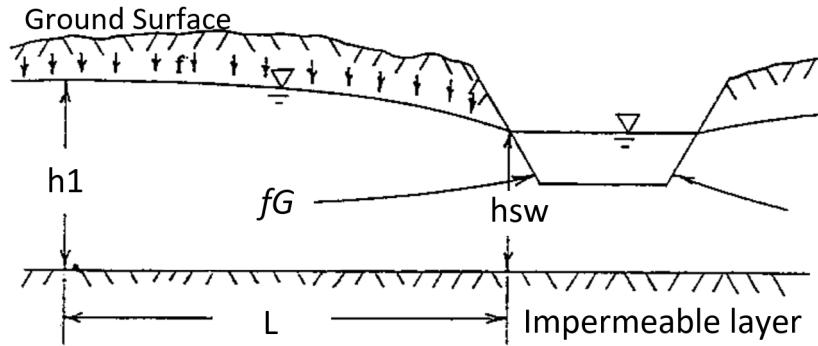


Figure 5.17: Dupuit-Forcheimer lateral seepage to adjacent channel. Modified from

Dupuit-Forcheimer Lateral Seepage equation (5.3) was chosen to calculate f_G . There was no indication that this choice was the best for the study area. The reasons why Dupuit-Forcheimer was chosen over the three examples described in SWMM's hydrology reference manual [25] were: 1. Unlike the linear reservoir method, all parameters of the equation could be estimated based on topographical data before calibration process; 2. Simpler than hooghoudt's method since the groundwater infiltration flow is evaluated only at one node. Dupuit-Forcheimer seepage scheme is depicted in figure 5.17.

$$f_G = A_1 \cdot d_L^2 - A_3 \cdot d_L \cdot h_{sw} \quad (5.3)$$

where:

f_G = Flux from saturated zone to receiving node [m/s]

$A_1 = -A_3 = 2K_s \cdot L^{-2}$ [s m]⁻¹

$K_s = f_\infty$ = Saturated hydraulic conductivity [m/s]

L = Distance between h_1 and h_{sw} [m]

d_L = Saturated zone elevation [m]

h_{sw} = Elevation of water inside the receiving node with same reference as the other elevations [m]

it was assumed that the water table elevation distribution is similar to the surface elevation spatial distribution to estimate L parameter. This is a rough estimate since there is no indication that the water table elevation follows the surface elevation in the study area. For this, the location of highest surface elevation point per each subcatchment was found using available DEM data and Qgis raster calculator. This point was assumed to be h_1 and its distribution is showed in figure 5.19. The chosen equation is valid only for flows entering the receiving node. Thus, flow occurs only when $d_L > h_{sw}$.

The distance between these two points were measured and assigned to L for each subcatchment groundwater component. The saturated hydraulic conductivity (K_s) was assumed to be constant for



Figure 5.18: Position of h_2 for each subcatchment

all subcatchments and its value was estimated in section 5.2.4. d_L is computed every time step of the simulation, but its initial value (d_{L0}) is required and it was set as the bottom elevation of each receiving node (h_B) as depicted in table 5.5. The aquifer bottom elevation used as reference for all the other elevations was also set as h_B since no information of the impermeable zone was assessed. h_{sw} value is obtained from hydraulic model flow routing as the wastewater depth plus h_B .

The aquifer parameters were estimated from soil properties data tables available in the literature ([22],[25]) s done for infiltration parameter estimation also using equation 5.2 to average the values from each subcatchment to the entire catchment area. The parameters and respective estimated values are as follows:

- porosity = 0.41 [m^3/m^3]
- field capacity = 0.27 [m^3/m^3]
- wilting point = 0.18 [m^3/m^3]
- Tension slope = 15
- hydraulic conductivity slope = 48.6
- Fraction of evaporation = 0.35 [fraction]
- Lower Evap. Depth = 5 [m]
- Lower GW Loss Rate = 1 [m/s]
- Initial Unsat. Zone Moisture = 0.27 [fraction]

As mentioned on section 5.1, there are streams crossing the delineated catchment area not included in the model. This can be observed when checking the two-aquifer model water budget since interaction between aquifer and surface water (rivers and streams) is not modeled omitting a probably

considerable baseflow loss. This loss is represented, in this study, by the seepage to deeper aquifers (f_L). The constant rate of seepage is a user-supplied parameter and was used as one of the calibration parameters to simulate long term groundwater table elevation.

