Tampere University

Moein Ashofteh Beyraki

AUTOMATED CATCHMENT DEFINITION FOR SIMULATIONS OF CITY-SCALE STORMWATER NETWORK

A CASE STUDY TO IMPROVE HELSINKI COMBINED SEWER SYSTEM MODEL

Master's thesis
Faculty of Engineering and Natural Sciences
Examiners: Visiting Teacher Niko Niemelä
University Instructor Annina Takala
Dec 2022

ABSTRACT

Moein Ashofteh Beyraki: Automated catchment definition for simulations of city-scale stormwater network
Master's thesis
Tampere University
Master of Science (Technology) Degree Programme in Environmental Engineering
Dec 2022

The growing impact of climate change on the frequency of extreme weather events and the quality of surface water has necessitated more accurate urban stormwater modeling. Establishing stormwater models can save resources and minimize detrimental environmental impacts when the most critical parts of the network are identified first. The EPA Storm Water Management Model (SWMM) is widely used to study runoff in urban areas. Rainfall-runoff modeling with SWMM requires precise characterization of sub-catchments. However, the delineation and parametrization of accurate sub-catchments for large urban areas and a city-scale model is a time-consuming and complex process, which makes it tedious and prone to designers' errors.

Manual catchment delineation and parametrization challenges indicate the need for automated tools to save modelers a significant amount of time and prevent manual errors. Nonetheless, automated methods can only be used if they are proven to demonstrate their ability to provide realistic results. Furthermore, selecting the spatial resolution of the sub-catchments remains a challenge for simulating the models without a high computational burden.

The main objectives of the thesis were to assess methods for automated delineation and parametrization of SWMM sub-catchments for city-scale modeling applications. The target was to avoid manual work as much as possible while keeping the results consistent using varying Geographic Information System (GIS) approaches and literature values. Testing the automated methods was investigated in two main steps. In the first step, four different GIS-based methods are used, namely: the old HSY method, QGIS, GISTOSWMM, and SCALGO. SWMM sub-catchments were created using these methods in four selected case areas within the Helsinki combined sewer network (CSN) in Finland. The methods were compared with each other, focusing on the fluency of the process, hydraulic results, spatial resolution, and the capability to be used in a city-scale model. In addition, the thesis discusses the impacts of using automation, a new imperviousness layer, and varying levels of detail in catchment definition. In the second step, the best method was used for the whole Helsinki CSN for evaluation in an extensive city-scale model.

The results indicate that the SCALGO method can be used to make hydrological models that range from small to city-scale due to its fast and accurate catchment definition, adjustable spatial resolution, and good model performance. It was found that including stormwater inlets in the SCALGO catchment definition method had a minor effect on the hydraulic results. The use of merged sub-catchments with a minimum adjustable area via the SCALGO toolbox was found practical for finding a suitable subcatchments size. Furthermore, the new data (LaserVesi) obtained from an automated imperviousness surface detection model was useful in estimating the sub-catchments imperviousness parameter. The results of this study make it easier to update sub-catchments for city-scale models. In Helsinki, this is particularly interesting as the network is upgraded annually and more separate sewers are built. While this study focused on automated catchment definition methods for city-scale networks, the findings provide in-depth information about SWMM models' automatic implementation for urban catchments without calibration.

Keywords: Stormwater modeling; automated catchment definition; Catchment delineation and parametrization; SWMM

PREFACE

This master's thesis was carried out as a collaborative project between Fluidit, the Helsinki Region Environmental Services Authority (HSY), SCALGO, and Tampere University. I want to thank all the partners for entrusting me with this project.

The people at Fluidit Company were the main reason I became interested in studying water resources engineering and chose it as my major. I want to thank all of them for providing me with an amazing work environment. I owe many thanks to my supervisor and team leader, Hannes Björninen, for providing patient guidance and feedback throughout the project and developing my skills for my future career path. Warm thanks to Markus, Pedro, Lauri, Mika and Janne for answering all my questions and helping me before and during this study with such enthusiasm.

I am grateful to HSY for this unique opportunity and for providing funding for this thesis project. Special thanks to Leena Sänkiaho for facilitating the data transfer and answering my question about the Helsinki sewer network.

I would like to express my sincere appreciation to my supervisors at Tampere University, Niko Niemelä and Annina Takala, for their advice and lessons during this work.

Lastly, I want to thank my family, especially my mother, for helping and encouraging me in any way they could throughout my studies.

Tampere, Dec 2022

Moein Ashofteh Beyraki

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS AND ABBREVIATIONS

CS	Combined sewer
CSO	Combined sewer overflow
CSS	Combined sewer system
DEM	Digital elevation model
DTM	Digital terrain model
DWF	Dry-weather inflow
EPA	United States Environmental Protection Agency
GIS	Geographic Information System
HSY	Helsinki Region Environmental Services Authority (In Finnish: Helsingin seudun ympäristöpalvelut)
LiDAR	Light Detection and Ranging
NSE	Nash-Sutcliffe efficiency
PFE	Peak flow error
RF	Random forests
RMSE	Root-mean-square error
SSS	Separate sewer system
SVM	Support vector machines
TIA	Total impervious area
WWF	Wet-weather inflow
WWTP	Wastewater treatment plant

A	Area [m²]
A	Flow cross-sectional area [m²]
A	Sub-catchment area [m²]
A_{sn}	Node's storage surface area [m²]
$\sum A_{sl}$	Sum of the surface area contributed by the links connected to the Node [m²]
d	Depth of ponding [m]
d_s	Depression storage [m]
ET	Evapotranspiration [mm/d]
e	Surface evaporation rate [m/s]
f	Infiltration rate [m/s]
G_{in}	Ground-water inflow [mmd]
G_{out}	Ground-water outflow [mm/d]
g	Acceleration of gravity [m/s²]
H	Hydraulic head of water in the conduit (Z + Y) [m]
i	Precipitation rate [m/s]
k	Dimensionless coefficient for width calculation
n	Manning's roughness coefficient [-]
L_{avg}	Average maximum overland flow path [m]
L_{pipe}	Length of the main drainage channel
P	Precipitation [mm/d]
$\sum Q$	Net flow into the node-link (inflow – outflow)
Q	Flow rate [m³/s]
Q	Stream outflow [mm/d]
Q_{sim}	Simulated discharge [m³/s]
Q_{obs}	Observed discharge [m³/s]
$\bar{Q}_{obs}$	Average observed discharge [m³/s]
q	Runoff rate [m/s]
S	Slope [m/m]
S_f	Friction slope (head loss per unit length)

t	Time [s]
W	Flow width [m]
x	Distance [m]
Y	Conduit water depth [m]
Z	Conduit invert elevation [m]
ΔS	Change of (liquid and solid) water storages [mm/d]
α	Dimensionless coefficient for width calculation

1. INTRODUCTION

Stormwater is surface water that originates from rainfall or melting snow and ice. Flash flooding from stormwater in heavy and sudden rainfall, is among the most frequently occurring natural disasters (UNISDR, 2015). Urban areas are particularly prone to floods, resulting in life losses and economic damage (Y. Chen et al., 2015). Besides, because of climate change and the frequency of extreme climate events, these consequences are anticipated to be intensified in many locations in the future (Mirza, 2003). Thus, urbanization, growing population, and climate change are three global trends pushing the importance of urban stormwater management in urban areas (Semadeni-Davies et al., 2008). One of the aspects of stormwater management is to decrease the overflow frequency in the combined sewer systems (CSS) in which sanitary and stormwater are collected and transported together (Fu et al., 2019). CSS encounter significant fluctuation, including high flows during storm events (Bareš et al., 2008). Sufficient precipitation in CSS can lead to combined sewer overflow (CSO), resulting in untreated sewage discharges into receiving water bodies and posing serious challenges to maintaining water quality criteria (Fu et al., 2019; Garofalo et al., 2017).

Urban development has induced land cover modifications and impervious surface expansion in cities (Dou & Kuang, 2020). This has led to more runoff generation and frequent overflows in cities' sanitary and combined sewer networks, often due to severe storms exceeding the network's capacity (Fletcher et al., 2013). To address this issue and prevent overflows, designers must analyze, forecast, and manage peak flows, which can be estimated only by using modeling on a large scale (Kong et al., 2017; Radinja et al., 2019). Physical-based hydrological models like the Stormwater Management Model (SWMM) are the efficient and cost-effective way to test the reliability and performance of stormwater systems (Autixier et al., 2014). In SWMM-based models, the study area must be divided into sub-catchments linked to each other and the drainage network. The shapes, areas, outlets, and parameters of the sub-catchments will affect the accuracy of the urban surface runoff and its concentration times (Rossman, 2016). As a result, a precise characterization of sub-catchments is critical for urban rainfall-runoff models.

However, the delineation and parametrization of accurate sub-catchments for large urban areas are tedious and complex processes (Niemi et al., 2019). Thus, building

hydrological models is time-consuming and prone to designer's errors. Manual catchment delineation and parametrization challenges highlight the necessity for automated methods (Krebs et al., 2014). Different methods with varying techniques, assumptions, type of data, and spatial distribution have been proposed by academia and commercial software developers. All of these methods integrate (Geographical Information System) GIS technologies in hydrological modeling to ease the development of hydrological models in urban environments but have not been tested or compared on a large scale.

1.1 Goals, research questions and scope

This thesis explores the hydrology setup and parameterization for the extensive urban combined sewer network (CSN) with SWMM. The goal of the research is: "Explore the best methodology for semi-automated or automated catchment delineation and parametrization with efficient and accurate results". To achieve this, different methods are tested on small areas, and an automated method with more features and better hydraulic results is adopted for updating the whole Helsinki combined sewer network. The following three research questions (RQ) sum up the goals of this thesis:

RQ1: Which method is more robust, automated, and systematic and leads to good enough hydraulic results, compared to the measurement?

- How much manual work is still needed for using each method and to process the initial data required?
- How do different data like the new automated imperviousness layer impact hydrological modeling?

RQ2: How do the increased level of catchment detail and spatial resolution affect simulation performance?

- Which of the methods enables the use of varying spatial resolution?
- What level of detail results in good enough accuracy while being efficient (considering model uncertainties)?

RQ3: How to best apply the automatic catchment delineation and parametrization on a large scale?

- Test selected methodology for the entire Helsinki combined sewer (CS) (approx. 17 km²)
- What are the most critical factors in the scaling-up process? Where are the bottlenecks?

- Recommendations for full implementation of the Helsinki CS catchment update using automatic processes.

In addition, the study scope is limited to investigating catchment delineation and parametrization approaches for SWMM modeling. Model calibration is not included in the study objectives.

1.2 Structure of the thesis

Chapters 2 and 3 provide background on the fundamentals of hydrology and modeling concepts for this study. Chapter 2 contains fundamental information on the hydrological water cycle and sewage systems in urban areas. Chapter 3 introduces the principles of rainfall-runoff modeling and discusses key points and challenges in building the SWMM model, as well as the effects of sub-catchment parameters on the results.

In Chapter 4, the study's methodology is presented by depicting the research process and demonstrating the tools used for each step of this study. Then, study sites, data sources, and catchment definition methods are described. This chapter will concern how the sub-catchments are generated and given the parameters for case areas and the city-scale model.

Chapter 5 presents the statistical and hydraulic results of models with different sub-catchment definitions. First, results and findings in case areas were discussed, and catchment definition methods were compared. Then, the results of the scale-up process were also covered in the same chapter. Finally, the study's findings and identification of worthwhile topics for future investigations will be summarized in Chapter 6.

2. URBAN HYDROLOGY

2.1 Regional water balance and catchment

A catchment (also known as a drainage basin and watershed) serves as the fundamental regional unit in the majority of hydrological analyses and modeling (Bren, 2015). It is defined as a region that contributes to the stream flows at a specific cross-section of a watercourse or a specific inlet to the stormwater network (Dingman, 2015). Lack of a topographic map, catchment size, and large flat areas are among the difficulties in catchment definition (Bren, 2015).

Catchments can be defined based on topography and the division is the boundary along which water may flow to either of two distinct catchments. In the past, catchments have been defined using topographic maps with elevation contours (Dingman, 2015). Digital elevation models (DEMs) and digital terrain models (DTMs), however, have recently taken over as the primary data source using GIS systems (Bren, 2015). The main difference between these two models is that DEM typically takes into account all persistent objects on the surface, such as buildings, trees, and other artifacts, whereas the DTM shows the bare earth surface (Podobnikar et al., 2000).

Each catchment is a system (see Figure 1), with a control volume governed by the regional water-balance equation:

$$P + G_{in} - (Q + ET + G_{out}) = \Delta S \quad (1)$$

Where:

P = precipitation [mm/d],

G_{in} = ground-water inflow [mm/d],

Q = stream outflow [mm/d],

ET = evapotranspiration [mm/d],

G_{out} = ground-water outflow [mm/d],

ΔS = change of (liquid and solid) water storages [mm/d].

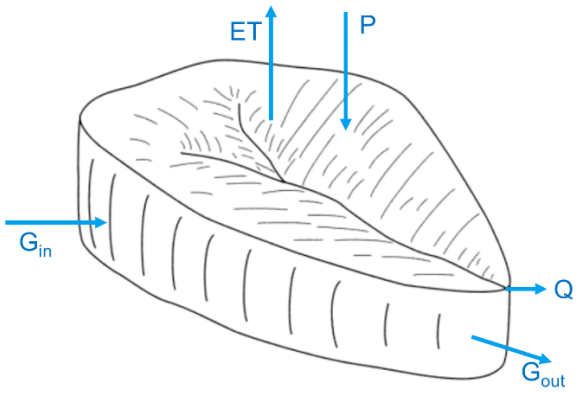


Figure 1. *Regional water balance in conceptual catchment model.*

Due to the conservation of water volume, the change in storage can be roughly calculated as zero over a long time. Evapotranspiration considers all actions that take place close to land or water surfaces that cause liquid water to evaporate into water vapor. It includes both the evaporation of water from the ground's surface and plants' vascular systems into the atmosphere. During the second process, plants take in water from the soil, evaporating it via tiny holes in their leaves. This loss is influenced by the vegetation type, timing, and kind of precipitation, among other things, although it is frequently not as significant. (Dingman, 2015)

2.2 Water balance alteration due to urbanization

Urbanization dramatically impacts a catchment's hydrological functioning by altering the natural water pathways, infiltration, and evaporation, as shown in Figure 2 (Dow & DeWalle, 2000). Although the impacts vary, the majority of the water-balance alterations are associated with replacing natural surfaces with impervious surfaces and artificial flow pathways in urban catchments, which disturb the natural hydrological cycle and reduce infiltration and water storage capacity (Fletcher et al., 2013; Sheng & Wilson, 2009). Most impervious surfaces are artificial surfaces, including roofs, pavements on roads, driveways, parking lots, playfields, sidewalks, etc. At the same time, many urban natural surfaces and non-sealed pavements are also highly compacted and have an impervious quality, which lowers their ability for infiltration. Furthermore, in cities, not only the soil surface layer has been altered and compacted. Generally, engineered soils have replaced thick layers of natural soils with compacted soil or other construction materials, changing the sub-surface infiltration characteristics in most areas.

Furthermore, in cities, not only the soil surface layer has been altered and compacted. Generally, engineered soils have replaced thick layers of natural soils with compacted

soil or other construction materials, changing the sub-surface infiltration characteristics in most areas.

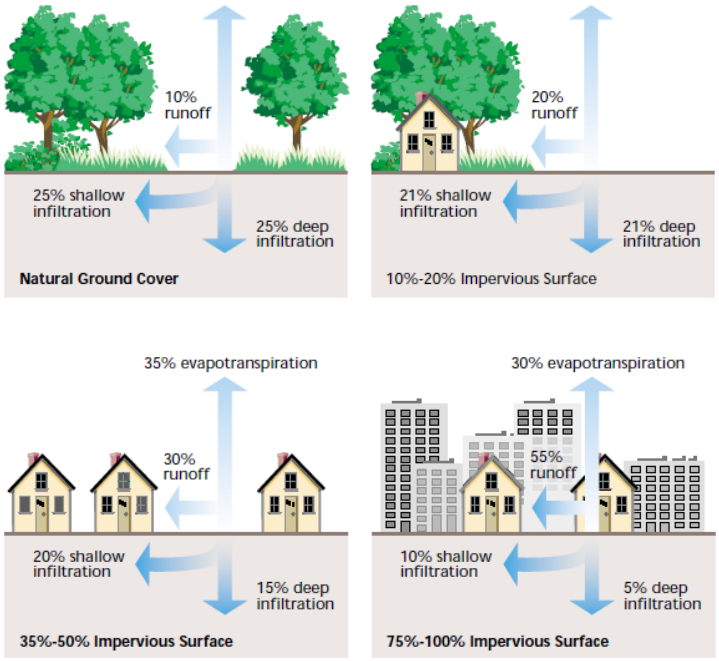


Figure 2. *Urbanization impacts on water cycle* (Bernard & Tuttle, 1998).

It is noteworthy that urban runoff is another term for stormwater from urban areas, and here, unless otherwise stated, the term stormwater is used to refer to urban runoff. When impervious surfaces do not allow rainfall to infiltrate, stormwater runs off until it finds its way to the drainage system, evaporates, or drains onto a pervious area where runoff can infiltrate into the soil.

Additionally, most impervious surfaces convey water more efficiently than naturally pervious surfaces, which enables runoff to flow quickly to stormwater inlets, channels, and low-lying areas (Shuster et al., 2005). Therefore, rainfall's runoff process becomes faster, causing higher peak flows and shorter recession and concentration time which leads to an increase in the total runoff volume (Fletcher et al., 2013; Shuster et al., 2005). Consequently, if stormwater volume and flow are high, exceeding the discharging capacity of the catchment, flooding may occur at the catchment outlet or in the conveyance system.

2.3 Combined sewer network and overflows

Generally, urban sewage network wastewater is divided into dry-weather inflow (DWF) and wet-weather inflow (WWF). When there is no precipitation or snowmelt, the DWF is defined as the regularly observed flow that contains effluent from human developments. Like freshwater consumption, the DWF commonly uses a cyclical rhythm. On the other

hand, the stormwater flows that occur during and after a rainstorm or snowmelt event are represented by the WWF. Stormwater is gathered and transported from urban areas through an underground sewer network. When stormwater is collected, it is either conveyed via a separate sewer system (SSS) (Figure 3(a)), where sanitary sewer and stormwater flow in separated pipes, or a combined sewer system (CSS), where sanitary sewer and stormwater flow in the same pipes, shown in Figure 3(b) (Al Aukidy & Verlicchi, 2017). Both sewage systems listed above have benefits and drawbacks.

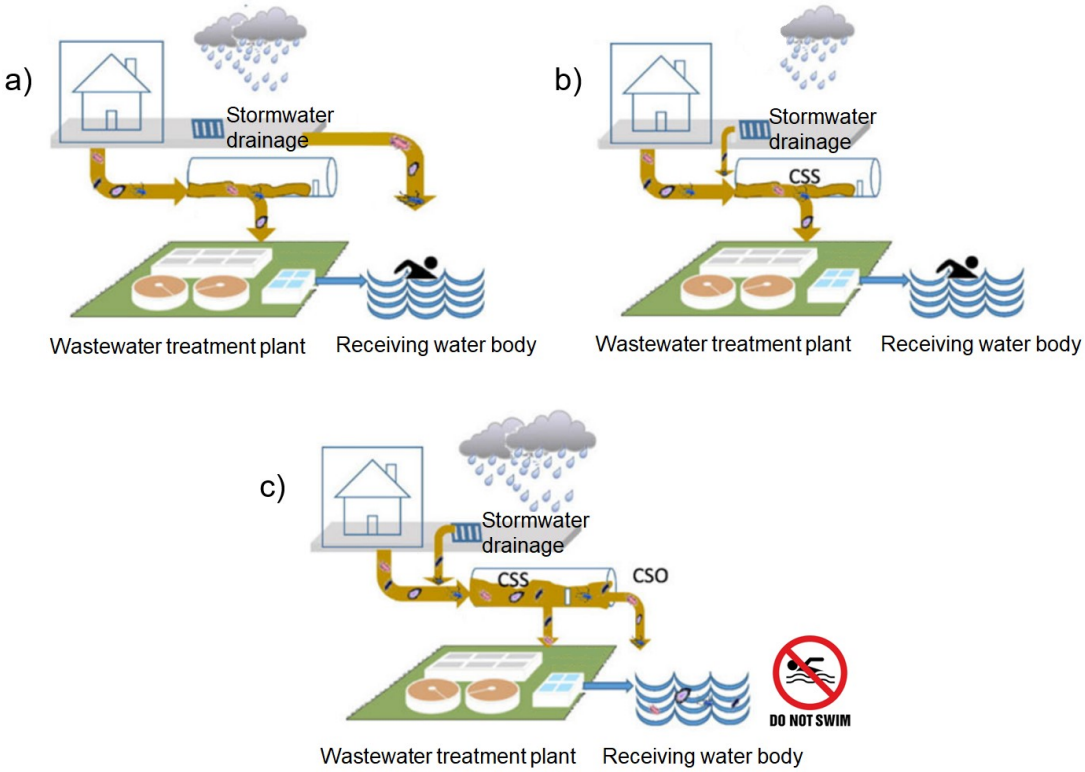


Figure 3. *Overview of the urban sewer network: (a) separate sewer system; (b) combined sewer system; and (c) combined sewer overflow (adapted from Al Aukidy & Verlicchi, 2017 with permission from Elsevier).*

The elder of the two systems, CSS, is widely used, for instance, in the center of old European cities where drainage systems were constructed in the 20th century (Semadeni-Davies et al., 2008). CSS was used in Helsinki, the capital of Finland, before the 1970s. The primary benefit of CSS is transporting rainwater and wastewater to a treatment plant, where they are treated before being discharged into the environment (Thorndahl et al., 2015). However, CSS is a challenging system to operate in treatment plants. In CSS, inflows to a treatment plant vary in quantity, quality, and temperature, making chemical and biological treatment processes difficult to control.