Table 5.5: Groundwater Flow parameter estimation

Subcatchment	h_B [m]	L [m]	$A1$ [$s \cdot m$] $^{-1}$
41	75.25	1108	4.3E-05
42	66.17	2026	1.3E-05
43	60.04	2040	1.3E-05
44	62.22	1397	2.7E-05
45	58.09	1324	3.0E-05
46	52.2	1797	1.6E-05
47	64.2	1319	3.0E-05

Calibration of L parameter may apply as h_1 were located, as expected, at the borders of the subcatchments with the rough estimation described above. It is likely that h_1 is located further than it is in reality. The calibration should then move towards decreasing the value of L . This decrease would result on an increase for $A1$ values resulting on an increase in the groundwater infiltration rate. A flow rate not greater than the observed is expected when using the initial values as suggested in table 5.5 assuming a constant saturated hydraulic conductivity (K_s). Therefore, f_G is likely closer to its lower limit range and calibration efforts should focus on increasing f_G by decreasing L .

First simulation was carried with the values presented on table 5.5. The result of the first simulation of groundwater table (d_L) from subcatchment 41 is plotted in figure ???. The variation of the simulated groundwater table did not follow completely the pattern observed from well HP 12 even though both simulated and measured data are from the same delineated area (subcatchment 41). There is an overall increase on the simulation from Jan to May. However, a decrease on d_L from February to March was simulated, but not observed.

It is also important to note that the average elevation of the bottom of the junctions in Jokela's SSN is higher than the groundwater table measured (figure 5.16 suggesting that the groundwater infiltration into the SSN is not caused by the water table elevation per se, but by the increase in the soil moisture during aquifer recharge periods (such as snowmelt periods). This is the reason the initial d_L elevations were set as the bottom elevation of the network's receiving node instead of the measured values. The groundwater elevation results from the model simulations is rather a representation of the increase in soil moisture caused by infiltration than the actual groundwater table levels of the aquifer in the sewershed.

5.2.6 Synthetic Unit Hydrograph: RTK

Synthetic Unit Hydrographs RTK were added to each of the seven pumping stations as a representation of the hydrological response of each of the delineated subcatchments as illustrated in figure 5.20.

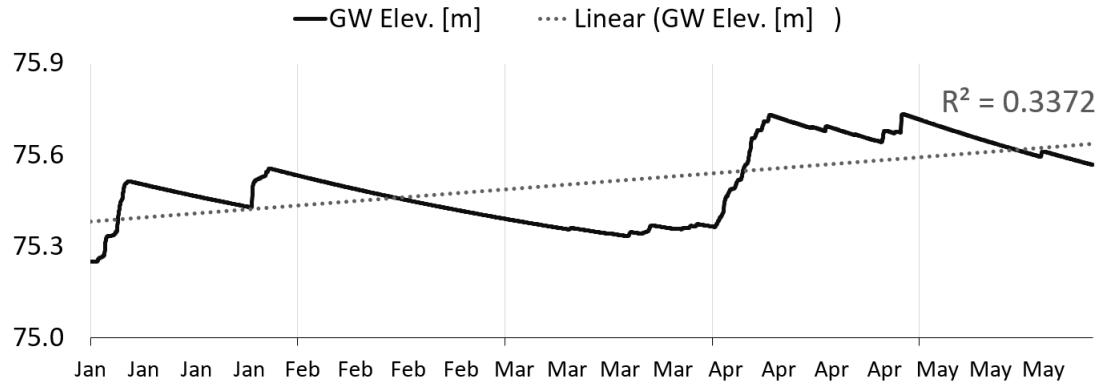


Figure 5.19: Groundwater table first simulation from January to May of 2018

The available parameters to be estimated for each sewershed are in total 15:

- $R_{short}, R_{medium}, R_{long}$
- $T_{short}, T_{medium}, T_{long}$
- $K_{short}, K_{medium}, K_{long}$
- $Dmax_{short}, Dmax_{medium}, Dmax_{long}$
- $Drec_{short}, Drec_{medium}, Drec_{long}$