Another important drawback of CSS is combined sewer overflows (CSOs) (Figure 3(c)) which occur when the capacity of the sewers or the treatment plant is surpassed during

periods of high rainfall or snowfall. When the sewer network is filled up during CSOs, the excess sewage spills out via overflow structures directly into receiving waterbodies to prevent the backflow of sewage and stormwater into buildings (Tibbetts, 2005). CSOs contain all typical harmful substances of stormwater and untreated domestic and industrial sewage that could directly contaminate receiving waters (Montserrat et al., 2015).

Due to the widespread concern about contaminated water, stormwater and sanitary sewer are governed by particular legislation in all developed countries (De Feo et al., 2014). For example, in Europe, the strict requirements of the EU environmental regulations require members to enhance their wastewater treatment facilities. Furthermore, in developed countries, stormwater and its overflows are recognized as the primary cause of contamination for urban receiving water bodies (Kostarelos et al., 2011).

Development of the SSS started because of the harmful impacts of CSOs. In the SSS, the stormwater is routed straight to a nearby receiving water body, while sanitary sewer is led and treated in WWTP (Thorndahl et al., 2015). This can enhance the water quality of nearby water bodies as less sanitary effluent is released into the receiving waterways. In Finland, in the capital, Helsinki, the first SSS was constructed in 1938. From then on, the SSS was used in most newly developed areas across the country (Sillanpää, 2013). Compared to CSS, in the SSS, the treatment process in the treatment facilities is more efficient since the conditions are more uniform and the inflow volume is lower. Although SSSs lower the possibility of CSOs, converting a CSS to a SSS is neither easy to implement in short-term nor cost-effective in old areas as it requires additional planning and extra pipes (Ahm et al., 2016).

2.4 Stormwater management

There are three common types of flooding, pluvial floods, fluvial floods, and coastal floods (Anees et al., 2020). Fluvial floods occur when the water level in rivers and streams rises, and intense windstorms from seawater or tsunamis cause coastal floods. Urban flooding, also known as pluvial flooding, describes inundation during heavy rainfall that causes runoff to surpass the drainage system's capacity and, consequently, accumulation of water on the built-up surface for a long time. Due to urban expansion and the rapid increase in impervious surfaces and urban expansion, flooding has become more common in urban settings (Sheng & Wilson, 2009). In addition, the likelihood of urban flooding is projected to rise because of climate change (Poff, 2002).

Hydrological events, such as rainfall or river flow, have occurrence probability, which is fundamental to designing flood estimation. In the past, "recurrence interval" and "return

period" were used to express the estimated average time between two same-size events, which resulted in misinterpretation. The probability of urban flooding corresponded to the rainfall event. Land cover changes, the state of the drainage channels, and other factors can affect urban flooding probability. The probability of urban flooding for urban areas vulnerable to river or coastal flooding depends on both the rainfall event and river stage (water level).

Forecasting peak flows, runoff volumes, and discharge hydrographs are among water resource engineers' primary tasks (Bedient et al., 2008). In urban areas, both the infrastructure and the people are at risk from flood flows and its repercussions might affect the environmental, economic, and social aspects of all societies. Sustainable stormwater management practices must be widely adopted to preserve the built environment and ensure public health.

In CSS, the quantity and quality of CSOs vary significantly based on precipitation. As CSOs are heavily polluted by heavy metals, organic waste and nutrients (Al Aukidy & Verlicchi, 2017), the sustainable management of CSOs is critical. Therefore, it is essential to have reliable stormwater systems to reduce the risk of CSOs and the environmental damage they cause.

Stormwater system designs vary based on their purpose and the country where they were first developed. This method for handling stormwater considers both future requirements and the preservation of natural resources (Hvitved-Jacobsen et al., 2010). The management strategies can be structural i.e., based on built systems, like rainwater detention basins, porous pavement, green roofs, low-impact development, and vegetated swales, or non-structural, like pollution control, political decisions, or street sweeping (Barbosa et al., 2012; C. Martin et al., 2007).

Holding runoff inside the catchment before being collected by the drainage network mitigates flood risks by reducing the runoff peaks reaching the catchment outlet and distributing the runoff volume over a longer period of time. Stormwater detention and retention ponds can be used the same way and they can delay the catchment runoff and reduce stormwater volume especially if they allow infiltration of stormwater into the ground. The detention ponds only contain water if there are precipitations. In contrast, retention ponds always have water inside, and the water level is up during storm events.

Furthermore, pervious pavements, grassed swales, and other low-impact development techniques can attenuate and decrease the runoff peak using infiltration and evaporation processes (Dietz, 2007). Moreover, such techniques and practices are essential for improving the quality of urban water bodies and reducing construction expenses in urban

development (Liu et al., 2021). Green roof construction is also popular, although its peak flow decrease is lower than other management techniques (Lee et al., 2010).

All means to attenuate stormwater flows mitigate the risk of heavy rainfall events overloading the drainage system (pipes, pumps, etc.). The best outcome is often obtained when various management techniques are combined and dispersed across the watershed (Lehtinen, 2014).

3. RAINFALL-RUNOFF MODELING USING SWMM

3.1 Rainfall-runoff modeling

Modeling is the process of simulating the natural world using a model representing a part of that world. Rainfall-runoff modeling simulates the hydrological response (runoff) to a particular input (precipitation) over a period of time. The urban rainfall-runoff model describes rainfall-runoff processes and provides data for urban flood disaster prevention (Ji & Qiuwen, 2015). This can be challenging due to the highly nonlinear processes, complicated relationships, and great spatial variability involved (Dingman, 2015).

There are different prediction mechanisms but mostly fall between these two poles, system view and physically-based models. In the system view (also known as black-box or conceptual models), the method looks for an abstract function connecting input and output functions. It has some significant disadvantages because of its simplicity. It may successfully work in small catchments (e.g., less than 80 ha) while is unable to take into consideration any storage or detention ponds. (Dingman, 2015)

In contrast, physically-based models provide a detailed description of the physical processes that are occurring underneath such as groundwater flow, evapotranspiration, infiltration, snowmelt, and overland flow. In physically-based modeling, the hydrological process of water movement is either approximated by empirical equations or finite difference approximation of the partial differential equation describing the mass, momentum, and energy balance (Mbajiorgu, 1995). Therefore, it is an efficient and cost-effective way to test the reliability and performance of urban stormwater systems (Salvadore et al., 2015).

3.2 Data acquisition for rainfall-runoff modeling

Developing rainfall-runoff models for large urban regions requires the use of suitably accurate spatial databases (Bach et al., 2013). In many countries, there are high-quality public open databases that include data, such as accurate DEM, land cover data sets and meteorological time series. In the best cases, temporal and geographic high-resolution hydrological data from real-time sensors are also available. Such data is rapidly improving and better available in the public domain. For instance, the Finnish government has made an online portal publicly (avoindata.fi) that aggregates open national databases under a single domain (Warsta et al., 2017).

Remote sensing and machine learning

Rapid urban development and associated surface elevation and land cover change highlight the need to understand ground surfaces accurately. Thus, there is a high need for generating such spatial databases for hydrological analysis and urban planning.

In recent decades, satellite remote sensing has become a popular and helpful method for providing various data. Remote sensing allows for the quick capture of information on the land cover at a fraction of the cost of other approaches such as ground surveys (Z. Chen & Wang, 2010). The advantages of satellite images for mapping land cover are their multi-temporal availability and spatial coverage (Talukdar et al., 2020). Thus, together, GIS and remote sensing data can be used to estimate the different parameters of an urban catchment that SWMM models need for simulating stormwater runoff (Jain et al., 2016). Even though more precise mapping remains challenging because of urban surface diversity, this has opened up significant potential for incorporating such data into urban hydrological assessments.

Among novel methods and algorithms, machine learning and computer vision have been widely used recently for detecting land use and land cover types using urban high-resolution remote sensing images. K-means, Deep learning, support vector machines (SVM), and random forests (RF) are all among the machine learning models that have been shown to be beneficial for remote sensing image categorization owing to their ability to learn features from large datasets automatically (P. Zhang et al., 2018). Deep learning has attracted international scientific attention due to its ability to evaluate large amounts of data for modeling a computer learning process from observations (P. Zhang et al., 2018). It has become the best-practice method for classifying remote sensing data and image classifications (H. Zhang et al., 2019). The usage of deep learning for different applications, such as automatically extracting features for categorizing land cover data like roads, building footprints, and grasslands, has mostly helped the remote sensing community (Mnih, 2013).

Furthermore, Deep learning has made remarkable progress in pattern recognition which opens up a lot of room for advancement in automated impervious area mapping (Huang et al., 2019). Convolutional deep neural networks, in particular, have been effectively used to image categorization because they integrate the context of a single-pixel prediction over many scales, using the information in the neighborhood of the predicted pixel and resulting in increased accuracy (P. Zhang et al., 2018). Numerous researchers and engineers have successfully adopted this strategy (e.g., Huang et al., 2019; Pan et al., 2020; H. Zhang et al., 2019).

In Finland Helsinki Region Environmental Services Authority (HSY) has created a precise open land cover dataset for the Helsinki region (HSY, 2017). The data was gathered using infrared aerial orthoimages and LiDAR (Light Detection and Ranging), a remote sensing method used to examine the Earth's surface and measure the ranges. The European Environment Agency has compiled the Urban Atlas data collection, which includes pan-European land cover information for urban zones with more than 50 000 inhabitants (EEA, 20177). SCALGO, for example, created a "Convolutional Neural Network" model based on the UNET method for mapping imperviousness in the Helsinki metropolitan area. The research evaluates the use of newly acquired data from the National Laser Scanning and Aerial Imaging Program for water management (LaserVesi Project, 2022).

Radar precipitation measurement

Precipitation is the primary component driving rainfall-runoff simulation (Rossman, 2016). To assess and model a catchment, precise precipitation observations are essential. Nevertheless, the uncertainties in the rainfall estimates can account for a significant portion of the modeling errors (Moulin et al., 2009). Measuring precipitation and methods for generalizing measured data to the studied area add uncertainty to modeling (Dingman, 2015).

Even though rain gages have been used as a source of precipitation data for centuries, they still continue to serve as one of the primary sources of rainfall data (Kidd et al., 2017). Automatic recording gauges can record the quantity of precipitation with a consistently higher temporal resolution (Niemi, 2017). Despite having an apparently straightforward method of action, rain gages are prone to some degree of inaccuracy due to the errors caused by wind, gage calibration flaws, and evaporation losses (Humphrey et al., 1997). The biggest issue with rain gage recordings is related to their areal representativeness, even if rain gages still provide the most precise measurements of surface rainfall volume and intensities at a particular location (Niemi, 2017). Since precipitation is known to vary greatly across both time and space, hydrologists prefer to estimate the areal precipitation for catchment studies instead of using point measurements (Peleg et al., 2013).

Weather radars, which have been actively developed since the end of World War II, area another important tool for measuring precipitation. The primary benefit of employing weather radar is its ability to deliver spatially continuous rainfall estimation at short temporal sampling intervals (5 min) and at a relatively high spatial resolution (Seo et al., 2010). Potential useful information can be obtained from radar rainfall data when high-

resolution rainfall data is required in urban hydrology, like urban flooding (Smith et al., 2001; Thorndahl et al., 2017). There are, however, errors in radar measurements of precipitation. This is due to the fact that weather radar measures precipitation indirectly by estimating how strong the signal is reflected by water droplets. Conventional observation is still necessary for validating radar data against ground observations (Dingman, 2015). Consequently, estimates from weather radars with extensive spatiotemporal coverage can be modified using more precise point measurements from rain gages to provide more accurate precipitation estimates than those obtained solely from rain gages or weather radars (e.g., Seo et al., 2010). Furthermore, it is important to note that the requirement to preprocess and validate radar data with massive datasets slows down modeling performance and restricts the method's adoption in the typical stormwater modeling workflow.

Dual-polarization radar rainfall estimations can help reduce inaccuracies and provide more precise snow observations compared to conventional single-polarization radars (Cifelli & Chandrasekar, 2010). As of June 2017, all ten of Finland's national weather radars can operate in dual polarization. However, the operational radar product has not yet made much use of the radars' dual-polarization capabilities (Gregow et al., 2017).

Water level and flow measurements

The environment in the sewer network is corrosive and humid. Besides, the water level varies at different times of the year and is usually carried in non-full open channel pipes. Therefore, it is challenging to measure the accurate flow inside the pipes while it is necessary for evaluating the hydrological behavior of the catchment (Quevauviller et al., 2007). While measuring the water level is often cheaper and installing the meters faster, flow is more often measured instead of the water level. The reason is that flow is more beneficial for understanding and monitoring the operation of the network, especially for analyzing the quantity and quality of drainage system discharges to receiving water bodies (Ahm et al., 2016). Ultrasonic and electromagnetic measurements are the most commonly possible methods for partially filled conduits (Godley, 2002).

An ultrasonic meter consists of two sensors that are installed outside the pipe and uses the Doppler effect for measuring water velocity assuming that the speed of sound in liquid is constant (Godley, 2002). The ultrasonic meter sends ultrasonic waves into the pipe at a certain angle, both upstream and downstream, and measures the time it takes for the waves to travel back to the meter. Ultra-sonic meters are accurate but sensitive to impurities in the water and variations in flow characteristics. They work better with

pipes that are full of water, sufficiently straight, and have a high flow rate. Ultrasonic meters are easily installed and can be portable (Ren et al., 2022).

Electromagnetic flow measurement is based on the electrical conductivity of the liquid and Faraday's law (Li et al., 2020). The water acts as a moving conductor in the pipe, into which a voltage is caused by the electric field, proportional to the flow rate (Doney, 1999). Then, the induced voltage can determine the water average velocity in the section. The accuracy of electromagnetic flow measurement is widely known for pressurized pipe and has been used for more than four decades (Quevauviller et al., 2007). Electromagnetic flow measurement is widely used in Finland at new or renovated wastewater pumping stations. Electromagnetic flowmeters have accurate measurements for partially filled open water but needed to be calibrated before as they require an undisturbed pipe section and a sufficiently high flow rate (Quevauviller et al., 2007). High accuracy, unaffected by high solid debris, flow range coverage, and no contact with water are among its advantages and demanding installation and high price were reported as its disadvantages (Doney, 1999; Quevauviller et al., 2007).

3.3 The EPA Storm Water Management Model (SWMM)

Several software programs are available to simulate physically-based rainfall-runoff models. The EPA Storm Water Management Model (SWMM) is among the most popular models (Madrazo-Uribeetxebarria et al., 2021; Shahed Behrouz et al., 2020). SWMM is free, open-source software provided by the United States Environmental Protection Agency (EPA) with a graphical user interface first developed in 1971. It is mostly utilized to simulate stormwater quantity and quality processes in urban areas (Peterson & Wicks, 2006).

SWMM source code is in the public domain and has been upgraded and re-written during the development process. It is used as a core engine and packaged with user-friendly interfaces in commercial software, e.g., Fluidit. It can model both single and multiple precipitation events (Rossman, 2016). The simulation period is divided into numerous time steps (typically minutes or hours), and SWMM can monitor the quantity and quality of runoff for each of these time steps. In urban areas, SWMM is commonly employed in the research and design of stormwater drainage systems. It has well-documented user manuals, and several previous case studies used SWMM as the core engine.

3.4 Governing equations in SWMM

Catchment hydrology

SWMM requires a detailed characterization of catchment land use, overland flow, and channelized flow paths to estimate the quantity of catchment runoff and create discharge hydrographs (Jain et al., 2016). Since the hydrological processes on the catchment surfaces are essential to rainfall-runoff modeling, the land surface must be divided into sub-catchments (O'Loughlin et al., 1996). Sub-catchments vary in characteristics and can receive water input from precipitation or as runoff from adjacent sub-catchments (Rossman, 2016). Stormwater can infiltrate, evaporate, and turn into a runoff in the catchment depending on the specific characteristics of the catchment. In SWMM, each sub-catchment must be linked to a *Rain Gage* component containing rainfall intensity or volume data at particular time intervals (Rossman, 2016). Each sub-catchment has a single discharge point in the model.

SWMM characterizes each sub-catchment through a set of geometrical and hydrological parameters such as imperviousness, Manning's roughness coefficient, and depression storage for calculating the depth of water and runoff's volumetric flow rate. SWMM engine calculates the runoff rate (q) from sub-catchments as follows:

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

Where:

A = area [m²],

W = flow width [m],

S = slope [m/m],

n = Manning's roughness coefficient [-],

d = depth of ponding [m],

d_s = depression storage [m].

A sub-catchment area may have both pervious and impervious surfaces. SWMM is given the different parameters for each portion and calculates the runoff rate separately. Runoff from these portions may flow over the other portion or reach the drainage system directly. (Rossman, 2016)

In addition, the sub-catchments are regarded as nonlinear reservoirs, which receive precipitation and produce runoff and pollutant hydrographs. The non-linear model takes into

account hydrological losses, such as evapotranspiration, infiltration and initial abstraction by depressions, as indicated in Figure 4, (Rossman, 2016).

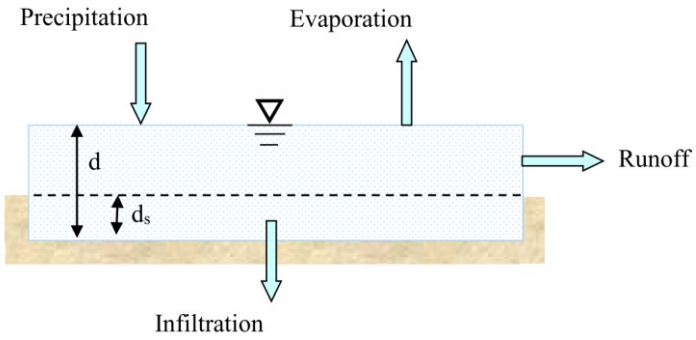


Figure 4. *Non-linear reservoir model (Rossman, 2016).*

According to the principle of conservation of mass, the difference between the input and outflow rates throughout the sub-catchment determines the net change in water depth d per unit of time t:

$$\frac{\partial d}{\partial t} = \mathrm{i} - \mathrm{e} - \mathrm{f} - \mathrm{q} \tag{3}$$

Where:

i = precipitation rate [m/s],

e = surface evaporation rate [m/s],

f = infiltration rate [m/s],

q = runoff rate [m/s].

The input of water onto saturated pervious or impervious surfaces causes ponding and surface runoff takes place if the ponding depth (d) grows higher than the depression storage capacity of the ground (d_s). Therefore, the surface tension's capacity to keep the water still is exceeded, and runoff flow over the area, as shown in Figure 4. However, not all the precipitations result in a runoff. Effective precipitation is the part of the rainfall that exceeds the loss rates on the catchment and creates runoff immediately or shortly after the event (Dingman, 2015).

Evaporation occurs in the sub-catchments if there is standing water on the surfaces. The user can define evaporation rates from air temperatures. Furthermore, other parameters such as the site's latitude and wind speed can affect the magnitude of evaporation.

For detailed modeling of the groundwater conditions, it is possible to incorporate optional *Aquifer* objects. Water can be exchanged between the drainage system and *Aquifers* via pipe leakages. When temperatures are low, precipitation has a chance of being collected in the snowpack object, which can model sub-catchment snow accumulation and melt (Rossman, 2016).

The infiltration rate is the velocity at which water enters the soil from the pervious parts of the catchments (Dingman, 2015). Impervious surfaces eliminate infiltration in urban areas, increasing surface runoff (Fletcher et al., 2013). Lower infiltration rates hinder groundwater recharge and decrease groundwater levels and stream base flows. Consequently, decreased groundwater levels could make it more challenging to use groundwater and sustain a healthy environment. There are three different built-in infiltration methods are available in SWMM including the Curve Number, the Green-and-Ampt, and Horton's equation (Heber Green & Ampt, 1911; Horton, 1933).

Basic hydraulic concepts

In urban areas runoff is typically captured and conveyed with an underground combined sewer or stormwater system. To simulate how catchment runoff is collected by drainage systems and transported to the discharge point by stormwater management infrastructure, the hydraulic network must be studied. Hydraulic network determines how runoff and potential external inflows are hydraulically routed through a system of pipes and channels called conduits. In SWMM, conduits are connected at nodes that can represent junctions, stormwater inlets, storage units or outfalls. Common conduit parameters are, length, slope, invert elevations, Manning's roughness coefficient n, and cross-sectional geometry. Similarly, the junction nodes have parameters, namely the invert elevation, the depth from the ground surface and junction diameter or volume.

To compute the flow in the conduits for each computational time step, SWMM provides users with the choice to use either a steady flow or an unstable flow routing method. In SWMM, the hydraulics of unsteady flow is determined using the Saint-Venant equations, a pair of partial differential equations for conservations of mass and momentum (Rossman, 2017). SWMM offers two primary alternative numerical solution methods for solving Saint-Venant equations: dynamic wave or kinematic wave analysis. Dynamic wave analysis provides the most theoretically correct solutions since it solves these equations in their complete form (Rossman, 2017). It can consider channel storage, backwater impacts, entrance and exit losses, flow reversal, and pressured flow. Any general network structure, even with diversions and loops, may use dynamic wave analysis since it calculates both water levels at nodes and flow in conduits. It is the method

of choice for networks susceptible to substantial backwater and flows regulation via weirs and orifices (Rossman, 2017). Results accuracy and simulation time are influenced by routing method selection (Rossman, 2017).

The conservation of mass and momentum for gradually varied unsteady free surface flow can be represented as:

$$\frac{\partial A}{\partial t} + \frac{\partial Q}{\partial x} = 0 \tag{4}$$

$$\frac{\partial Q}{\partial t} + \frac{\partial (Q^2/A)}{\partial x} + gA\frac{\partial H}{\partial x} + gAS_f = 0 \tag{5}$$

Where:

A = flow cross-sectional area [m²],

Q = flow rate [m³/s],

t = time [s],

x = distance [m],

H = hydraulic head of water in the conduit [m],

Z = conduit invert elevation [m],

Y = conduit water depth [m],

S_f = friction slope [m/m],

g = acceleration of gravity [m/s²].

Equations (4) and (5) are utilized to determine the discharge (Q) in the conduits. The nodes' head values (H) are calculated based on flow conservation. The nodes' head values (H) are calculated based on flow conservation. The principle of flow conservation is dictated in each assembly, consisting of the node itself and half the length of each connecting link. The changes in volume with respect to time must be equal to the difference between the flow entering the node and the flow exiting the node. This formula computes the flow Q in the conduits and the head H at the nodes in each time step. (Rossman, 2017)

$$\frac{\partial H}{\partial t} = \frac{\sum Q}{A_{sn} + \sum A_{sl}} \tag{6}$$

Where:

$\sum Q$ = net flow into the node-link (inflow – outflow),

A_{sn} = node's storage surface area [m²],

$\sum A_{sl}$ = sum of the surface area contributed by the links connected to the node [m²].

Moreover, flows exceeding the drainage system's capacity may result in ponding. This enables the temporary storage of extra water at a particular junction and allows this pond to dry out when system capacity is once again available. Ponding may also be disabled by the user, resulting in the system overflowing and losing all extra volume of water from the model (Rossman, 2017).

3.5 SWMM catchment definition for large urban area

In SWMM-based models, the main watersheds (catchment) must be divided into detailed sub-catchments with distinct characteristics which are connected to one another or to the underlying stormwater network. The definition of the sub-catchments determines the simulation rainfall-runoff process and affects the accuracy of the runoff simulation and flow concentration times. Thus, SWMM requires a detailed and correct determination of its sub-catchment parameters for characterizing urban catchment runoff.

SWMM sub-catchments are often made up of a variety of land cover types. As a consequence, sub-catchment parameters are unique, demonstrating the spatially heterogeneous hydrographic properties of the urban area (Dongquan et al., 2009). The shapes, outlets, and parameters of the sub-catchments determine the simulation rainfall-runoff process and affect the accuracy of the runoff simulation and flow concentration times. Based on the surface properties, all sub-catchment parameters can be categorized into physical (geometry) and hydrological parameters. The physical parameters cover the sub-catchment's width, slope, outlet as well as the shape and area, while the hydrological parameters include the imperviousness, depression storage, and surface. Therefore, an infinite number of different sets of parameters can be used to determine the overall system and replicate the catchment response. The challenge for the catchment modeler is to choose a set of parameter values that are appropriate at the local scale while remaining suitable for scaling up to the system level. For this, multiple sources of information on the urban catchment are required or helpful to increase the model's accuracy. However, detailed data regarding the urban surface area is frequently limited, which challenges parameter estimation.

Calibration has been utilized extensively in research to determine appropriate values for sensitive hydrological parameters when high-resolution data is inadequate or unavailable (Choi & Ball, 2002). SWMM calibration implies carefully modifying control parameter

values until the results from simulation and observation are consistent. However, SWMM calibration requires a lot of time and resources, requiring automation in the process. Calibration is only available for gauged catchments; however, adequately gauged catchments are rare in urban areas. Initial parameters can be identified successfully without calibration. However, this will require data to be collected, processed, and estimated as the model parameters, which require some manual work for large urban areas.

This laborious parameterization of the catchment demonstrates the need for automated techniques, especially if applied at a large scale (Krebs et al., 2014). Most of the initial model parameters can be derived by GIS tools that utilize topographic data, including DEM, land use, and soil type. A number of prior research have shown the application of GIS to SWMM input data preparation (Jain et al., 2016; P. H. Martin et al., 2005). Thus, automatic parametrization with GIS tools can give some geometry parameters of sub-catchment and help to assimilate surface parameters like manning's coefficient and depression storage to the literature-suggested values.

3.6 Catchment physical (geometrical) parameters

Catchment delineation (area/border)

Catchment (sub-catchment) delineation is a requirement for hydrological evaluation and runoff estimation (Ray, 2018). Delineating and demarcating are the first steps of the catchment definition. Delineation is crucial since it establishes a control volume within which the runoff equations are utilized and the external forcing, such as precipitation and evaporation, are applied (Jankowfsky et al., 2013).

Sub-catchment delineation is more complicated in urban areas as drainage networks, cut and fill, and road curb and gutter construction may change flow direction and sub-catchment borders. Before the advent of remote sensing and DEMs, catchments were delineated mostly by manual drawings using contour maps and visual interpretations (Salih & Hamid, 2017). In large metropolitan areas, difficulties caused by the extent of the area and complicated surface flow warrant the need for automated catchment delineation with a systematic approach. Despite fine details like road curbs, DEM-based delineation with the required resolution has been proven to produce acceptable results compared to visual interpretation (Krebs et al., 2016).

Sub-catchments are typically delineated as irregular polygons, which can be stored in the format of polygon feature layers by GIS. Such layers can contain sub-catchment parameters and attributes required for further GIS processing.

While previous research has advanced the use of GIS tools to analyze flow routing (Lhomme et al., 2004), Dongquan et al. (2009) were among the first to use a Digital Elevation Model (DEM) to delineate catchments in ArcGIS automatically. The approach combines cells that are part of the same drainage basin, independent of land cover, storm network configuration, or flow routing between sub-catchments. Furthermore, Sanzana et al. (2017) introduced Geo-PUMMA, a semi-automatic method for generating well-shaped vectorial meshes or Urban Hydrological Elements.

The catchment definition is done based on the main hydraulic networks. Consequently, to enhance precision and enable realistic DEM-based catchment delineation in urban areas, all conduits are recommended to be included in the pre-processing of a DEM, even if not explicitly modeled (Gironás et al., 2010). Ji & Qiuwen (2015) offer a technique for incorporating stormwater network into DEM-based catchment delineation through the spatial analysis method in ArcGIS. A similar technique known as burning the DEM has been shown to be efficient for defining catchments (Almeida Silva, 2019). These details, however, are not widely known by water companies. Several commercial modeling software (e.g., InfoSWMM, PCSWMM) provide GIS tools for automated DEM-based catchment definition; nonetheless, only a few consider network inlets in the process.

Warsta et al. (2017) introduced a free-to-use open-source (GisToSWMM5) tool for automated raster-based sub-catchment generation by using a uniform computation grid (cell-based) and automated sub-catchment connection. The method considers each raster cell as one sub-catchment, which results in small enough catchments to warrant the use of homogeneous land use types. However, the drawback of the high-resolution cell-based catchment approach is the excessive resolution and long simulation times that make the method unfeasible for large areas. Niemi et al. (2019) enhanced the tool (GisToSWMM5) by including an adaptive sub-catchment generation capability and generating sub-catchments based on real surface flow paths and land use patterns. The tool merges minor sub-catchments that have a common outflow and a homogenous land cover, significantly reducing the number of sub-catchments. This approach still produces a large number of sub-catchment due to the diversity of flow direction and land use in urban areas.

Subcatchment Width

Sub-catchment's width is a particular hydrologic parameter that influences the time of concentration and defines the shape of the runoff hydrograph. In SWMM, irregular sub-catchments are transformed into roughly equal rectangular planes with a slope and width that represents the width of the overland flow. This parameter specifies the physical width

of one of the edges of this rectangle to which surface flow is directed perpendicularly. Sub-catchments, however, are rarely perfectly rectangular since they are often defined using topography/contour maps and land use layers. Thus, the width cannot be estimated in a straightforward manner, and it is one of the least concrete SWMM parameters. It is commonly implemented as a calibration parameter as it can significantly impact the runoff hydrograph properties (Dell et al., 2021; Rossman, 2016).

There are techniques to determine an initial estimate of catchment width even without calibration. The most common method suggested in the SWMM user manual is to find the average maximum length of overland flow by dividing the area by this length, as shown in equation 7 (Rossman, 2016). The maximum length of overland flow is the length of the flow path from the sub-catchment's farthest drainage point before the flow gets channelized. The characteristic width should be computed by averaging the maximum lengths of a variety of feasible flow routes. Another practical estimation recommended by the SWMM user manual is selecting one to two times the length of the main drainage network throughout the sub-catchment (Rossman, 2016).

$$W = A/L_{avg} \tag{7}$$

$$W = \alpha L_{pipe} \tag{8}$$

Where;

A = sub-catchment area [m²]

L_{avg} = average maximum overland flow path [m]

L_{pipe} = length of the main drainage channel

α = dimensionless coefficient

Reviewing other methods suggested by other scholars indicates other ways of estimating the hydraulic width, depending on available data and catchment shape (e.g., Krebs et al., 2014; Nowogoński et al., 2019) including equations described in the formulas:

$$W = k\sqrt{A} \tag{9}$$

$$W = 1.5\, L_{avg} \tag{10}$$

Where;

k = dimensionless coefficient

Slope

The slope parameter defines the inclination of conceptual rectangular planes. It should represent the average slope along the overland flow route to the drainage system inlet. In reality, the slope varies within each sub-catchment, especially in large heterogeneous sub-catchments. Hence, the most feasible way to derive sub-catchment slopes would be to calculate from DEM or DTM.

The most suitable data for calculating the slope can be DTMs since they do not include objects on the surface, like buildings, representing the natural and surface inclination. DEMs can be used for calculation, but some DEM data have major flaws that should be addressed: the terrain should be cleared of the buildings. Excluding buildings remove the unrealistic values since buildings lead to significant vertical discontinuities in the terrain. Therefore, the slope at cells close to rooftops can be even hundreds of percent. If these cell values are left untouched, mean sub-catchment slopes can unrealistically be increased. Therefore, buildings' rooftops should be removed for slope calculation since the majority of the potential energy is usually lost in a rainspout due to turbulence instead of changing response time.

The slope percentage for each pixel can be derived from the neighboring cell, which can be averaged at each sub-catchment to compute the mean slope required for the runoff estimation. Furthermore, the slope can also be calculated longest flow path in the sub-catchment since flow is mainly concentrated in the stream cells while only small amounts of runoff are transferred from other upstream cells. The burnt DEM/DTM should not be used for slope calculations since the cells around the burnt channels would result in incorrect slopes.

3.7 Catchment surface (hydrological) parameters

The sensitivity analysis or parametric research demonstrates the importance of sub-catchment parameters. Hydrological parameters are typically derived from land cover classification created using remote sensing techniques.

Imperviousness

In SWMM, the imperviousness parameter expresses the percentage of impervious surfaces according to the sub-catchment total area (Rossman, 2016). Although impervious areas mainly include roads and streets, their physical definition is hard since many surfaces are only partially impervious. The imperviousness can be determined in various different techniques, including calibration, estimation using land cover information,

imperviousness layer and automatic or manual image pro-cessing of aerial or satellite orthophotos.

Impervious surfaces have been highlighted as an indicator of the impacts of urbanization on water resources. Runoff volume and flow rates are very sensitive to estimates of imperviousness; thus, they should be calculated more accurately (Rossman, 2016). Total impervious area (TIA) is a characteristic defining the extent of imperviousness in a watershed. However, impervious areas directly connected to the drainage system (DCIA) (e.g., surrounded by pervious surfaces) have been reported to be a better parameter for representing impervious surfaces in urban areas (Ebrahimian et al., 2016). In addition, effective impervious area is the primary contributing area in rainfall events (Shuster et al., 2005).

If the sub-catchments have been delineated homogenously, a specific value of imperviousness can be set for all sub-catchment of the same land cover. Each land cover's imperviousness can be estimated from recommended values in the SWMM user manual and literature. On the other hand, if sub-catchments are composed of different land cover types, imperviousness can be estimated based on the area-weighted average of land cover types. As a result, each land cover type percentage is multiplied by the associated percentage of imperviousness in each land cover type to determine sub-catchment imperviousness.

Manning’s roughness coefficient

Manning's roughness coefficient defines the frictional resistance of water when flowing over land surfaces (Rossman, 2016). It is important to note that the overland flow patterns differ significantly from open channel flow, as the water depth values might be a few millimeters or less (Sanz-Ramos et al., 2021).

Manning's roughness coefficient values for overland flow are not as well-known as those for channel flow. This uncertainty is due to the substantial variation in land cover, transitions between laminar and turbulent flow, and relatively shallow depth. Based on previous studies, a typical range for each land cover was suggested in the SWMM manual, but there is no consensus regarding the particular value illustrating the uncertainty around these estimations (Rossman, 2016). Different values are set for the pervious and impervious areas in SWMM sub-catchments. Good estimation can be made for both impervious and pervious portions of the sub-catchment by selecting the values representing all land cover types within each portion.

Depression storage

Depression storage is the depth or volume of ponded water that must be filled on the surface before stormwater runs off (Viessman & Lewis, 2003). In other words, it describes the part of rainfall that ponds in depressions on the ground or built surfaces with no possibility of escape and form surface runoff.

Initial abstraction caused by such phenomena consists of two parts: the volume of water for wetting and saturating the surface and the amount of water that is stored as surface ponding. In ponding, the smallest depressions are filled first, and after they are filled, the water flows into the larger depressions and if the intensity of the rain exceeds the depression capacity of the surface, the water ends up in the rainwater drainage network or in an open ditch, forming an effective surface runoff at the discharge point of the catchment area. On pervious areas, water in depression storage can infiltrate and evaporate; thus, storage capacity may be refilled frequently (Rossman, 2016). However, on impervious areas, water can only evaporate, requiring a longer time to regain its total capacity (Rossman, 2016). Depression storage is a sensitive parameter for small storm events as the volume stored directly removes the volume available for generating runoff (Xu et al., 2019).

Different values of depression storage can be utilized for the pervious and impervious areas within a sub-catchment. In addition, depending on whether having depression storage or not, the impervious area may be separated into two types of subareas. Thus, some urban land surfaces, including buildings or steep roofs, may be regarded as impervious surfaces free of depression storage if precipitation on them is quickly drained off. In SWMM, a part of the impervious area can be set to not having depression storage with the percent *"% Zero-Imperv"* parameter to stimulate immediate runoff (Rossman, 2016). The sub-catchment fraction covered by buildings is usually recognized as the proportion of area with no depression storage which could be estimated using land cover data.

The amount of depression storage depends on the surface material, its condition and the slope of the surface (Arnell, 1980; Viessman & Lewis, 2003). Thus, depression storage can be dependent and calculated from the slope, as shown in Figure 5, or different built surfaces (Arnell, 1980). Furthermore, some recommended values for depression storage from literature are available in the SWMM user manual, such as suggesting values up to 2.5 mm for impervious surfaces and values from 2.5 to 5.1 mm for lawns (Rossman, 2016). The numbers are not very accurate, and depression storage is one of the typical calibration parameters (Swathi et al., 2019).

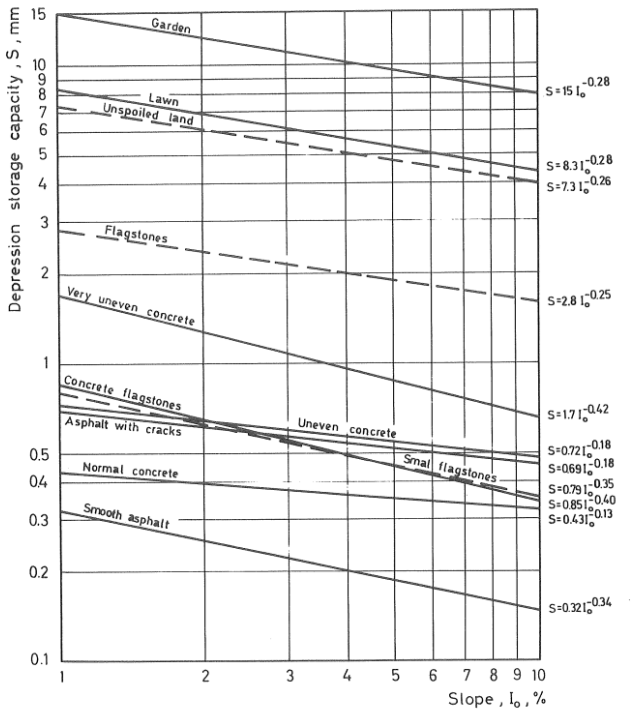


Figure 5. *Depression storage capacity as a fraction of surface slope (Arnell, 1980).*

3.8 Spatial resolution impact on rainfall-runoff modeling

The catchment's high-resolution description details are important in urban environments to make an accurate representation (Q. Chang et al., 2019). Recently, the accessibility of high-resolution data has increased, combined with the growth in computing resources. As a result, more detailed high-resolution models are developed. Although there is an unavoidable tendency to simulate SWMM models at a higher resolution the advantages and disadvantages of doing so should be studied in greater detail. One of the worries was that the high resolution might raise an uncertainty or lead to overparameterization.

Creating a model with a greater resolution may depict more processes occurring on catchment surfaces. However, defining the proper scale and balancing model features with an efficient and manageable computational load and keeping a certain level of modeling accuracy have proven to be challenging (Q. Chang et al., 2019). Previous studies on changing the resolution of the SWMM models showed that while a high-resolution model has a relatively low impact on total runoff volumes, it is highly effective on simulated peak flows (Q. Chang et al., 2019; Ghosh & Hellweger, 2012; Krebs et al., 2014). According to Q. Chang et al. (2019) findings, using a higher-resolution model increased peak flow. Warsta et al. (2017) experience the same results using grid cells as subcatchments. However, Krebs et al. (2014) findings showed that a higher-resolution model has lower flow peaks.

Furthermore, since sub-catchments are defined and simulated with the hydraulic networks, hydraulic results might be affected by drainage network details (T. Chang et al., 2018; Jang et al., 2018). Recent studies investigated the effects of the inclusion of stormwater inlets in flood analysis in urban areas. Their results highlighted the importance of inlets for correctly representing flood extent as they determine actual drainage capacity.