A literature review was carried to find a suitable range for the estimation of parameters, but they may vary greatly according to the type of watershed being studied. As an example, one may expect a different range of parameters between highly urbanized and semi-urban or rural areas. As stated by Vallabhaneni and Burgess [32] R-values are proportional to the sewershed delineated area. As a hypothetical exercise one can image the differences in the are of the two delineation methods depicted in figure 5.12 with D3 having a larger area. The fraction of rain falling over the D3 area that finds its way into the sewer system is smaller than the fraction of D1. As the R-values are calibrated to be further multiplied by the sewershed area resulting the desired volume entering the pipe network. Precipitation input data is another factor influencing the estimation of R-values. In case precipitation measurements used for calibration of R-values fails to capture the realistic volume falling over the sewershed, estimated R-values will also represent erroneous fractions of the precipitation entering the sanitary sewer network. Although distinct precipitation measurements lead to different R-values, rain gauges and radar in some cases do capture similar rainfall pattern as concluded by [?]. Therefore, both rain gauge and radar may yield similar T and K estimated values. In summary, R-values are influenced by sewershed area, precipitation input, network and catchment's characteristics Barden et al. [6] applied RDII UH for continuous simulation of a catchment with residential dominant land use and obtained a range of 4 to 26% for R-values for 14 months recorded period. The study obtained the best results when applying seasonal RTK-values with monthly IA parameters and well performing

model for moderate and large storms with seasonal RTK and IA parameters. Vallabhaneni and Burgess [32] proposes a range for RTK parameters in urban sewersheds depicted in table 5.6. This range was added as limits for the DDS optimization algorithm.

Table 5.6: Range of RTK parameters by Vallabhaneni and Burgess [32]

Curve	T [h]	K
Short-Term	0.5 - 2	1 - 2
Medium-Term	3 - 5	2 - 3
Long-Term	5 - 10	3 - 7

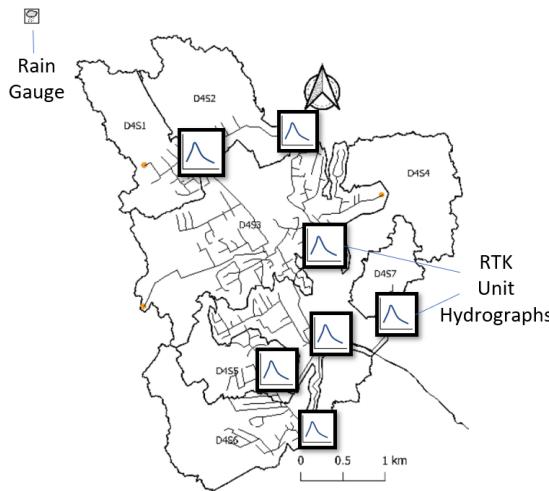


Figure 5.20: Representation of Jokela's Hydrological Model using RTK UH method

5.3 Calibration

5.3.1 Calibration of physically based model

[to do]

- include the table with the manually calibrated values for all parameters
- Maybe include the results of optimization algorithm results (if time allows).

First calibrations were carried for long-term period of the first five months of 2018 (from Jan to May). The groundwater flow into the SSN was lower as expected once the flow parameters were set to the minimum proposed range. Parameter $A1$ value was then increased to approximated one order of magnitude higher. Figure 5.21 depicts the results with the increased values of $A1$.

The time of flow increase was successfully achieved. The snowmelt routine successfully utilizes the weather input parameters provides water that infiltrates recharging the aquifer and increasing

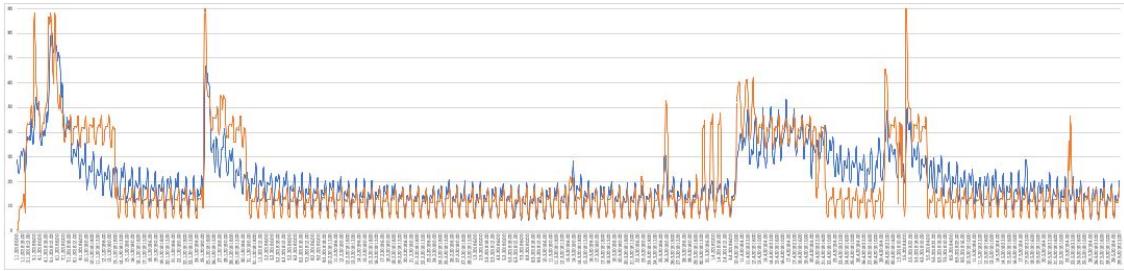


Figure 5.21: Long-term simulation of winter period evaluated at Jokela Pumping station using Dupuit-Forcheimer and unidirectional flow condition. Simulated = Orange, Observed = Blue

the flow into the SSN. Sharp peaks were also simulated successfully by the runoff module in the same period as the observed values. However, the flow magnitudes were still over estimated and the groundwater flow was abruptly interrupted by SWMM routine that sets the GWI flow to zero once the groundwater table drops below the level of the water in the receiving node. SWMM uses this routine whenever $A3$ value parameter is set to different than zero.