4. MATERIALS AND METHODS

The study's main objectives were largely defined by the HSY, the client of the study. Research questions were formed from the main objectives as presented in section 1. The study's main goal was to update the sub-catchments for the HSY CS model where the effects of automated catchment definition on hydrology and network hydraulics were studied. Since various methods were chosen for comparison, including the new method by SCALGO, particular test models were developed for four different case areas. After identifying the best method for scaling up, its performance on a large-scale area was investigated.

4.1 Research process

The investigation began with structuring and understanding the problem. It was attempted to gain as much reference as possible from similar studies to investigate essential issues within the project's scope. The study work's methodology is depicted in a flowchart form in Figure 6.

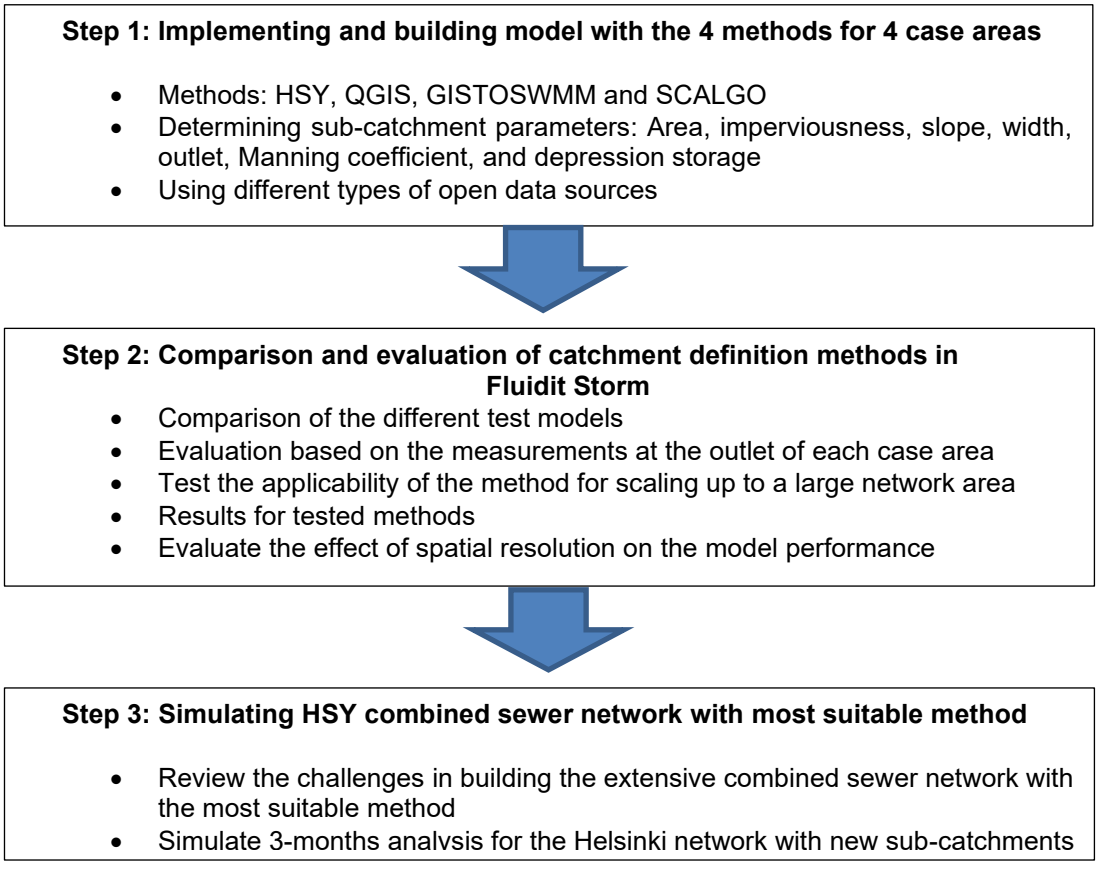


Figure 6. *Research process.*

According to the goal of the thesis, this study was conducted in three main steps. In the first step, sub-catchment was defined in four case areas with four catchment definition methods that claim to be suitable for large urban areas. In each method, various GIS spatial and observation data analyzes were performed. Within each method following parameters should be defined: boundary (delineation or area), imperviousness, slope, width, outlet, Manning coefficient, and depression storage (and area without depression storage). The methods might have different ways of calculating some parameters while sharing the same for other parameters. The result of each method will be exported in GIS format in which each sub-catchment has an attribute table for each parameter.

In the second step, the sub-catchments generated using each tested method were applied to a rainfall-runoff model to evaluate their performance. Since HSY has already adopted the Fluidit Storm network modeling system for their combined sewer system simulations, the testing and analysis were carried out on the same platform. Fluidit Storm, part of Fluidit Oy's product family software[1], is based on the widely adopted open-source EPASWMM simulator. However, unlike the open-source version, the Fluidit Storm software has a powerful GIS user interface and enhanced data management features such as support for multiple scenarios and python scripts. In addition, the scaling-up part (step 3) was not feasible in the open-source version due to its limited features. At the end of this step, methods will be evaluated and compared based on the simulation results, pros and cons, spatial resolution, and uncertainties.

In the third step, the most suitable method was used for testing and evaluating extensive urban areas. Sub-catchments generated in this stage were utilized to simulate the long-term simulation of the extensive Helsinki CS area described in Section 4.3.

4.2 Study site analyzed in this study

The city of Helsinki

Helsinki, the capital, is the biggest and most populous city in Finland. Helsinki is located in southern Finland on the coast of the Baltic Sea with a population of over 650,000 people The city has a humid continental climate with warm summers. The average annual air temperature is 5.9 °C, and the average annual rainfall is 655 mm, according to the Finnish Meteorological Institute's climatological 30-year summary from 1981 to 2010 (Pirinen et al., 2012). In Helsinki, intense convective summer rains with a small duration

[1] Fluidit Ltd, www.fluidit.com

are often the rain events that cause excessive surface runoff in the metropolitan area (Aaltonen et al., 2008).

In the 1870s, when the water utility started operations and wastewater volume began to expand, the building of the municipal CSS started. Since the 1960s, SSS has been constructed in new suburbs; nowadays, only SSS are legal to be built. Currently, the HSY's Viikinmäki WWTP sewer system, which was studied in the thesis, has an extensive CSS (1733 ha) in the city center and the adjacent older parts of the city. In addition, surrounding areas have a SSS (2460 ha) that partially connects to the Viikinmäki WWTP via the CS trunk lines. Furthermore, wastewater is conveyed from other municipalities via the northern and eastern sewage tunnels to this WWTP. The strategic and long-term goal is to replace CSS with SSS, but the process may take many decades since it is very costly and requires massive construction work in crowded and dense areas. The current investment plan includes the separation of 45 km of the total 220 km CSN in 10 years. Such constant changes in the network impose the need to automatically update the model. Figure 7 shows the extent of the Viikinmäki sewer model.

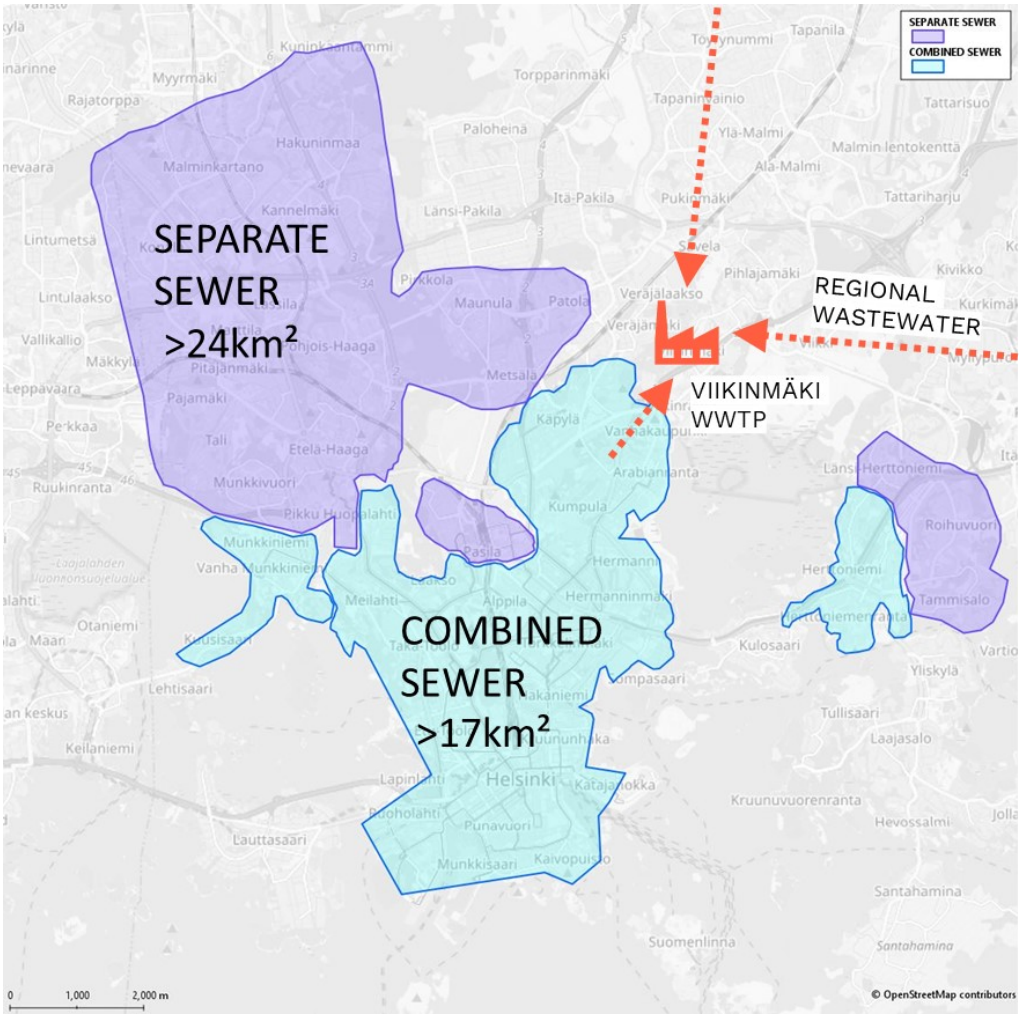


Figure 7. *Overview of the extent of the Viikinmäki sewer model.*

Population expansion and development in the CSS area led to a rise in the volume of wastewater and runoff. Heavy storms have caused several CSOs across the network, which might negatively affect the Baltic Sea. HSY continuously reports the overflows every quarter of the year through the hydrological model results. There have been various development projects to improve the model performance and increase the accuracy of the CSO reporting.

Stormwater runoff monitoring at discharge points of four small study catchments was studied initially to determine the best method for scaling up to the whole HSY catchments. The case areas Brahenkenttä, Taivallahti, Herttoniemi, and Munkkiniemi (illustrated in Figure 8) were selected as the region of interest for implementing the catchment definition methods for several reasons. First, all four areas are located within the city of Helsinki. Second, there have been data measurements in these case areas. Third, the areas are located approximately in different parts of Helsinki's network, representing different land cover and imperviousness, ranging from the dense urban area in the first case area (a) to the low dense urban area in the fourth case area (d).

Brahenkenttä area

The 40 ha Brahenkenttä area (see Figure 8(a)) is a residential area with apartment buildings and total imperviousness of more than 60%. It is characterized by concrete buildings, representing a city center area with a large fraction of impervious surfaces. In the area, new buildings have a direct link to the combined sewer while old buildings release their runoff into impervious surfaces which are directly connected to the CSS.

Taivallahti area

The 50 ha Taivallahti area (see Figure 8(b)) is a medium-density residential area with apartment building type with fragmented green area and yards. It is located on the west side of Helsinki center with more than 40% impervious area. The measurement device was installed before the weir, which discharges overflows to the sea.

Munkkiniemi area

The 50 ha Munkkiniemi area (see Figure 8(c)) is located in northwest Helsinki. It is a less intensively built or medium-density residential area (imperviousness is around 30%) residential area consisting of mostly single-family houses and areas large of vegetation. Some of the areas in the south have a separate stormwater network that discharges the water to the sea.

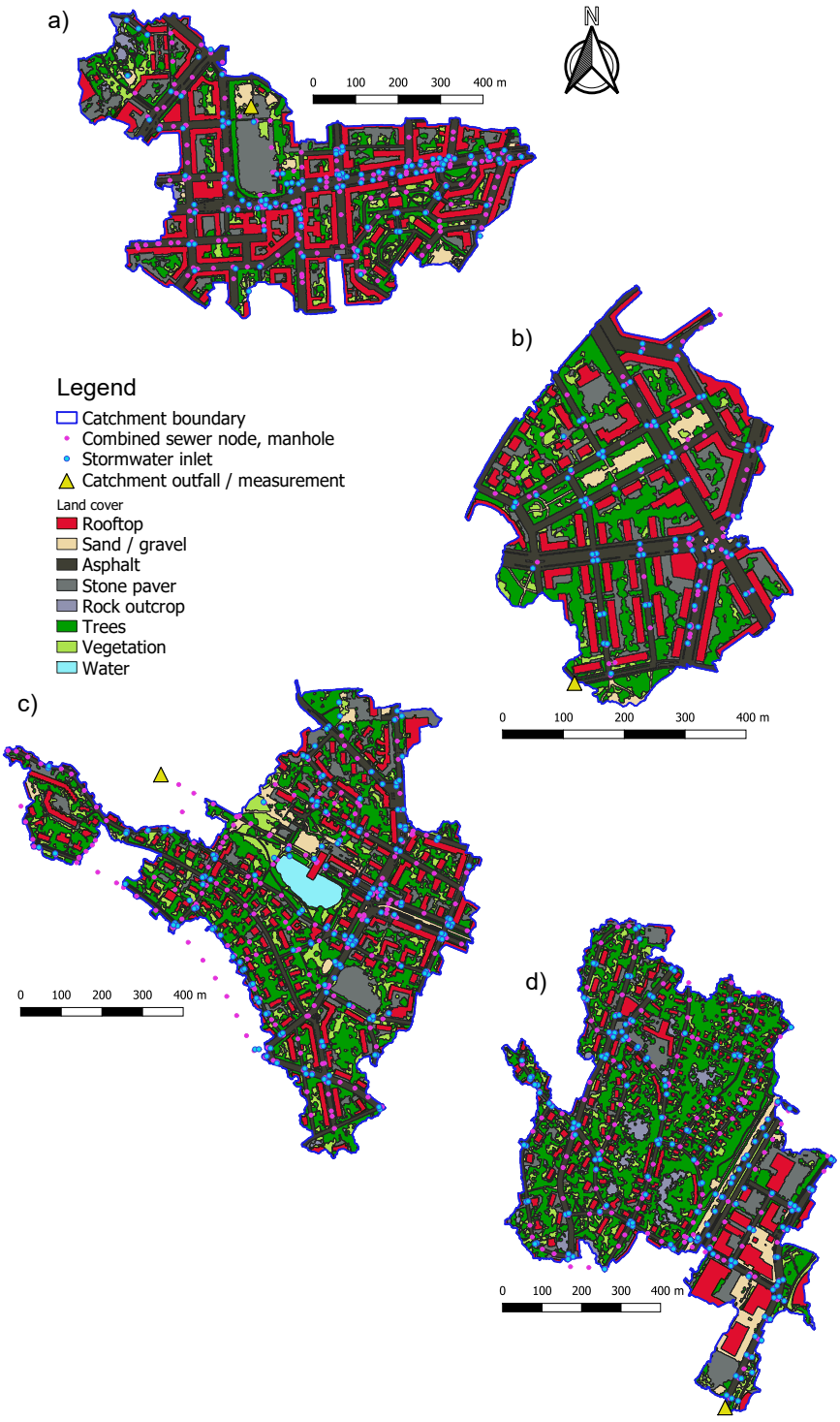


Figure 8. *Land cover characterization based on HSY dataset in (a) Brahenkenttä area, (b) Taivallahti area, (c) Munkkiniemi area, and (d) Herttoniemi area.*

Herttoniemi area

The 50-ha Herttoniemi catchment (see Figure 8(d)) is a low-density part of the Herttoniemi neighborhood. The area has been inhabited for a long time, forming a populated community in the 1910s. In the 1950s, the plans for construction were adopted, and many of the houses in this area were built during the same period. Herttoniemi is a

partially industrial area that is currently being developed for residential use. It is located about 7 kilometers east of the city center, running from the inner-city area through the eastern suburbs. The area is located on an elevated and naturally rocky area. It is considered a low-density urban area with relatively green parts and a 40% share of impervious surfaces. The areas located upstream of the measurement point do not cover the entire Herttoniemi neighborhood and have CSS. However, throughout the whole Herttoniemi neighborhood, there are some separate sewer pipes and stormwater management systems. Most of the older apartment blocks, surrounded by large, continuous areas of vegetation, are not directly connected to the CSS and drain onto adjacent pervious lawn areas.

4.3 Hydraulic network model

HSY had a hydraulic model of the Viikinmäki WWTP area prior to this study. The model has varying detail and parts, including combined sewer (CS) area, SSN area, and sewer conveyed from other municipalities via either measured or synthetic time series. The hydraulic network was sourced in SWMM based model. Additionally, they constructed the hydrologic part of the SWMM model for the combined sewer region using sub-catchments. Furthermore, some calibrations have been done to improve the model's results. Before this study, the HSY Combined Sewer Model was transferred from the US EPA SWMM platform to a fully compatible but vastly enhanced Fluidit Storm modeling platform that uses the SWMM engine and concept for 1D modeling in 2020 and has been updated yearly.

Household sewer inflow is estimated based on their water consumption bills and is updated once in a while to increase input data accuracy. Table 1 shows the total main drainage network length, average daily water consumption in each area, and the total Helsinki combined sewer network.

Table 1. *Sewer networks on case area.*

	Case area				**Helsinki combined sewer**
	Brahenkenttä	**Taivallahti**	**Munkkiniemi**	**Herttoniemi**	
Main drainage network (m)	5620 m	2830 m	1234	8630 m	220 km
Number of stormwater inlets	121	70	157	153	-
Average water consumption	6.5 l/s	6 l/s	6 l/s	6.1 l/s	1050 l/s

4.4 Data for rainfall-runoff modeling

This study relies on freely accessible spatial data sources. Most of the spatial data utilized in this study were obtained from the open data file download service of the National land survey of Finland (NLS, or Maanmittauslaitos in Finnish).

Digital elevation model

The digital elevation model (DEM) is a raster dataset with a 1 m grid cell size. Each cell has an average ground surface elevation of the cell.

Land cover data

The descriptions of land cover were derived from maps obtained from the City of Helsinki, Bing aerial images, Google Maps, and Google Street View. All of these are publicly accessible. The source data are from the 2019–2020 period and the final raster data has a resolution of 1m. The Helsinki region's land cover categorization datasets have been developed since 2014 to help urban planners with stormwater design, land use planning, and the monitoring of land cover changes. The datasets are updated approximately every two years. The land cover data is divided into 11 groups based on high-resolution color infrared aerial images and spatial data of roads and building footprints.

However, distinguishing between asphalt and sand/gravel surfaces is challenging due to their spectral properties. Incorrect interpretation of pervious (sand/gravel) and impervious (asphalt) may significantly impact hydrological runoff results required for stormwater assessments (Kokkonen et al., 2019).

High-Resolution Imperviousness Layer (LaserVesi)

The High-Resolution Imperviousness Layer has been generated in a collaborative project by SCALGO, HSY and The Finnish Environment Institute (SYKE). The project's goal was to provide data on artificial impervious surfaces in Greater Helsinki Region to support surface water management applications. The data was derived by SCALGO using machine learning and convolutional neural network model. The model was trained with high-quality land cover polygons created by Water Utilities. Such polygons were produced from National Laser Scanning and Aerial Imaging program data. The model generates a raster in which each pixel (of size 0.2m) reflects the probability of the pixel constituting an impervious surface (LaserVesi Project, 2022).

Runoff data

Combined sewer discharge data recorded at the catchments was gathered by recording devices, installed at the catchment outfalls inside combined sewer pipes. Discharge flow

was measured at Munkkiniemi and Herttoniemi from 29 March 2019 to 25 June 2019, using an ultrasonic flow measurement device with a one-minute resolution. In addition, flow data for Brahenkenttä and Taivallahti were obtained from 5 July to 26 August 2021 with a two-minute resolution.

Precipitation data

Precipitation is one of the most important inputs and sensitive parameters for hydrological models. Finnish Meteorological Institute's 5-minute rainfall radar data served as the basis for HSY's precipitation time series. Precipitation data were included in the HSY Combined Sewer Model utilizing *Rain Gage* components, which generated a 250-meter grid across the research region via an automatic weather station. The closest *Rain Gage* component was liked to each catchment.

Some events were selected from the period during which measurement was available in each case area to simulate the results, as shown in Table 2,

Table 2. *Summary of studied rainfall events.*

Area	Event code	Date	Event duration (h)	Rainfall depth (mm)	Peak rain intensity (mm/h)	Flow volume (m^3)	Peak flow (l/s)
Brahenkenttä	B1	21.8.2021	2	16.3	44	4150	1550
Brahenkenttä	B2	19.8.2021	3	6.2	16	1650	500
Taivallahti	T1	28.7.2021	5	11.3	19	1400	440
Taivallahti	T2	19.8.2021	3.5	11.5	42	1700	600
Munkkiniemi	M1	2.9.2019	4	5.2	17	1800	300
Munkkiniemi	M2	23.8.2019	6	34.5	39	9500	705
Herttoniemi	H1	11.5.2019	7	8.6	12	2000	360
Herttoniemi	H2	9.6.2019	2	2.7	9	800	330

4.5 Catchment definition methods

HSY method

Description

This method incorporates GIS tools and manual effort in catchment definition and was used originally to generate the sub-catchments for the model before this study. This approach is included in the study to allow a comparison of the new and existing methods. The process was created by a consultancy in 2016, but the original purpose was not the CSS model. Later on, sub-catchment parameters have been improved through additional calibration projects.

Implementation

Sub-catchments were delineated using some GIS tools but then finalized and merged with manual processes into smooth sub-catchments considering property boundaries and based on the modeler's judgment. The slope has been estimated based on the average slope within the catchment. The width parameter was calculated by dividing the area by the longest flow path which was the longest distance between the drainage system inlet and a corner of a catchment.

Some surface parameters like depression storage and Manning surface roughness coefficient were set as a constant value. For all sub-catchments, depression storage and Manning coefficient were set to 0.572 mm and 0.01 for the impervious sub-catchment area and 5.71 mm and 0.1 for the pervious area, respectively. Furthermore, imperviousness was estimated using landcover imperviousness percentage, shown in Table 3.

Table 3. *Impervious area as a percentage of land cover.*

Land cover	Percent impervious area
Road (paved)	80
Road (unpaved)	60
Buildings	90
Other impervious surfaces	90
Vegetation	10
Forested area	10
Open rock	40
Bare soil	40
Water	0

QGIS method

Description

QGIS a FOSS GIS software (QGIS Development Team, 2022) is a geographic information system (GIS) platform that offers users a free and open-source environment to work with and analyze geographical data QGIS offers tools, such as GRASS and SAGA, to integrate spatial and non-spatial data for automated catchment delineation and hydrological analysis. It was selected for this study because it is a freely available open-source application with well-documented user instructions that have been utilized in past research and real-world applications (Almeida Silva, 2019). The primary objective of this method is to subdivide study catchments and identify the watershed parameters utilizing software QGIS internal toolbox with less manual work.

It is important to note that QGIS was also used for pre-processing and GIS editing in the other methods. However, catchment definition and hydrological analysis are the focus of this section.

Implementation

The DEM with a resolution of one meter was utilized to identify all catchments draining into measurement locations in each case area. Initially, a depressionless DEM was generated using the *SAGA sink removal* tool because depressions could hold overland flow and alter the catchment delineation.

Utilizing DEMs without pre-processing for catchment delineation leads to stream networks that are inconsistent with reality. The reason is due to the neglect of the stormwater/sewer network and flow obstacles in urban areas. For example, the sub-catchment of an area may overlap with the pipe network from another area. This problem can be overcome by integrating the original DEM with vector sewer network data. Using the process known as 'stream burning,' the vector data describing the pipe network could be "burned" into the DEM raster data by decreasing the elevation of cells representing the sewage network. This will guarantee that the flow never leaves the channel and follows it to the discharge point (Saunders, 2000). Thus, the stream network data was burned on the DEM raster file using the *GRASS r.carve* tool. In the next step, catchments can be automatically delineated by the *GRASS r.watershed* tool. Figure 9 depicts the differences in sub-catchment delineation using the burning technique on the DEM.

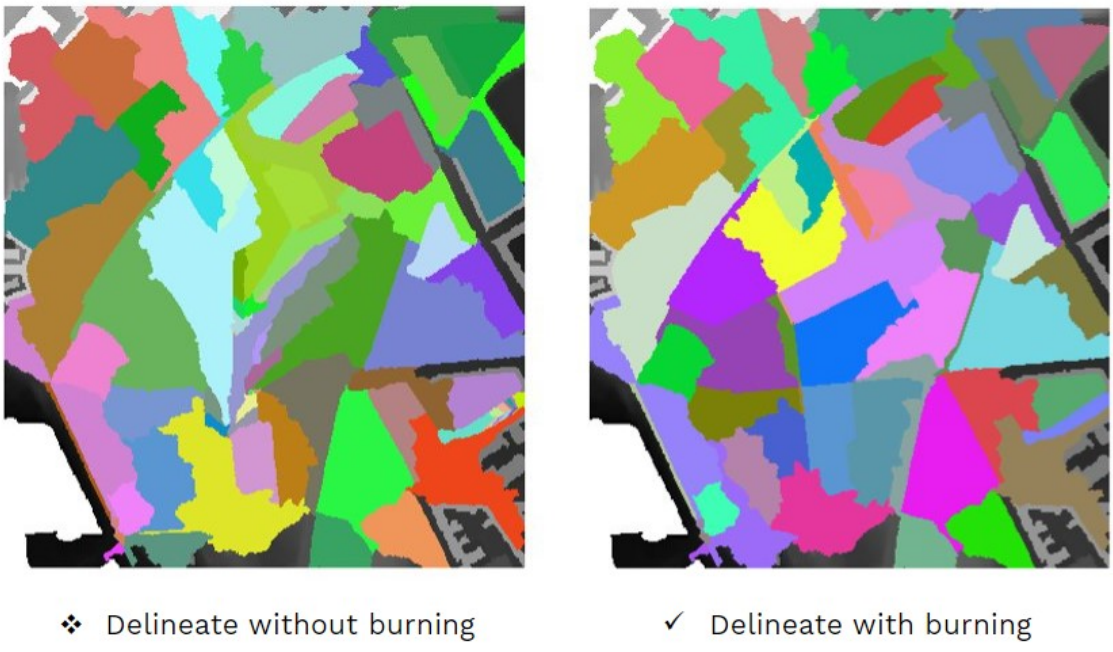


Figure 9. *Comparison of sub-catchment delineation with and without Burned DEM.*

The slope varies within each sub-catchment. To compute the average slope of each watershed, the following procedures should be taken to omit the building. First, as the high-resolution land cover data is available, buildings can be captured from this data. Second, a buffer should be applied to ensure the disabling effect of the high slope caused by the elevation difference between buildings and the surrounding area. Third, buildings can be clipped from the DEM. Finally, average slopes can be calculated for each of the sub-catchments with the *Zonal Statistical* tool.

Imperviousness was calculated using the same technique as in the previous method using the percentage of imperviousness in each land cover type from Table 4. Manning's roughness coefficient and depression storage were estimated based on the land cover data. As a result, the dominant land cover type on each sub-catchment has a significant impact on the ultimate average value. The relation between land cover type and these parameters is presented in Table 4. The coefficients were selected from a suitable range in the SWMM user manual (Rossman, 2016). For this reason, the percentage of each land cover type within each catchment was calculated using the *Zonal Histogram* tool, and an area-weighted average calculation was carried out to estimate the corresponding value. The calculation was done for both impervious and pervious subareas since separated values are given for the depth of depression storage. The width parameter was calculated by fitting each sub-catchment in a rectangular using the *Oriented minimum bounding box* tool and assigning the rectangular width for the sub-catchment.

Table 4. *Estimates of Manning's roughness coefficient and depression storage for land cover types.*

Land cover	Manning coefficient	Depression storage (mm)
Road (paved)	0.011	0.42
Road (unpaved)	0.03	2.4
Buildings	0.012	0
Other impervious surfaces	0.015	1
Vegetation	0.2	5
Forested area	0.09	10
Open rock	0.035	0.3
Bare soil	0.03	3
Water	0.011	0.1

Two particular challenges were identified in the process of connecting the sub-catchment to the hydraulic network. Initially, assigning a correct outlet node for each sub-catchments is a time-consuming manual task in the catchment definition, requiring some level of automation in the process. Furthermore, usually not all the sub-catchments generated

with selected DEM belong to the studied area. In other words, they might discharge their runoff to another outlet point rather than the network in the studied area.

To address challenges, the network's point components representing manholes were added to the QGIS for the case area and surrounding areas. Accordingly, while the outlet of the sub-catchments is defined, the outer boundary of the study catchment is shaped when unwanted sub-catchments for the case area are connected to the other network.

However, this is not the case for all studied areas as they do not have neighbor areas with drainage networks and are covered by forest or sea. Some manual adjustment was still required to delete the areas that were not being drained to the selected region.

GISTOSWMM method

Description

GisToSWMM5 is an automated sub-catchment generator and SWMM model builder, developed by Aalto university. The tool uses elevation, land use, flow direction, and network data to generate sub-catchments and identify surface flow paths either between catchments or into the stormwater network (Niemi et al., 2019). The tool is freely available under MIT-license from the GitHub online repository. The tool also provides means to export sub-catchments in GIS format.

All the input raster data must have the same resolution, grid, and coordinate reference system. The tool initially divides the area into a matrix of cells based on the data grid size. Each raster cell forms an individual sub-catchment and contains only one type of land cover, highlighting why data should be in the same grid and resolution. Flow direction data defines the cells' outlet that could be another cell or a stormwater network inlet. In the merging process, cells with the same land use and ultimate outlet are merged to form a sub-catchment. Furthermore, rooftop cells are either routed to the nearest inlet in the stormwater network or to the neighboring area if the rooftops are not linked to the network.

This method follows the methodology proposed by Krebs et al. (2014), where the sub-catchments are small enough to have one homogeneous land use type. Therefore, the tools generate a large number of sub-catchments with a high spatial resolution and a homogenous land cover for each.

Implementation

For this method, DEM and land cover data were provided in ASCII Grid format. For flow direction data, DEM has passed through the pre-processing to achieve a better result from the toolbox. This is due to the sensitivity of the result to the flow direction data. To

connect the catchments to the manholes, sub-catchments should route into the pixels around the manholes or other sub-catchment which has a route to the drainage network. Otherwise, the sub-catchment will be ignored during the merging process. Therefore, the final results would only include the sub-catchments connected to the network. Accordingly, the SAGA sink removal tool and GRASS r.carve tool were used to remove depression and burn the network into the DEM before deriving flow direction data from it. Finally, the SAGA channel network tool was used to calculate the flow direction file. The average slope of each cell is determined by calculating the arithmetic mean of the slopes between the current cell and the eight adjacent cells, and when cells are merging, the mean slope of the contributing cells is given to the merged catchment. Furthermore, when calculating the average slope, slopes between roofs and other types of land cover are disregarded to avoid high slopes. The toolbox automatically calculates the width of generated sub-catchments based on their area according to the following formula:

$$W = 0.7\sqrt{A} \tag{11}$$

The HSY land cover classes were slightly reclassified or renamed into the corresponding landcover suggested by the previous paper, which used this method for HSY landcover (Kokkonen et al., 2019). Furthermore, the suggested sub-catchment parameters for each land cover in the toolbox are demonstrated in Table 5 (Kokkonen et al., 2019). As the sub-catchments have homogenous land cover, artificial surfaces were considered to be completely impervious, whereas agricultural, forest, semi-natural, and wetland surfaces were deemed to be pervious.

Table 5. *Sub-catchment parameters according to (Kokkonen et al., 2019).*

Land cover	Imperviousness (%)	Manning coefficient	Depression storage (mm)
Asphalt	100	0.011	0.42
Rock	100	0.03	2.49
Rooftop	100	0.012	0.87
Sand and gravel	33	0.03	2.49
Stone paver	50	0.02	0.39
Vegetation (trees+grass)	0	0.238	4.22
Water	100	0.011	0.1

In addition, the generated sub-catchment can be extracted in GIS formats with toolbox utilities. Further information about the toolbox, data formats and implementation can be found in the toolbox GitHub online repository.

SCALGO method

Description

The SCALGO (SCALGO Live) is a commercial software which includes a dynamic catchment delineation toolbox. This toolbox was enhanced during this study through feedback loops and test-driven development. Software developers used comments from the author and Fluidit company to improve the merging algorithm and provide the required parameters for the SWMM sub-catchment. Throughout the project, many models were constructed to test and evaluate the toolbox subcatchments delineation, parametrization, and hydraulic results. Weekly or monthly meetings were held with stakeholders to discuss the development process and resolve bugs in the toolbox. The newly developed techniques were piloted in this project as HSY has a SCALGO Live license.

The toolbox can generate sub-catchments based on a topographic dataset and the stormwater network spatial data – more specifically, the stormwater inlet locations. With depressionless terrain data, the toolbox delineates sub-catchments contributing to each stormwater inlet based on flow paths on the terrain. The toolbox can also merge generated catchments according to the network to achieve a suitable catchment delineation for SWMM modeling applications. Moreover, a radius can be set for each inlet to capture the flow paths that likely should enter the inlet; however, they narrowly miss it due to the terrain model resolution. In this way, flows are better captured by inlets near the curb line, which help to generate more realistic catchments in urban areas.

Using the slider in the toolbox, the minimum size of the generated sub-catchments can be defined. In practice, catchments that are close to each other and have the same network can be merged together to achieve the minimum size. This enables the modeler to choose the level of spatial resolution of the sub-catchments for their model.

During the merging process, the new outlet will be set automatically according to the geometry of the new sub-catchment. The toolbox returns the land cover distribution within each sub-catchment according to the most updated HSY land cover dataset. Furthermore, it can also return the imperviousness of each sub-catchment based on the LaserVesi project. For each sub-catchment, the tool calculates the land cover percentage and the slope based on the start and end points of this longest flow path.

Implementation

The toolbox only needs pipe and node data to be imported in GIS format to delineate each area's sub-catchments. Most updated data, such as DEM and land cover, already exists in the toolbox for Nordic countries. All relevant stormwater and combined sewer

inlets can be considered since the toolbox can delineate sub-catchments based on inlet points, unlike other methods tested. However, using stormwater inlets can be challenging because they are not usually modeled and do not exist in the current HSY hydraulic network. In addition, even it is possible that stormwater inlet data is not available. In this project, stormwater inlet data was provided by HSY as a GIS layer and Google Street view was used to clarify unclear locations. Twice the length of the drainage network within each subcatchment was selected for the width parameter. Similar to the QGIS method, Manning's roughness coefficient and depression storage were estimated based on the land cover data.

Four scenarios were tested on each case area to analyze the impact of stormwater inlet inclusion, sub-catchment size, and different parameters. One scenario was built with only a combined sewer network, similar to the previous method, and another scenario was built with accounting for the inlets in the network. Furthermore, a scenario without a merging process was generated to evaluate the merging process and the impact of sub-catchment sizes on the model performance. This is because of the size of the results files, which should be considered beforehand to find a good balance between performance and resolution. Finally, since imperviousness is sensitive data, the third scenario for each case area was built to compare the new imperviousness data with the old land-cover method. The summary of and differences between the scenario built with the SCALGO toolbox is provided in Table 6.

Table 6. *Summary of models made with SCALGO.*

Name	Description	Purpose
SCALGO-Base	Catchment minimum size: 5000 m^2. Imperviousness data: LaserVesi Stormwater inlets: Not included	Evaluate the SCALGO catchment toolbox in typical use case.
SCALGO-Inlet	Catchment minimum size: 5000 m^2. Imperviousness data: LaserVesi Stormwater inlets: Included	Evaluate the inclusion of stormwater inlets in the models.
SCALGO-Small	Catchment minimum size: No merging Imperviousness data: LaserVesi Stormwater inlets: Included	Evaluate the impact of spatial resolution and catchment sizes on the hydraulic results at catchment discharge points.
SCALGO-Coeff	Catchment minimum size: 5000 m^2. Imperviousness data: Landcover data + HSY Coefficient Stormwater inlets: Not included	Testing the SYKE LASERVESI imperviousness data and comparing it to the old method with coefficient.

4.6 SWMM model simulation setups for case areas

SWMM model setting

The parameters selected for the simulation of the case area are shown in Figure 10. The *Report Step* was set to 2 minutes to provide higher resolution results for analyzing shorter-duration storm events. Additionally, 1 hour and 1 minute were selected for the Dry and Wet steps, respectively. *Report Averages* setting was allowed to produce accurate average results and smooth results curves. Moreover, the dynamic wave routing method was chosen for this study's simulations. Both pervious and impervious subareas were setup to discharge directly to the outlet.

Properties			
Report Step	00:02:00	Infiltration Model	Green ampt
Units	I/s	Lengthening Step	5
Allow Ponding	☑	Link Offsets	Elevation
Dry Days	0	Maximum Trials	16
Dry Step	01:00:00	Minimum Slope	0.001
Wet Step	00:01:00	Minimum Surface Area	0.503
Flow Routing	Dynamic wave	Minimum Step	0.1
Head Tolerance	0.01	Routing Step	5
Headloss Formula	Darcy–Weisbach	Skip Steady State	☐
Surcharge Model	Original EXTRAN Method	Lateral Flow Tolerance	5
Ignore Groundwater	☐	System Flow Tolerance	5
Ignore Quality	☐	Supercritical Flow	Slope + Froude number
Ignore Rainfall	☐	Sweep Start	Jan 1
Ignore RDII	☐	Sweep End	Dec 31
Ignore Routing	☐	Threads	12
Ignore Snow Melt	☐	Variable Step	0.75
Inertial Terms	Dampen	Report Averages	☑

Figure 10. *Model Properties for case area simulations.*

Generated sub-catchments from different methods were imported into Fluidit Storm in separate model scenarios. Python codes were developed for each method specifically. These codes were used before running a simulation to assign or calculate parameters from the GIS file attribute table and connect the sub-catchment to the nearest Rain Gage. Furthermore, they were utilized to provide sub-catchment statistics.

Evaluation method

The simulated results were compared to observed values at the catchment outfall visually as well as Nash-Sutcliffe efficiency (NSE) (Nash & Sutcliffe, 1970), root-mean-square error (RMSE), and peak flow error (PFE), as follows:

$$NSE = 1 - \frac{\sum_{i=1}^{n}(Q_{obs} - Q_{sim})^2}{\sum_{i=1}^{n}(Q_{obs} - \bar{Q}_{obs})^2} \quad (12)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Q_{sim} - Q_{obs})^2} \quad (13)$$

$$PFE = 100\frac{Q_{sim,max} - Q_{obs,max}}{Q_{obs,max}} \quad (14)$$

where Q_{obs} is the observed discharge at the time t_i, Q_{sim} is the simulated discharge at the same time, $\bar{Q}_{obs}$ is the average of all Q_{obs} values, $Q_{obs,max}$ is the maximum of all Q_{obs} values, $Q_{sim,max}$ is the maximum of all Q_{sim} values and n is the length of the data.

In a model with zero estimates error variance, the NSE is equal to 1. In contrast, a model with equal estimating error variance and observed time series variance has an NSE equal to zero. In other words, NSE = 0 means that the model has the same prediction skill as the measurement mean in terms of the sum of the squared error. An NSE coefficient higher than 0.75 is considered good, while a value below 0.3 is poor. A value in between is regarded as satisfactory.

4.7 Scaled-up method for Helsinki combined sewer model

Based on the evaluation of the tested methods and discussion in section 5.2, the SCALGO method was deemed to be the best method for the scale-up process. One of the challenges of the scaling-up process was the complexity of the HSY combined sewer network; for example, there are areas where separate stormwater networks overlap and connect with the combined sewer system. While the separate storm-water network was not included in the hydraulic model, individual separate storm-water network lines were imported to the SCALGO toolbox for better delineation results. Furthermore, networks were reviewed to detect the nodes which are likely gathering stormwater runoff and should be enabled as inlets in the toolbox.

Additional pre-processing of the network data was required for using the SCALGO toolbox on such an extensive area. For example, particular components like pumps or weirs might add more complexity resulting in the wrong catchment definition. Finally, the radius of the nodes for capturing the inlets was set to 8 meters, and the minimum catchment size of 5000 m^2 was selected.