The runoff block parameter $\%routed$ was calibrated as described in section 3.2.1 to simulate losses caused by non-modeled streams and stormwater sewer network. To avoid sudden drops of groundwater flow, the $A3$ parameter was set to zero modifying the firstly proposed Dupuit-Forcheimer formula 5.3 excluding the second term to equation 5.4.

The flow characteristics of the winter period was much better simulated with groundwater infiltration following the pattern of observed data as depicted in figure 5.22. Some adjustments to runoff parameters are still necessary as well as groundwater flow parameters for periods with severe snowmelt.

$$f_G = A_1 \cdot d_L^2 \quad (5.4)$$

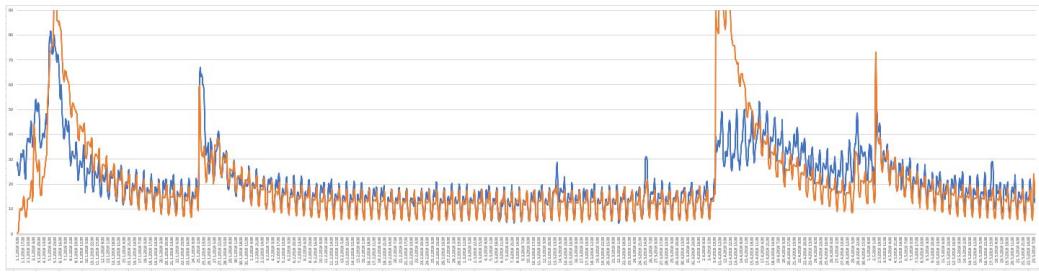


Figure 5.22: Long-term simulation of winter period evaluated at Jokela Pumping station using different GWI equation and unidirectional flow condition. Simulated = Orange, Observed = Blue

5.3.2 Calibration of RTK Model

[to do]

- better description of the calibration algorithm used and the search space.
 - include calibration results of other storms chosen for parameter calibration. Not figures, but maybe a table with nash-sutcliffe/RMSE and errors in peak and volume of the respective storm) .

The first storms used to investigate the performance of RTK model with DDS optimization algorithm occurred in 2018 summer from July 1st to July 6th. The reason why this period was chosen is that different storm intensities were recorded causing sharp increase in the sanitary sewer flows as depicted in figure 5.23.

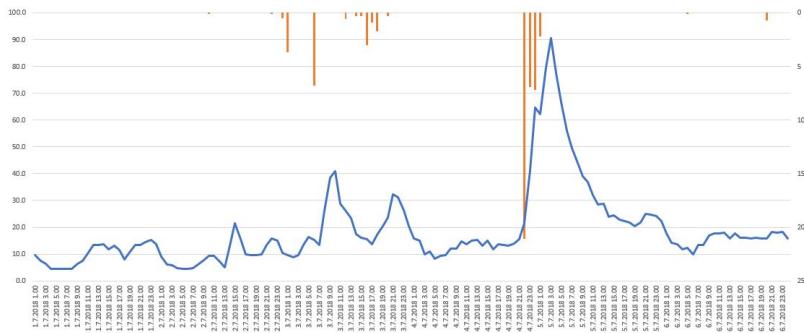


Figure 5.23: Rain 04

Figures 5.24, 5.25, 5.26, and 5.27, have the plot of simulated versus observed results where DDS optimization algorithm was set to use respectively 100, 300, 600, and 1000 iterations.

Table 5.7 depicts the results evaluated by nash-sutcliffe and root mean square error coefficients for pumping station 1 and pumping station 2.