The resulting catchments were applied to the HSY sewer model and used to simulate a 3-month period with the entire model (Q3/2021). The same simulation settings were adopted for the 3-month simulation, except for Report Step, which was set to 1 hour to

limit the result file sizes. Furthermore, storm events with CSOs were simulated again with 2 minutes Report Step and compared with the old catchments in more detail.

5. RESULTS AND DISCUSSION

This chapter presents the final catchment definition results and hydraulic outputs individually for each case area and the entire Helsinki sewer model. The results obtained using the four methods, as well as several scenarios for the SCALGO method, are reported below. First, the statistical results of different methods in the case areas are compared. Then, the performance of methods for simulating precipitation events is discussed in detail by analyzing the hydrographs at each case area outfall.

Furthermore, during the sub-catchment definition process, different values for some parameters were tested and simulated multiple times to achieve the best outcome. However, this chapter does not include a detailed presentation of the development stages required for each method.

Some of the results reported in this section, such as the computation time, may vary depending on the hardware used to run the model. In this study, simulations were calculated on a standard laptop computer running Windows 10 Pro. It uses an Intel 4-core i7-6820HQ processor, operating with 16 GB of RAM, and an AMD Radeon R M70 graphics card.

5.1 Catchment definition results on case areas

As a result of different catchment definition methods, sub-catchment with a variety of shapes and parameters were generated. Table 7 presents the sub-catchment area statistics for different sub-catchment delineation methods in each case area.

Observations on the size and number of sub-catchments

Results indicate the GISTOSWMM method had the most differences in statistical numbers compared to the other methods. The GISTOSWMM method resulted in an excessive number of small sub-catchments for each case area, even though it used adaptive sub-catchment generation to merge the cell-based sub-catchments. The minimum size of the sub-catchments in case areas, around 10 m^2 or 0.001 ha, was in line with the results from a paper by the author of the toolbox paper (Niemi et al., 2019).

Generally, input data resolution determines the minimum sub-catchment size in the GISTOSWMM method. Accordingly, the algorithm makes an independent sub-catchment for each cell if the cell is routed to a cell with a different land cover. The sub-catchments' mean and maximum size were found to be smaller than the other methods too. The

maximum size of the sub-catchments in this method is highly dependent on the number of land cover types and the homogeneity of the study side.

Table 7. *Sub-catchment statistics, area, and Total Impervious Area (TIA) for each case area and method.*

Area	Method	Number of sub-catchments	Area (ha)				TIA (ha)
			Min	Mean	Max	Total	
Brahenkenttä	HSY	36	0.3170	1.0238	2.39	36.86	26.16
	QGIS	41	0.1535	0.9339	2.37	38.29	24.87
	GISTOSWMM	29958	0.0001	0.0012	0.37	37.36	28.53
	SCALGO-Base	40	0.5161*	0.9862	2.66	39.45	27.15
	SCALGO-Inlet	35	0.5682*	1.0960	2.60	38.36	26.31
	SCALGO-Small	121	0.0006	0.3170	1.62	38.36	26.31
	SCALGO-Coeff	40	0.5161*	0.9862	2.66	39.45	26.10
Taivallahti	HSY	15	0.4410	1.3675	2.11	20.51	13.24
	QGIS	23	0.1854	0.8648	2.14	19.89	11.95
	GISTOSWMM	17375	0.0001	0.0012	0.35	20.35	13.29
	SCALGO-Base	21	0.5001*	0.9914	1.78	20.82	11.77
	SCALGO-Inlet	22	0.5084*	0.9458	2.03	20.81	11.81
	SCALGO-Small	70	0.0203	0.2973	1.37	20.81	11.81
	SCALGO-Coeff	21	0.5001*	0.9914	1.78	20.82	12.64
Munkkiniemi	HSY	27	0.1510	1.9957	6.11	53.89	31.89
	QGIS	53	0.1308	0.9651	2.73	51.15	25.08
	GISTOSWMM	50386	0.0001	0.0011	1.13	57.49	30.96
	SCALGO-Base	47	0.5238*	1.1392	4.47	53.54	21.43
	SCALGO-Inlet	40	0.5000*	1.2271	3.16	49.09	20.18
	SCALGO-Small	157	0.0010	0.3126	3.04	49.09	20.18
	SCALGO-Coeff	47	0.5238*	1.1392	4.47	53.54	25.89
Herttoniemi	HSY	23	1.1950	3.2857	9.92	75.57	37.09
	QGIS	60	0.0779	0.9465	3.76	56.79	26.03
	GISTOSWMM	55606	0.0001	0.0010	0.50	56.38	28.79
	SCALGO-Base	50	0.5158*	1.0422	3.92	52.11	20.99
	SCALGO-Inlet	50	0.5080*	1.0647	3.26	53.24	21.44
	SCALGO-Small	153	0.0010	0.3407	2.40	52.12	20.69
	SCALGO-Coeff	50	0.5158*	1.0422	3.92	52.11	24.20

* Minimum size of 0.5 ha was used for sub-catchments.

An example of routing between the cells in the sub-catchment generated by GIS-TOSWMM is presented in Figure 11. Based on the method's adaptive sub-catchment generation, if the upstream cell had the same land cover for each cell, the cell merged with the sub-catchment upstream of the cell. However, if the upstream cell had a different land cover, the upstream cell created a separated sub-catchment. Due to the high-resolution land cover data, assuming sub-catchments homogenous leads to the creation of numerous small sub-catchments (i.e., 1x1 m²). Using more land cover classes clearly leads to more sub-catchments in the GISTOSWMM method.

The high number of sub-catchments not only resulted in longer simulation time but also slowed down the software when working with the model. Using a coarser resolution for input data was tested by Warsta et al. (2017), and the most detailed resolution was considered the most accurate representation of the sites. However, even if inputs with coarser resolution were used, sub-catchment numbers would still be higher than in the other methods.

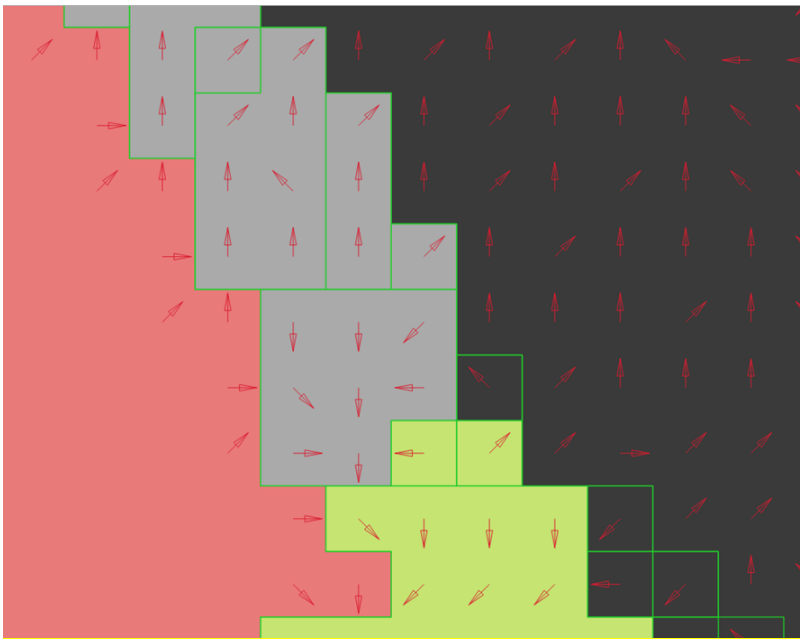


Figure 11. *Example of the flow direction effect (cell routing direction) in adaptive sub-catchment generation. The arrows represent each cell routing direction.*

While the GISTOSWMM method resulted in the largest number of catchments by a substantial margin, the SCALGO-Inlet method was second in terms of the number of sub-catchments. This SCALGO-Inlet method uses the contributing drainage area of each stormwater inlet, and therefore the number of stormwater inlets within each case area, to determine the total number of sub-catchments. Figure 12 shows how SCALGO merged sub-catchments based on stormwater inlet locations. During the merging process, the toolbox is able to return the name of the correct representative outlet. The algorithm enables the modeler to set the minimum size of sub-catchments according to the required spatial resolution. In general, all SCALGO methods with a minimum size set and the QGIS method resulted in approximately the same number of sub-catchments. The HSY method resulted in the largest sub-catchment area among the models with a minimum number of sub-catchments.

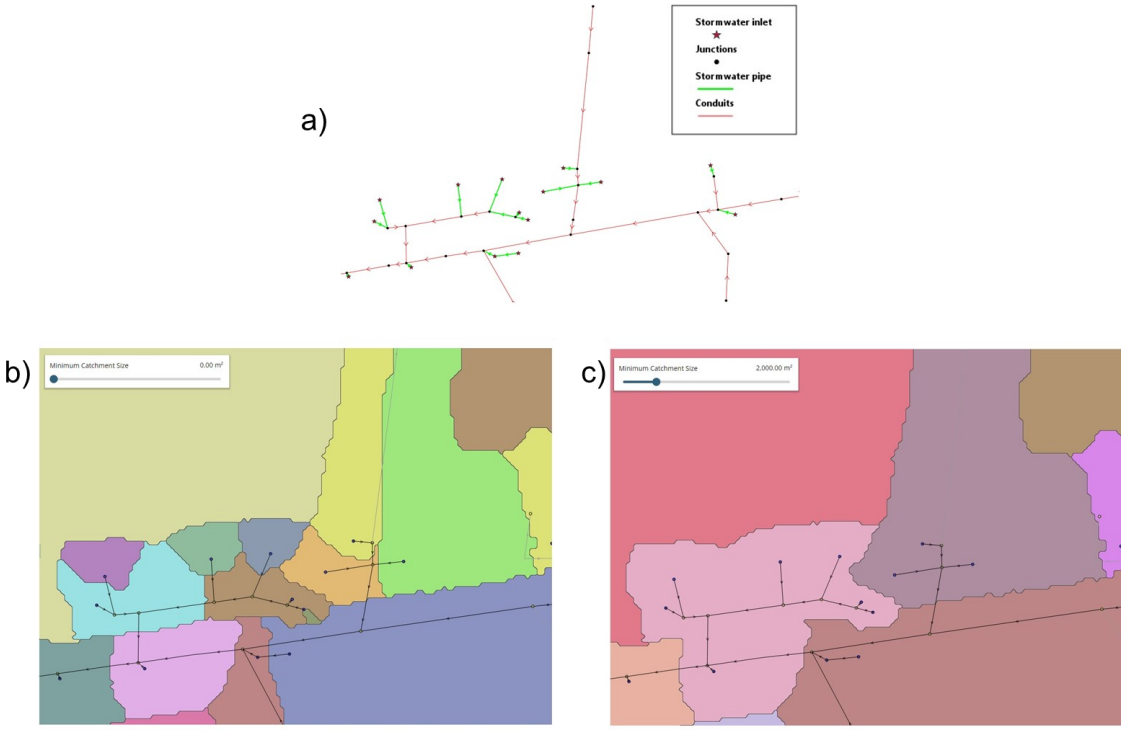


Figure 12. *Example of SCALGO sub-catchment delineation results. (a) Combined sewer components; (b) SCALGO sub-catchment for each inlet; (c) SCALGO dynamic merging algorithm results.*

Observations on the effects of methods on catchments' outer boundaries

All studied methods resulted in different outer boundaries, as shown in Figure 13. The outer boundary of the delineated sub-catchments demarcates the contributing area of each case area outfall, yielding differences in the total area for each method. In the same way, pervious or impervious areas near the boundary may be included or excluded using a different method, resulting in a different TIA. A significant difference is evident between the HSY method, and the other methods tested. Comparing the various scenarios of the SCALGO method showed that the impact of including stormwater inlets was minor. Some CSN pipes only carry sanitary sewer and have no stormwater inlets upstream (Figure 12(a), top middle). The outer boundary could only be changed if all nodes connected to those CSN pipes were considered stormwater inlets.

The total areas of case area 1 (Brahenkenttä) increased compared to the old sub-catchments, while the total area of case area 2 (Taivallahti) decreased significantly. There was a significant difference in Herttoniemi (case area 4), where the newly tested methods overlapped with only a portion of the old (HSY) sub-catchments. The best overlap was found for case area 1, where all methods shared more than 95% of their total area with other methods. The reason is that case area 1 is surrounded by the neighboring network, which helps to define the outer boundary of the catchment. However, the outer boundary

was more challenging in low-density areas where the outer boundary changed with different methods. The detailed delineation and the surface discretization of each case area are presented in Appendix 1-4.

Figure 13. *Method's sub-catchment outer boundaries for case area 2.*

Observations on total impervious area (TIA)

Comparisons demonstrated that different delineation and imperviousness calculation techniques have an impact on the TIA. Since impervious surfaces dominate the runoff generation process, the TIA is the most effective factor affecting simulated runoff dynamics (Krebs et al., 2016).

Comparison of TIA values derived from different sources (imperviousness data and land cover-based imperviousness) reveals the similarities of the TIA result in three case areas (case areas 1,2,4). In the first two areas, the total area is relatively similar in all models, which resulted in a reasonably similar TIA. Similarities between the imperviousness estimation data sources were anticipated since both could detect roads and buildings, which are major parts of TIA. However, in the case area three (Munkkiniemi), different sources came up with different percentages of impervious surfaces. This shows that the

high-resolution imperviousness layer underestimated a considerable portion of impervious surfaces in this low-density case area. One of the possible justifications is the challenging detection of other impervious surfaces from sand/gravel surfaces. SCALGO-Base, SCALGO-Inlet, and SCALGO-Small scenarios resulted in relatively similar estimates of TIA since they share the same area and imperviousness data. GISTOSWMM had the highest TIA compared to the other QGIS and SCALGO however, only part of it was hydraulically effective. In contrast, in QGIS and SCALGO sub-catchments, impervious areas were assumed to connect directly to the network.

The average values of the width, depression storage and Manning coefficient, along with characteristics of slope within each sub-catchment, are given in Table 8.

Table 8. *Slope statistic, and mean width, depression storage and Manning coefficient values for case areas using different methods.*

Area	Method	Slope (%)			Width (m)	Avg flow length(m)	Mean D_s		Mean n	
		Min	Mean	Max	Mean	Mean	Imp	Per	Imp	Per
Brahenkenttä	HSY	1.55	10.20	18.73	148.22	65.42	0.57	5.71	0.010	0.100
	QGIS	0.19	6.85	16.83	51.88	176.29	0.58	4.45	0.012	0.117
	GISTOSWMM	0.10	20.82	1330.0	1.33	2.72	0.56	4.04	0.017	0.216
	SCALGO-Base	0.60	3.78	7.80	265.45	66.68	0.66	4.73	0.013	0.116
	SCALGO-Inlet	1.10	3.95	8.70	382.08	54.09	0.66	4.73	0.013	0.117
	SCALGO-Small	0.00	3.90	23.00	110.52	110.40	0.58	4.23	0.012	0.112
	SCALGO-Coeff	0.60	3.78	7.80	265.45	66.68	0.66	4.73	0.013	0.116
Taivallahti	HSY	2.88	5.95	9.78	185.97	72.30	0.57	5.71	0.010	0.100
	QGIS	0.19	4.13	13.91	48.93	170.69	0.71	4.84	0.013	0.108
	GISTOSWMM	0.10	13.53	1145.1	1.34	2.73	0.42	3.95	0.017	0.206
	SCALGO-Base	0.50	2.93	7.40	240.40	60.91	0.72	4.86	0.013	0.101
	SCALGO-Inlet	0.50	2.44	7.70	301.57	56.49	0.71	4.87	0.013	0.101
	SCALGO-Small	0.50	2.70	9.40	94.78	71.61	0.62	4.18	0.012	0.099
	SCALGO-Coeff	0.50	2.93	7.40	240.40	60.91	0.72	4.86	0.013	0.101
Munkkiniemi	HSY	2.56	7.91	11.91	206.12	87.81	0.60	5.89	0.010	0.100
	QGIS	0.19	6.72	15.92	53.87	170.63	0.78	4.66	0.014	0.127
	GISTOSWMM	0.10	17.51	1284.5	1.33	2.73	0.43	3.99	0.018	0.211
	SCALGO-Base	0.40	4.19	8.50	373.10	38.43	0.76	4.70	0.014	0.120
	SCALGO-Inlet	0.60	3.79	8.60	392.61	44.99	0.78	4.71	0.014	0.119
	SCALGO-Small	0.00	3.91	14.00	100.03	173.99	0.72	4.50	0.013	0.115
	SCALGO-Coeff	0.40	4.19	8.50	373.10	38.43	0.76	4.70	0.014	0.120
Herttoniemi	HSY	4.79	11.47	16.86	283.59	108.54	0.59	5.49	0.010	0.100
	QGIS	0.77	10.89	22.99	49.38	176.63	0.72	4.75	0.013	0.113
	GISTOSWMM	0.10	16.55	1102.8	1.26	2.58	0.88	3.95	0.020	0.205
	SCALGO-Base	0.60	4.99	10.10	245.73	58.58	0.69	4.77	0.013	0.110
	SCALGO-Inlet	0.70	4.57	12.70	327.36	43.51	0.69	4.77	0.013	0.109
	SCALGO-Small	0.50	4.85	27.30	103.90	100.59	0.63	4.74	0.013	0.104
	SCALGO-Coeff	0.70	4.57	12.70	327.36	43.51	0.69	4.77	0.013	0.109

Differences in the range of sub-catchment slopes are due to different methods used for calculation. In GISTOSWMM, while cells with building landcover were not included in the slope calculation, some small (1×1) sub-catchments had a very high slope, which can be explained by local errors in the landcover data (Niemi et al., 2019). These local errors are usually for small built areas, like a shed, which were not detected as buildings in landcover data. However, according to GISTOSWMM's sub-catchment size, its effect is negligible. The same error might have occurred in the QGIS method, but with buffering and filtering techniques, slopes became smoother.

The width and average flow length of sub-catchments vary in different models as they directly relate to the area. The average values of depression storage and Manning coefficient are relatively similar for QGIS, and SCALGO, as they were assigned based on the land cover data and similar initial values. Regardless of the small-scale errors, values were in an acceptable range. However, each set of parameters from different catchment definition methods should be evaluated from their resulted hydrograph results.

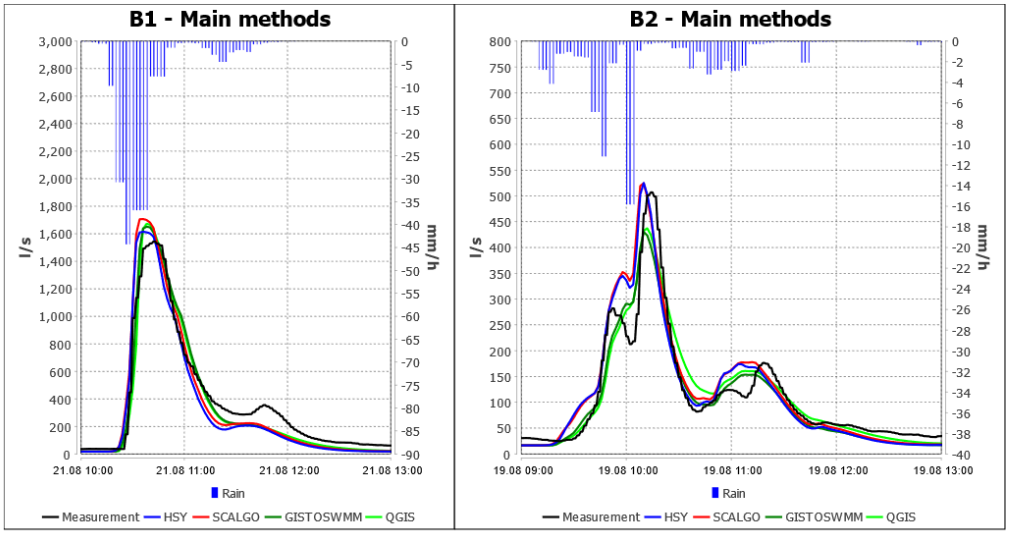
5.2 Hydraulic simulation results for case areas

Brahenkenttä

Figure 14 illustrates the discharge simulation results for models using different catchment definition methods for the Brahenkenttä case area. Each event is presented in two graphs for visual clarity purposes. Table 9 summarizes the corresponding performance statistics for each method. The assessment shows how well the measured event's temporal dynamics matched the simulated sub-catchment runoff's temporal dynamics at the area outfall.

Since the TIA is relatively uniform in all methods, the discharge hydrographs generally have a similar shape. However, slight differences in simulated flows were caused by differences in other hydrological parameters. Thus, it is hard to define which model yielded better performance. Furthermore, the results of the simulated discharge for SCALGO models are almost identical for the two events.

a)



b)

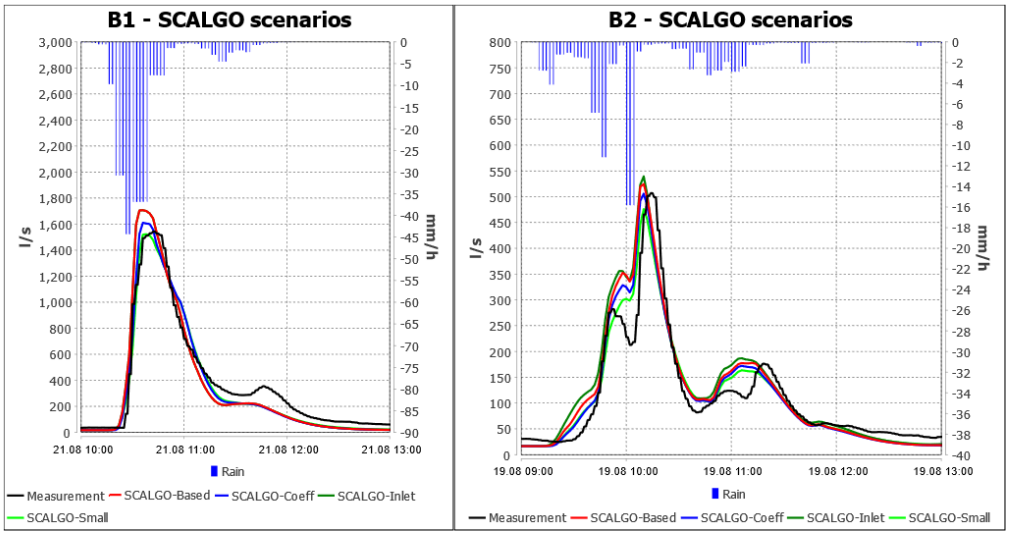


Figure 14. *Comparison of measured vs. simulated flows for the selected two events in Brahenkenttä with: (a) main methods ;(b) SCALGO scenarios.*

Table 9. *Models' performance for the Brahenkenttä study Catchment.*

Event	Method	NSE	RMSE	PFE
B1	HSY	0.923	121.49	4.60 %
	QGIS	0.952	95.24	8.32 %
	GISTOSWMM	0.954	93.11	6.98 %
	SCALGO-Base	0.925	119.63	10.57 %
	SCALGO-Inlet	0.925	119.64	10.58 %
	SCALGO-Small	0.961	86.23	-1.37 %
	SCALGO-Coeff	0.955	92.51	4.33 %
B2	HSY	0.864	40.06	3.60 %
	QGIS	0.913	32.00	-13.74 %
	GISTOSWMM	0.912	32.35	-15.17 %
	SCALGO-Base	0.853	41.76	3.51 %
	SCALGO-Inlet	0.817	46.95	6.43 %
	SCALGO-Small	0.909	32.81	-6.07 %
	SCALGO-Coeff	0.885	36.79	-0.20 %

Taivallahti

The measured and simulated discharge time series and performance statistics for the Taivallahti area are presented in Figure 15 and Table 10, respectively. Similar to the Brahenkenttä area, simulated discharge results for models were almost identical in pattern and magnitude for events. The SCALGO scenarios showed promising results since they performed well for peaks compared to the other methods. The peak flow in the first event was overestimated by old sub-catchments compared to other methods. However, in the second event, all the methods failed to simulate the second peak flow. Such drastic deviations in modeled discharge values from measurements were observed in the longer events, as in the previous case area. The amount of runoff generated by pervious surfaces during the second rainfall peak could justify the extra flow in the measurement data. The simulation results can be improved by adjusting the infiltration parameters, which is out of the scope of this study.

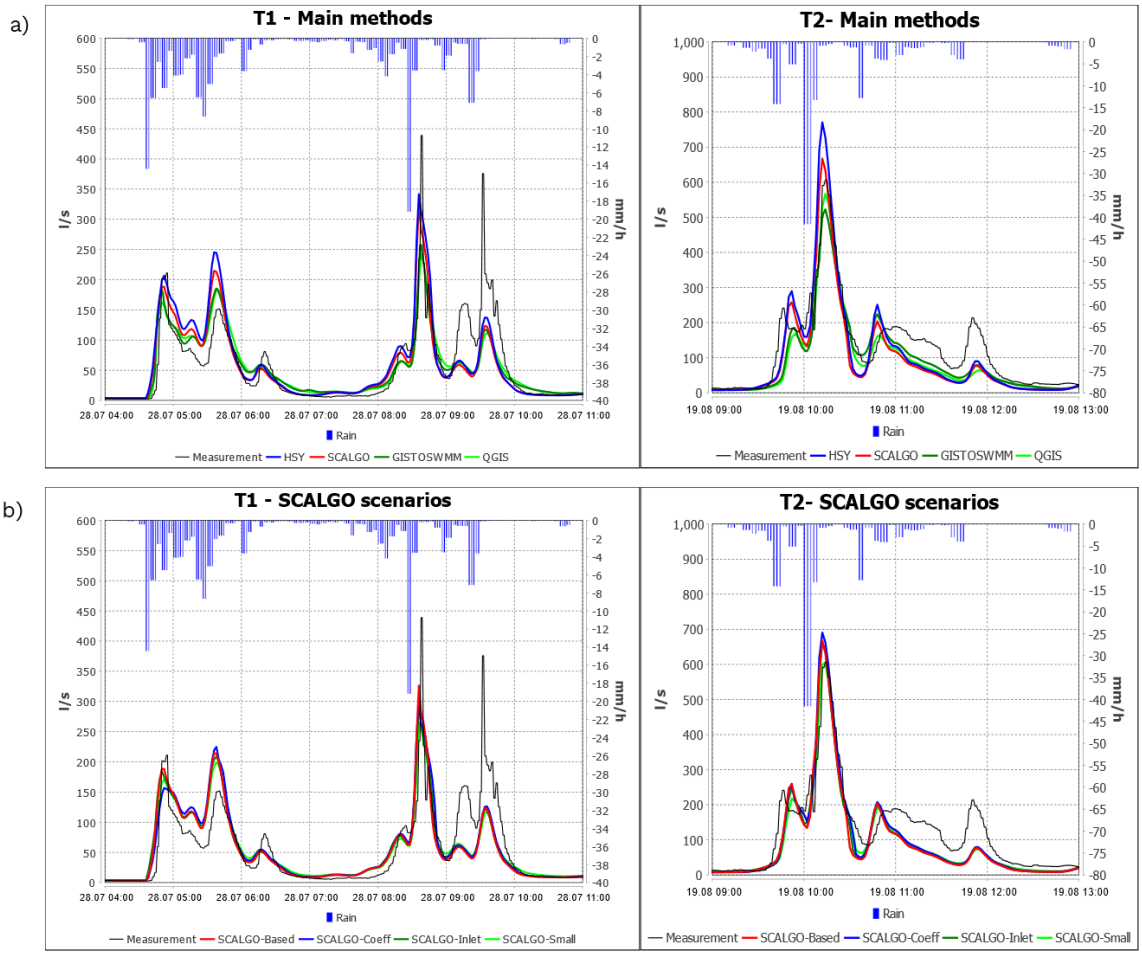


Figure 15. *Comparison of measured vs. simulated flows for the selected two events in Taivallahti with: (a) main methods ;(b) SCALGO scenarios.*

Table 10. *Models' performance for the Taivallahti study Catchment.*

Event	Method	NSE	RMSE	PFE
T1	HSY	0.587	42.67	-22.17 %
	QGIS	0.675	37.77	-44.89 %
	GISTOSWMM	0.682	37.39	-41.25 %
	SCALGO-Base	0.625	40.58	-25.63 %
	SCALGO-Inlet	0.630	40.32	-30.87 %
	SCALGO-Small	0.665	38.36	-39.53 %
	SCALGO-Coeff	0.624	40.64	-31.14 %
T2	HSY	0.714	65.26	27.04 %
	QGIS	0.809	54.76	-6.54 %
	GISTOSWMM	0.818	52.88	-13.87 %
	SCALGO-Base	0.767	60.00	10.09 %
	SCALGO-Inlet	0.788	57.24	9.52 %
	SCALGO-Small	0.819	52.85	-0.33 %
	SCALGO-Coeff	0.784	57.12	13.88 %

Munkkiniemi

The Munkkiniemi had the greatest difference in TIA despite having a nearly similar total area, especially between the HSY (old) and new sub-catchment proposed by this study. In the first event, results show that the older sub-catchments had better performance in medium events, as the other methods underestimated the total runoff volume.

In the second event, the total volume of the stormwater looks more similar. However, due to the smaller size of the combined sewer area, Munkkiniemi suffers from hydraulic bottlenecks and overflows in the manholes, even in a medium event. In addition, the second event was one of the heaviest storm events in recent years, and many overflows were observed in the network. Accordingly, part of the excess stormwater may have exited the network in overflows, and the final hydrograph is highly influenced by hydraulics, water ponding area, and overflows within the network. Therefore, evaluating the sub-catchment performance with this event was deemed misleading and the usability of the results is limited.

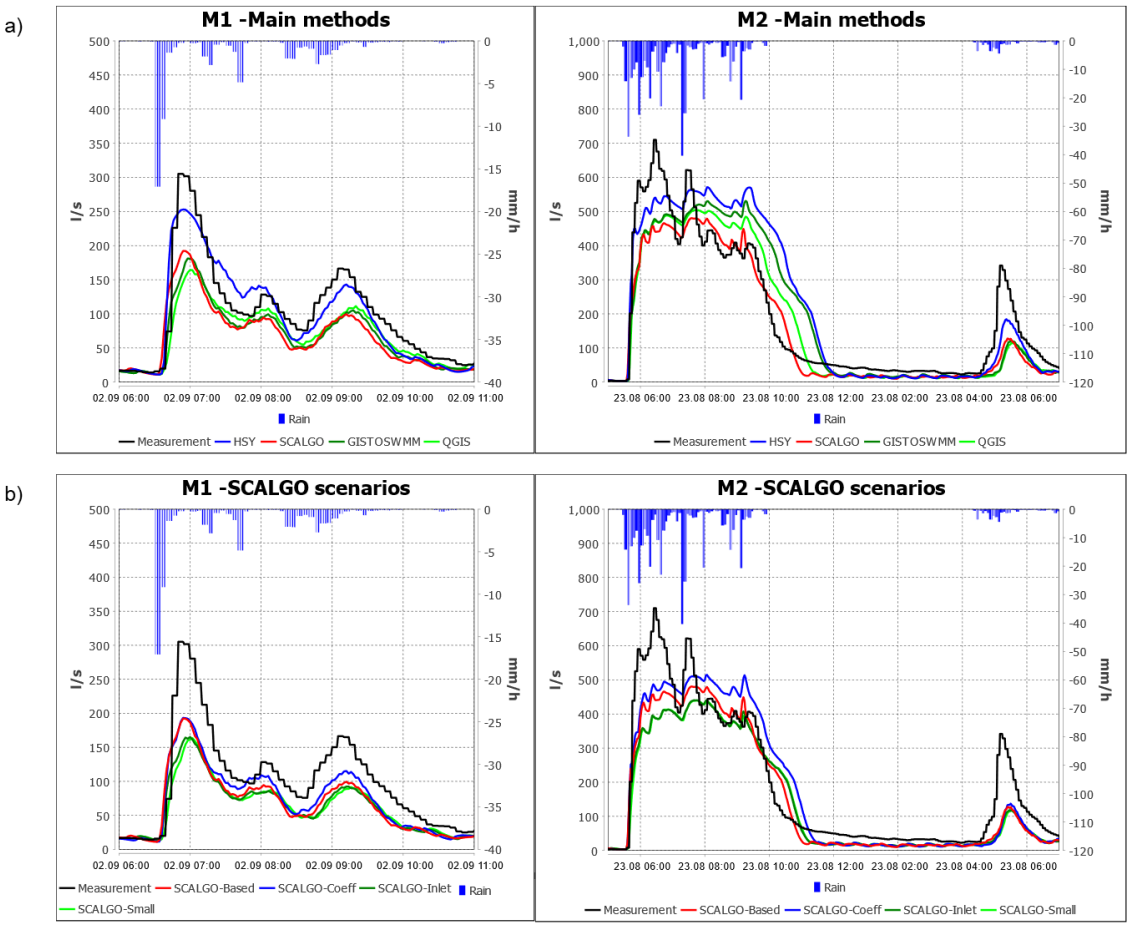


Figure 16. *Comparison of measured vs. simulated flows for the selected two events in Munkkiniemi with: (a) main methods ;(b) SCALGO scenarios.*

Table 11. *Models' performance for the Munkkiniemi study Catchment.*

Event	Method	NSE	RMSE	PFE
M1	HSY	0.869	25.77	-17.11
	QGIS	0.634	46.51	-46.15
	GISTOSWMM	0.665	45.13	-40.56
	SCALGO-Base	0.657	45.80	-37.00
	SCALGO-Inlet	0.587	51.48	-45.86
	SCALGO-Small	0.575	52.37	-46.43
	SCALGO-Coeff	0.718	40.32	-36.66
M2	HSY	0.721	106.16	-19.56
	QGIS	0.816	85.71	-29.00
	GISTOSWMM	0.739	101.72	-25.19
	SCALGO-Base	0.858	76.05	-32.30
	SCALGO-Inlet	0.802	90.59	-37.96
	SCALGO-Small	0.793	92.75	-38.07
	SCALGO-Coeff	0.834	81.36	-27.48

Herttoniemi

Figure 17 shows the measured discharge hydrograph from the Herttoniemi case area and compares it to simulated values in various methods. The simulated peaks using the HSY method were generally higher than the measurements and the results from other methods because of the larger TIA as noted in Table 7. Although GISTOSWMM had a higher TIA than QGIS and SCALGO, its hydrograph was almost similar to theirs. This is because not all the impervious surfaces were directly connected to the network. As a result, water that ran off from some impervious surfaces went through pervious areas and got infiltrated before draining into the network. Table 12 shows that models based on the SCALGO methods most closely match the flow measurements in the two events. GISTOSWMM and QGIS performed acceptably in these two events.

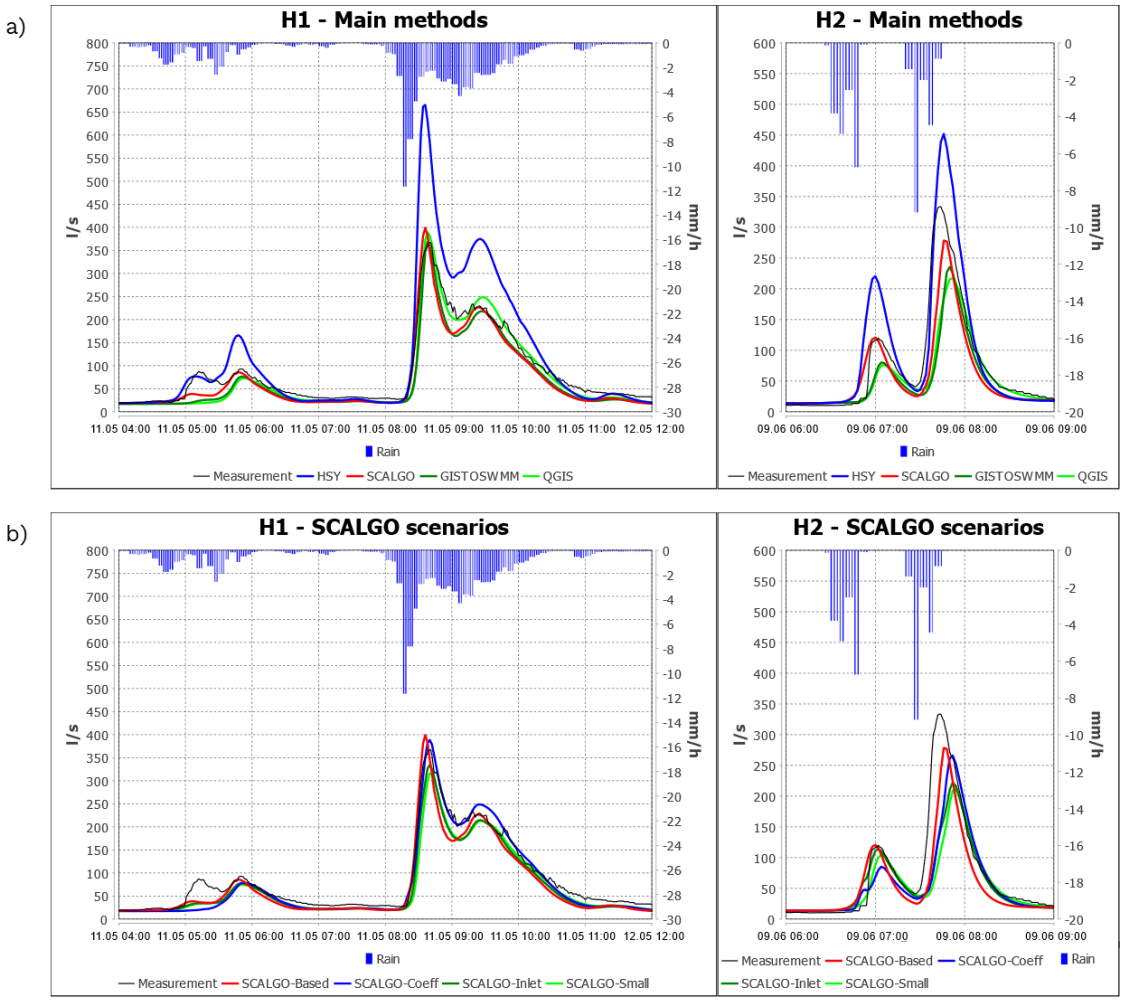


Figure 17. *Comparison of measured vs. simulated flows for the selected two events in Herttoniemi with: (a) main methods ;(b) SCALGO scenarios.*

Table 12. *Models' performance for the Herttoniemi study Catchment.*

Event	Method	NSE	RMSE	PFE
H1	HSY	0.399	66.26	81.19 %
	QGIS	0.940	19.84	6.01 %
	GISTOSWMM	0.9171	23.78	0.14 %
	SCALGO-Base	0.932	21.35	8.67 %
	SCALGO-Inlet	0.939	20.14	-8.99 %
	SCALGO-Small	0.919	23.23	-14.07 %
	SCALGO-Coeff	0.952	17.60	5.79 %
H2	HSY	0.725	45.77	35.48 %
	QGIS	0.653	51.81	-35.02 %
	GISTOSWMM	0.6239	54.70	-29.33 %
	SCALGO-Base	0.790	39.93	-16.50 %
	SCALGO-Inlet	0.654	51.17	-33.40 %
	SCALGO-Small	0.546	59.00	-36.65 %
	SCALGO-Coeff	0.667	50.16	-20.21 %

5.3 Method comparison and discussion

Overall, this study's results suggest that instead of the old manual method, automated sub-catchment definition is recommended for urban areas, using terrain topography, high-resolution land cover data, and detailed stormwater drainage system data. Automated methods seemed to work better in dense highly impervious urban areas compared to low-density areas. For example, in regions with large green areas, the runoff volume can vary considerably due to the routing of runoff from impervious areas to pervious areas. Small-scale features of the urban surface, such as curbs, are easily neglected in automated delineation methods (Krebs et al., 2016). However, using higher-resolution data (DEM, DTM, and land cover) can improve the results. The importance of detailed land cover types and high-resolution terrain elevation in SWMM model performance was highlighted by previous studies (Petrucci & Bonhomme, 2014).

Selecting the catchment definition method for an extensive city-scale model depends on many factors. Some of the most important factors are how fast the methods work for generating sub-catchments and how they affect the simulation of the model. Figure 18 compares the required time for building and simulating the model with each catchment definition method.