Table 5.7: RTK rain 04 simulation results evaluation. Nash-Sutcliffe = 1 means a perfect model

Number of Iterations	Nash-Sutcliffe		RMSE	
	PS1	PS2	PS1	PS2
100	0.82	0.81	6.49	4.59
300	0.85	0.85	6.02	4.08
600	0.86	0.86	5.68	3.98
1000	0.88	0.79	5.39	4.75

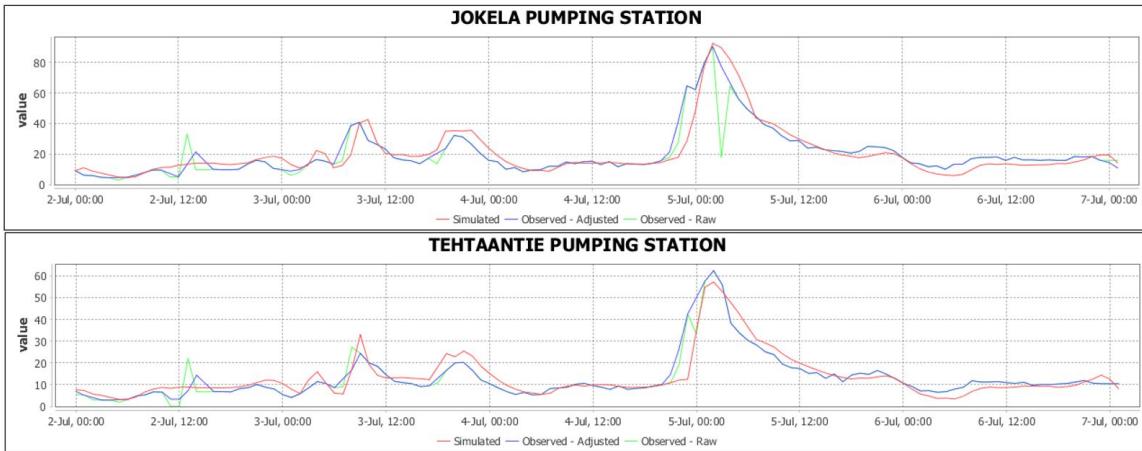


Figure 5.24: Calibration results of RTK parameters using Rain04 and 100 iterations of DDS optimization algorithm

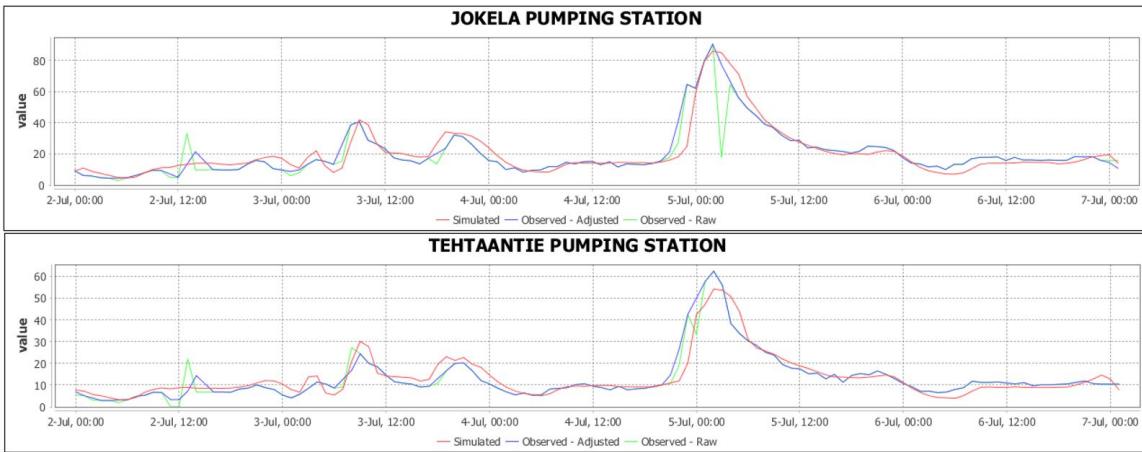


Figure 5.25: Calibration results of RTK parameters using Rain04 and 300 iterations of DDS optimization algorithm

5.4 Validation

5.5 Validation with historical Meteorological Data

[to do]

5.6 Validation with historical Weather Forecast Data

[to do]

- flow data of 2019 that will be used for the validation using weather forecast received today

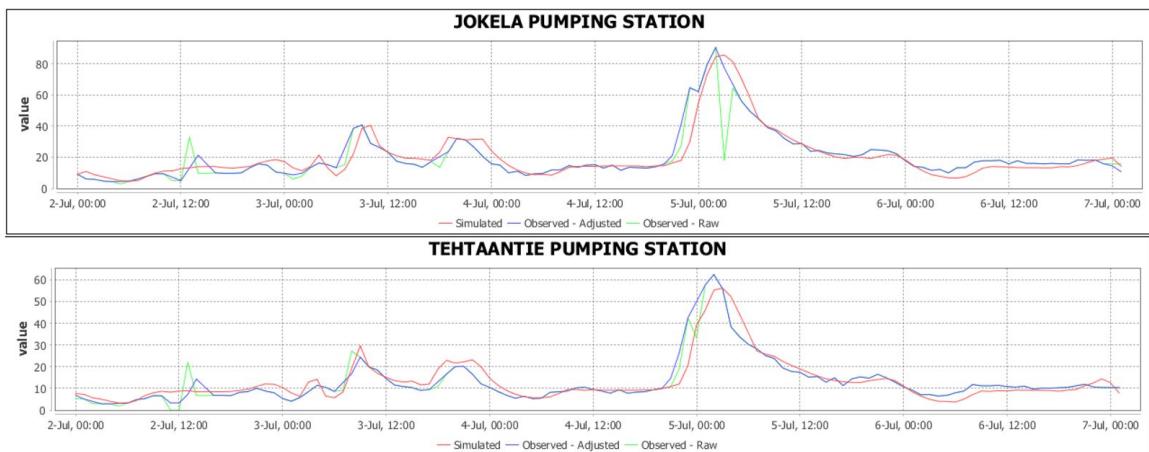


Figure 5.26: Calibration results of RTK parameters using Rain04 and 600 iterations of DDS optimization algorithm

(15.07.19).

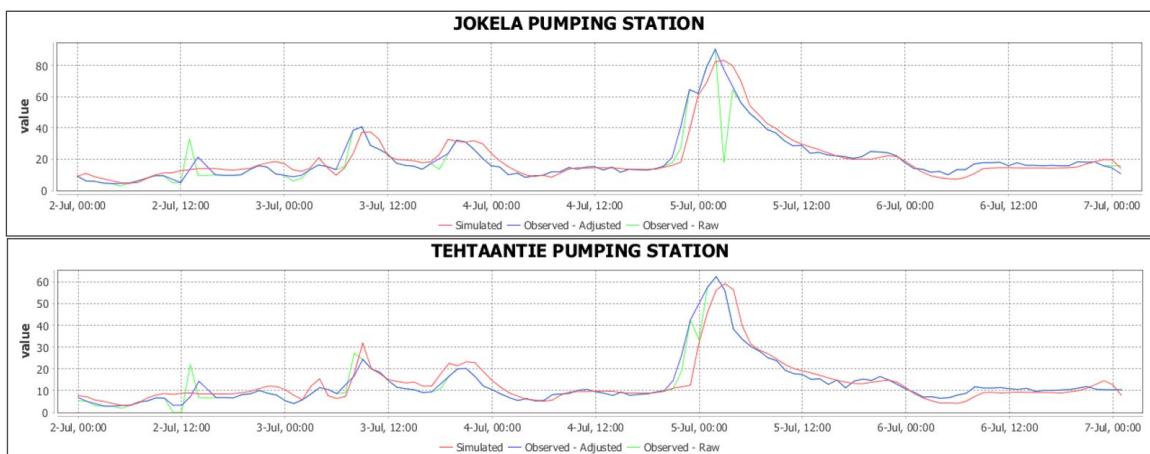


Figure 5.27: Calibration results of RTK parameters using Rain04 and 1000 iterations of DDS optimization algorithm

6 RESULTS AND DISCUSSION

[To Do]

- results of long-term simulations x short term simulations with update of initial condition parameters
- final conclusion of the performance of both methods and discussion of their advantages and advantages for real-time applications

6.1 Water Balance

[To Do]

- water balance of long-term simulation versus "short term" simulation RTK and Physically based

7 CONCLUSION

7.1 Conclusion

[to Do]

- conclusions of the different model creation, parameter estimation, and simulation results for the proposed real time modelling application.

7.2 Recommendations

[to Do]

- Data input routines (frequency to retrieve data from APIs)
- use of multi-objective function for DDS calibration algorithm
- Definition of parameter range to reduce search space and speed up calibration routines - Use of open data (free services and their limited quality)

BIBLIOGRAPHY

- [1] A. Osman Akan. Horton Infiltration Equation Revised. *Journal of Irrigation and Drainage Engineering*, 118(5):828–830, 1992.
- [2] R.J. Akan, A.O. and Houghtalen. *Urban Hydrology, Hydraulics, and Stormwater Quality: Engineering Applications and Computer Modeling*. Wiley, 2003. ISBN 9780471431589.
- [3] Eric Anderson. Snow accumulation and ablation model-SNOW-17, NWSRFS user manual. Technical Report January, 2006. URL http://www.nws.noaa.gov/oh/hr1/nwsrfs/users{_}manual/part2/{_}pdf/22snow17.pdf.
- [4] Eric A . Anderson. National Weather Service River Forecast System- Snow Accumulation And Ablation Model. Technical report, Washinton, DC, 1973. URL <https://repository.library.noaa.gov/view/noaa/13507>.
- [5] Richard Arsenault, Annie Poulin, Pascal Côté, and François Brissette. Comparison of Stochastic Optimization Algorithms in Hydrological Model Calibration. *Journal of Hydrologic Engineering*, 2013. ISSN 1084-0699. doi: 10.1061/(asce)he.1943-5584.0000938.
- [6] Greg Barden, Tim Fallara, Hunter Kelly, Fang Cheng, Benjamin J. Sherman, and Edward Burgess. Comparison of RDII Unit Hydrograph Approaches for Continuous Simulation using SWMM 5. *Journal of Water Management Modeling*, 6062:209–222, 2011. doi: 10.14796/jwmm.r241-12.
- [7] David Bennett, Reggie Rowe, Marie Strum, David Wood, Nancy Schultz, Kelly Roach, Mike Spence, and Virgil Adderley. Using flow prediction technologies to control sanitary sewer overflows. Technical report, 1999.
- [8] Nam Jeong Choi. *Understanding sewer infiltration and inflow using impulse response functions derived from physics-based models*. PhD thesis, University of Illinois at Urbana-Champaign. URL <https://www.ideals.illinois.edu/handle/2142/90540>.
- [9] Ana Maria de Roda Husman, Ton de Nijs, Ankie Sterk, Jack F. Schijven, and Heleen de Man. Climate change impact on infection risks during bathing downstream of sewage emissions from CSOs or WWTPs. *Water Research*, 105:11–21, 2016. ISSN 00431354. doi: 10.1016/j.watres.2016.08.053.
- [10] S Dent, R Blair Hanna, and L T Wright. Automated Calibration using Optimization Techniques with SWMM RUNOFF. *Journal of Water Management Modeling*, pages 220–238, 2004. ISSN 2292-6062. doi: 10.14796/JWMM.R220-18. URL www.chijournal.org.

- [11] Y. Djebbar and P. T. Kadota. Estimating sanitary flows using neural networks. *Water Science and Technology*, 1998. ISSN 02731223. doi: 10.1016/S0273-1223(98)00752-5.
- [12] EPA. Review of Sewer Design Criteria and RDII Prediction Methods January 2008. Technical report. URL <https://nepis.epa.gov/Adobe/PDF/P1008BP3.pdf>.
- [13] FMI. Open data - Finnish Meteorological Institute. URL <https://en.ilmatieteenlaitos.fi/open-data>.
- [14] Hazem Gheith. A Detailed Procedure for Separating RDII Stages and Generating a Single Set of RTK Hydrographs for Continuous Simulation. *Journal of Water Management Modeling*, 6062: 187–208, 2011. doi: 10.14796/jwmm.r241-11.
- [15] GTK. Geological Survey of Finland (GTK) Spatial Data Products. URL <https://hakku.gtk.fi/en/locations/search>.
- [16] Milja Heikkinen, Paula Schönach, and Ilmo Massa. Politicization of wastewater overflows: The case of the Vantaa River, Finland. *Water Policy*, 18:1454–1472, 2016. doi: 10.2166/wp.2016.011. URL https://helda.helsinki.fi/bitstream/handle/10138/187724/Wastewater_{_}overflow_{_}pre_{_}copy_{_}edit_{_}final.pdf?sequence=1.
- [17] Joong Gwang Lee, Christopher T. Nietch, and Srinivas Panguluri. SWMM Modeling Methods for Simulating Green Infrastructure at a Suburban Headwatershed : User's Guide. (October), 2017.
- [18] Jiuling Li, Keshab Sharma, Yiqi Liu, Guangming Jiang, and Zhiguo Yuan. Real-time prediction of rain-impacted sewage flow for on-line control of chemical dosing in sewers. *Water Research*, 149:311–321, 2019. ISSN 18792448. doi: 10.1016/j.watres.2018.11.021. URL <https://doi.org/10.1016/j.watres.2018.11.021>.
- [19] Charles Mosley, Shawn Dent, Paul Kadota, Leonard T. Wright, and Yassine Djebbar. Comparing Rainfall Dependent Inflow and Infiltration Simulation Methods. *Journal of Water Management Modeling*, (January), 2001. doi: 10.14796/jwmm.r207-16.
- [20] M K Muleta and P F Boulos. Analysis and Calibration of RDII and Design of Sewer Collection Systems. World Environmental and Water Resources Congress. Technical report, 2008. URL <https://archive.innovyze.com/news/showcases/RDIIAnalystPaperASCE2008.pdf>.
- [21] NSL. National Land Survey of Finland. URL <https://tiedostopalvelu.maanmittauslaitos.fi/tp/kartta?lang=en>.

- [22] Walter J. Rawls, Donald L. Brakensiek, and Norman Miller. Green-ampt Infiltration Parameters from Soils Data. *Journal of Hydraulic Engineering*, 109(1):62–70, 1983. ISSN 0733-9429. doi: 10.1061/(asce)0733-9429(1983)109:1(62).
- [23] Brent Robinson. Modeling Sanitary Sewer Groundwater Inflow Rehabilitation Effectiveness in SWMM5 Using a Two Aquifer Approach. *Journal of Water Management Modeling*, (Sanipor 2014):1–12, 2015. doi: 10.14796/jwmm.c385.
- [24] Larry A Roesner. Storm Water Management Model Manual. *Usepa*, I(July), 2009.
- [25] Lewis A Rossman and Wayne C Huber. Storm Water Management Model Reference Manual. Volume I - Hydrology (Revised). EPA/600/R-15/162A. I:231, 2016. URL <https://www.epa.gov/water-research/storm-water-management-model-swmm>.
- [26] Markus Sunela. *Real-Time Control Optimization of Water Distribution System with Storage*. 2017. ISBN 9789949831654. doi: 10.3917/lett.051.0109.
- [27] SYKE. Finnish Environment Institute (SYKE) Open Environmental Information Systems. URL https://www.syke.fi/fi-FI/Avoin{_}tieto/Ymparistotietojarjestelmat.
- [28] Hessam Tavakol-Davani, Erfan Goharian, Carly H. Hansen, Hassan Tavakol-Davani, Defne Apul, and Steven J. Burian. How does climate change affect combined sewer overflow in a system benefiting from rainwater harvesting systems? *Sustainable Cities and Society*, 27:430–438, 2016. ISSN 22106707. doi: 10.1016/j.scs.2016.07.003.
- [29] Henri Tikkanen. *Hydrological modeling of a large urban catchment using a stormwater management model (SWMM)*. PhD thesis, 2013.
- [30] Bryan A. Tolson and Christine A. Shoemaker. Dynamically dimensioned search algorithm for computationally efficient watershed model calibration. *Water Resources Research*, 2007. ISSN 00431397. doi: 10.1029/2005WR004723.
- [31] C. Valeo and C.L.I. Ho. Modelling urban snowmelt runoff. *Journal of Hydrology*, 299(3-4): 237–251, 2004. ISSN 0022-1694. doi: 10.1016/J.JHYDROL.2004.08.007. URL <https://www.sciencedirect.com/science/article/pii/S0022169404003683>.
- [32] Srinivas Vallabhaneni and Edward H Burgess. Computer Tools for Sanitary Sewer System Capacity Analysis and Computer Tools for Sanitary Sewer System. *Environmental Protection*, (October):1–104, 2007.
- [33] David Walker. an Artificial Neural Network-Based Rainfall Runoff Model for Improved Drainage Network Modelling. (1), 2014.

- [34] M. Wallner, U. Haberlandt, and J. Dietrich. Evaluation of different calibration strategies for large scale continuous hydrological modelling. *Advances in Geosciences*, 2012. ISSN 16807340. doi: 10.5194/adgeo-31-67-2012.