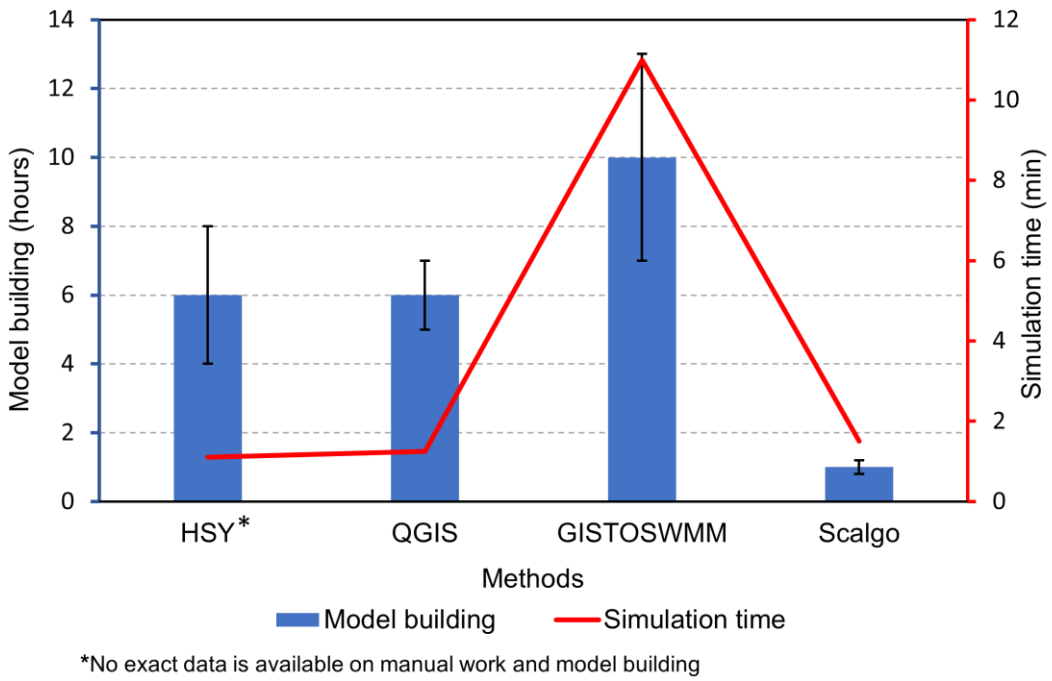


Figure 18. *Model building and simulation times for different catchment definition methods.*

Building times noted in Figure 18 refer to the time devoted to delineating sub-catchments, estimating parameters, and importing them into the existing hydraulic model. Thus, this is the author's estimation for building a model should the method be used in a new area to the same extent as the case area. The results include errors regarding each method since building time can vary a lot by modeler skill, computer hardware, and study area size. The results clearly show that the SCALGO method requires the least amount of building time due to less manual work and less dependency on the computer hardware since it is a Web application. The GISTOSWMM method required a longer building time, which varied greatly depending on the size of the area. This is because of the time required to create suitable input files and for the toolbox to merge the cells. Although there is limited data about the techniques used in the HSY method, it was anticipated that most tasks would be similar to those in the QGIS method. Therefore, according to the manual work of these two methods, they were estimated to have almost equal building times.

The term "simulating time" refers to the time it takes the SWMM simulator to compute and load the results. This includes the time it takes to run the event day and the day before it to balance the network. The simulation times observed show that the number of sub-catchments directly correlates with the simulation time required. The QGIS, SCALGO, and HSY methods have almost the same simulation time, while the same simulation took almost six times longer for the GISTOSWMM method. This is because GISTOSWMM had, on average, 600 times more sub-catchments. As a result, changes

in the simulation time can only be noticeable if the number of sub-catchments changes dramatically. The first reason is that sub-catchments are only part of the total components of the models. The second reason is that in SWMM, hydrological processes take less time to compute than hydraulic calculations.

Table 13 summarizes each method's general advantages and disadvantages in the sub-catchment delineation process. Modelers can decide on the best method for their purpose by comparing the hydrographs, computational load, simulation time, and pros and cons of each method. Therefore, the final decision for selecting the methods depends on many factors, such as the extent and complexity of the network, data availability, data resolution, the importance of the results, computer hardware, the budget, and the skills of the modeler. However, the conclusion and discussion in this study are based on looking for an automated catchment definition for a city-scale network model. This conclusion might be more in line with what utilities, cities, and consulting companies desire than what academics choose.

Table 13. *Summary of methods' advantages and disadvantages.*

Method	Advantages	Disadvantages
HSY (old)	• Include extra data e.g., Network information system with manual delineation	• Large manual work
QGIS	• Free software • Available manual and books • Question and answer website	• Availability of some tools in different versions • One-pixel or very small sub-catchments
GISTOSWMM	• Free open-source tool • High-resolution model • Tested in some scientific papers • Previous implementation in Finland	• Hard implementation (data format) • Require programming skill • Too many sub-catchments for practical use • Uncertainty leads to wrong results • Both long simulation and building time
SCALGO	• Realistic terrain (culvert, underpass) • Terrain tools (e.g.inlet radius) • Point based delineation • Works in complex network • Merging sub-catchments according to the network • Fast update for large area if any changes happen	• Commercial software

QGIS

QGIS has been reported to be used by many authors for sub-catchment definition purposes at a variety of scales. This is because of its free access and the many available manuals and question-and-answer websites. Such features enable this method to be used for small-scale models. The method was successful in defining outer boundaries for case areas, and the output hydrographs had good consistency with measurements. However, pre-processing of data and working with multiple toolboxes are still slow for a very extensive network.

In the QGIS method, delineating sub-catchments and selecting discharge points for sub-catchments is very demanding, especially in complex areas where both CSS and SSS exist. Moreover, the delineation sometimes generated very small sub-catchments, which meant that more error-prone manual work had to be done to fix the problems. Therefore, for a city-scale model, it still requires a lot of time from the de-signer to build the model.

GISTOSWMM

The main difference between the GISTOSWMM and other methods was its ability to generate high-resolution cell-based sub-catchments and find the water route between sub-catchments. In GISTOSWMM, instead of using only DEM, both land cover data and DEM were used in the sub-catchment delineation and sub-catchment routing. In other methods, routing between different land covers and, more importantly, between impervious and pervious areas was not established in such detail.

However, creating a very high-resolution model with GISTOSWMM did not result in any improvement in hydraulic results and model-building workload. Furthermore, no study site larger than 1 km^2 has been reported using this method. While the merging process reduced the number of sub-catchments by 90% compared to uniform 1 x 1 m^2 sub-catchments, the total number of sub-catchments was still excessive (Niemi et al., 2019). Consequently, based on the previous studies and the number of sub-catchments in the case areas, it can be predicted that a 17 km^2 urban area (e.g., the Helsinki CS area in this study) will have more than 1 million sub-catchments. Applying such a huge number of catchments into the hydraulic model would result in a high computational burden on the modeler’s device and lead to a very long simulation time. Long simulations are one of the major problems with models, especially if they are used to simulate a large area, like a city, or they are used for real-time application. While the GISTOWMM cell-based method is somewhat suitable for defining sub-catchments for small urban areas, it is almost impossible to use for a city-scale model. Hence, GIS-TOSWMM is not

recommended for studying city-scale stormwater systems since there are several methods that are easier to use and more efficient while producing similar hydraulic modeling output.

The GISTOSWMM was regarded as the most difficult model to construct. Basic knowledge of programming is necessary to understand how the toolbox and its utility programs work. In addition, GISTOSWMM's high-resolution outputs could sometimes lead to errors during the catchment definition. For example, since the method uses both land cover data and DEM in the routing procedure, it does not create a sub-catchment that routes water from the street level to rooftops. This assumption, however, may appear reasonable only if the DEM and landcover map matches perfectly. As a result, if a pixel near a house was wrongly mapped with the building landcover type, all of the sub-catchments that went to the pixel would be left out of the final result. For the same reason, this tool did not generate a sub-catchment for central building yards. It is concluded that such a high-resolution model does not provide the model improvements desired. Furthermore, the method introduces more uncertainties and difficulties to the model implementation process.

SCALGO

The SCALGO toolbox provides the modeler with numerous tools for delineating sub-catchments. Using water flow paths to define the contributing drainage area for each point, along with practical changes to the terrain (like culverts and underpasses), can help make the sub-catchment delineation more accurate. With this option, the modeler can define sub-catchments even for networks that are located close to each other. According to further results from case areas, the time for building models was greatly reduced with SCALGO compared to other automated methods. Furthermore, the method's lighter computational burden and shorter simulation time allow model construction for large urban catchments.

Discharge results computed with the SCALGO method were better or not worse than those with the previous HSY method. This is encouraging for relieving the burden associated with the sub-catchment delineation on a large scale. By comparing the different scenarios with SCALGO, practical results were found for stormwater modeling aspects like including stormwater inlets and the size of the sub-catchments, as well as using new imperviousness data.

SCALGO models with high-resolution impervious data (LaserVesi) showed the best overall hydraulic performance of the case areas, but they underestimated the TIA in the Munkkiniemi case area. Using sub-catchments with a minimum size of 5000 m^2

produced good enough results. Therefore, there is no need to use smaller catchments with this method. In addition, inlets can be used to better define catchment boundaries between the main catchment areas. However, there is no need to take into account all the inlets since their inclusion is time-consuming and offers no improvement to the hydraulic analysis at a large scale. As a result, the SCALGO method with a minimum size of 5000 m^2 and an inlet radius of 8 meters was selected as the most suitable method for applying to the Helsinki sewer model. Stormwater inlets and separate stormwater networks can be included to improve simulation results at a local scale.

Spatial resolution

In this study, various spatial resolutions were investigated, ranging from detailed cell-based sub-catchments to merged sub-catchments with the SCALGO toolbox. Models made with SCALGO showed that merged sub-catchments with a minimum size of 5000 m^2 had almost the same volume and peak as sub-catchments with a smaller minimum size. Due to the sensitivity of the SWMM's performance to the width parameter, it is important to choose the right value for a new merged sub-catchment. Consequently, a reliable result could still be reached if the sub-catchment outlet and width parameters were set correctly during the merging process.

It is possible to use a larger minimum sub-catchments size (more than 5000 m^2) for extensive models to optimize stormwater model performance. However, larger sub-catchments may ignore the rainfall's spatial variability and lead to incorrect local flows. Therefore, using large sub-catchments with high-resolution radar rainfall data is worthless since precipitation is assumed to be spatially uniform within each SWMM sub-catchment. When choosing an optimal spatial resolution for the sub-catchments, it is important to consider the resolution of the source data, for example, how detailed the rainfall data is.

Further discussion

SWWM sub-catchment works like a function that, when given rain data, returns the runoff response. Therefore, even though there were some differences in how sub-catchment parameters were selected, the general pattern of the results was similar. This could be justified by the identical rainfall data for sub-catchments. However, the differences in simulated runoff volumes and peak flow between the hydraulic results can be explained by the sub-catchments' total area, total impervious area, flow width, and slope. Slight underestimation of flow was expected because of the known underestimation of hydrograph tails and low flows by SWMM as encountered by previous studies (Guan et al., 2015; Niemi et al., 2019).

SWMM utilizes the Manning equation to define the surface runoff for each calculation time step. Parameter changes influence the dynamics of runoff generation based on their sensitivities. In general, the performance of the methods was at least acceptable, given that parameters from the literature were used with SWMM models that had not been calibrated.

Events with overflows

The high volume of overflow in the Munkkiniemi area limited the number of events that could be meaningfully analyzed. The intense events in this area usually exceeded the capacity of the network, causing overflows at several locations. Due to the flooding and ponding water in the area, pipes may still be carrying stormwater hours after the event. Enabling the ponding area in the model can avoid losing water in overflows by temporarily storing extra water at a particular junction. However, in reality, a part of the stormwater will not reach the outfall, and determining the amount of water overflowing from the manholes is challenging. The application of 2D modeling could reduce uncertainties in such an event, which should be considered if simulations of precise runoff are necessary. However, since the objective of the thesis was to compare the performance of the automated sub-catchment definition, the main conclusion can be drawn from the event without overflow.

Infiltration

Throughout the study, it was repeatedly observed how important a role the pervious areas have on the final hydrographs. While impervious surfaces dictated the volume of discharge for the smaller storm event, naturally impervious areas (i.e., rock) and saturated pervious surfaces generated significant stormflow for the larger event, especially on medium and low-density areas. In addition, a saturation of the pervious area was observed in high-density areas too when there were multiple peaks in precipitation data. While the hydraulic performance of the models was consistent during the first peaks, it was underestimated during the subsequent peaks.

It is possible to achieve more realistic results by allowing the depression storage to fill up and letting pervious surfaces become temporarily saturated in long storm events. The saturation time of the pervious area is determined by the parameters of depression storage and infiltration. The results of the testing in the area showed the potential of a pervious area to improve the result if lower infiltration was used. The reason might be low infiltration in some pervious areas. For example, a green area with naturally rocky parts and naturally impervious areas may have less infiltration capacity. Furthermore, the surfaces located under the trees' canopy might be impervious or have low infiltration. As a

result, a low infiltration parameter should be used for such pervious areas. However, calibration is required because no uniform value can be used in all areas. Because infiltration and calibration are outside the scope of this study, future research could investigate the spatial variability of infiltration parameters in pervious areas.

Precipitation variability

The rainfall data can explain a significant portion of the errors in hydraulic simulation results as it is the most sensitive input into the SWMM model (Arnaud et al., 2011; Niemi, 2017). The current radar rainfall data grid covers a total area of 6.25 ha (250 m by 250 m), while the average and maximum areas of sub-catchments are around 1 ha and 2 ha, respectively. Furthermore, most of the events simulated in this study happened in the summer. A typical feature of summer rains is their spatial-temporal variability. The alteration of rainfall from one *Rain Gage* to another was observed many times during the simulation of the events. Such variation in rainfall data can have a great impact on urban catchments' runoff (Ochoa-Rodriguez et al., 2015). Using higher resolution for the radar data corresponding to the size of the sub-catchment can affect the hydrograph, especially in intense short summer events. Many researchers have stated the effect of rain's spatial variability on hydrological model performance (e.g., Berne et al., 2004; Krajewski et al., 2003). Furthermore, the temporal resolution of rainfall was among the most important parameters reported by previous authors (Berne et al., 2004; Bruni et al., 2015; Ochoa-Rodriguez et al., 2015). Schilling (1991) argues that a temporal resolution of one minute is required for urban hydrological model applications. Another study by Ochoa-Rodriguez et al. (2015) suggested that finer resolution data (i.e., 1–5 minutes) is necessary to accurately capture the variability of precipitation data. Future studies can look at how spatial and temporal resolution affects hydraulic results and CSOs.

5.4 Modeling results of the Helsinki combined sewer network

The catchment definition toolbox developed by SCALGO was applied to the CS area of the Helsinki sewer model to test the feasibility of the approach at a large scale. Sub-catchment delineation for the Helsinki sewer model was more challenging compared to the case areas where importing the pipes and nodes to the toolbox sufficed to delineate the sub-catchments. The city of Helsinki has both combined sewer and separate storm-water networks. The city center uses mostly the old, CS pipes. However, as the development is ongoing in the city center, some areas already have separate sewer and storm-water networks in which stormwater is collected and conveyed to the Baltic Sea. Figure 19 shows a part of the Helsinki sewer network located close to the sea, where both types of network work to drain the stormwater.

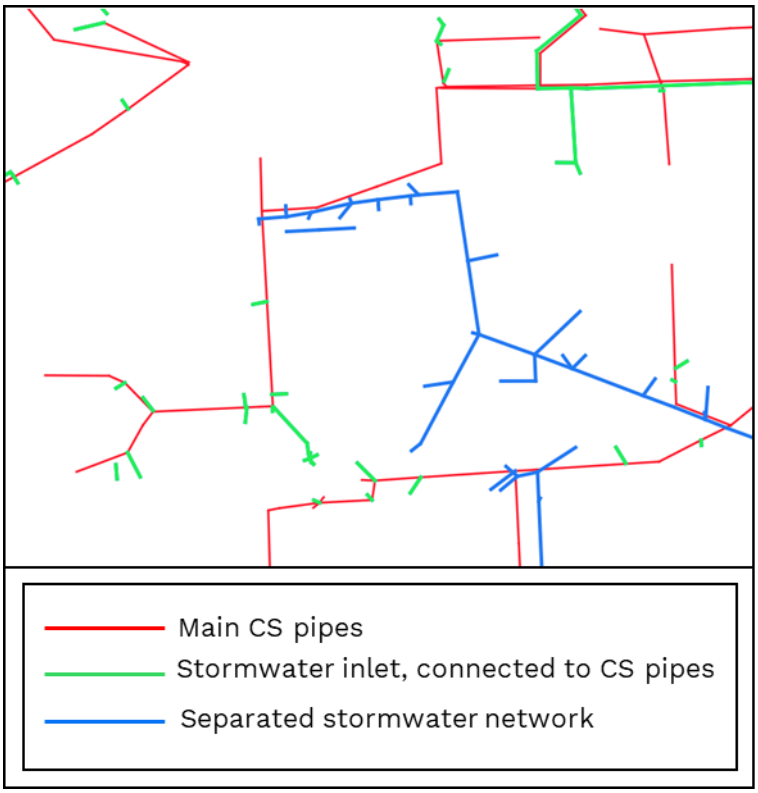


Figure 19. *An example of complex Helsinki sewer network.*

In some areas, like that shown in Figure 19, it was hard to define the area draining to the combined sewer network. Accounting for stormwater pipes is critical for achieving a realistic presentation of sub-catchments. However, only the main CS areas exist in the Helsinki sewer model, shown with the red line.

Parallel Model

According to the need for stormwater pipes in the catchment definition, such pipes must have been added to the model from the shape files. In addition, more changes were needed in the model, including 1) many other components and complexities, such as pumps, weirs, and overflow pipes, that had to be simplified for the toolbox. 2) The model had more than 7000 nodes, each of which tagged either an inlet or junction. 3) Some assumed pipes needed to be added in order to generate sub-catchments for large private properties without network data.

To have all such changes and pre-processing in one place and keep the main model untouched, a copy of the main model was created as a child scenario, called a "parallel model." Thus, the parallel model was safe to modify or add components as it would not be used for simulation. This model can be used in the future for updating purposes.

Delineation report

Despite the network's complexity, the SCALGO toolbox was successful in generating sub-catchments for an extensive city-scale model using the manual alterations done in the “parallel model”. Since there was no information about how long it took to manually build the old sub-catchments in the Helsinki CS sewer model, it was not possible to compare building times between the old and new sub-catchments. But by comparing models from the testing area, a rough estimate was made that the time it took to build a model with SCALGO was at least five times shorter than the old method. Consequently, the manual sub-catchment definition for the city-scale model, which can take several weeks, only took a couple of days with SCALGO.

Figure 20 depicts the model's new and old sub-catchments, as well as the location of the four case areas. The new catchment definition method resulted in a total of 1400 sub-catchments, connected to the combined sewer network. The number has increased by 40% compared to the old catchments (933). Moreover, the total area of the sub-catchment decreased from 1703 ha to 1580 ha. This was because some areas that were thought to be linked to the separated storm network were left out (suggested by HSY). The TIA of the new sub-catchments was 932 ha (59% of the total area), while the old sub-catchments had 1071 ha (62% of the total area). While the TIA decreased due to a reduction in the total area, the TIA percentage remained almost constant, showing the similarity of the imperviousness calculation of the method on a large scale. Based on new sub-catchments and their land cover data, in the Helsinki combined sewer area, buildings account for the largest fraction with 23.2%, followed by roads and forested areas with 22.1% and 20.7%, respectively.

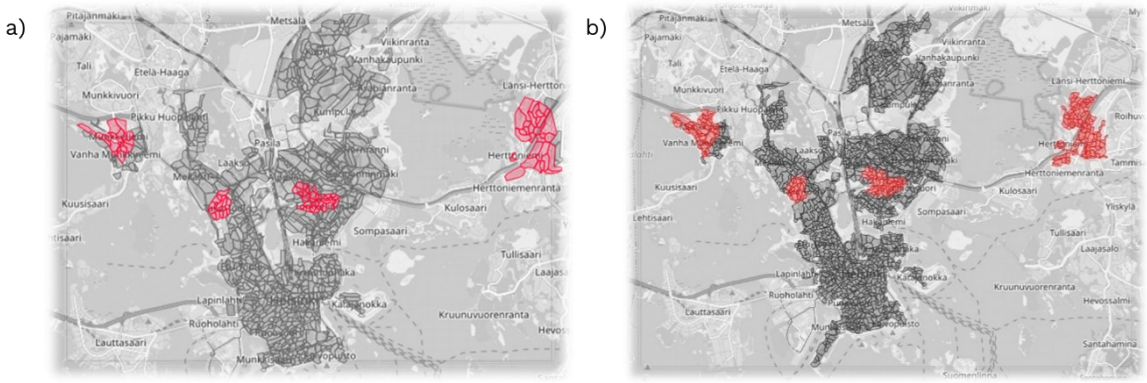


Figure 20. *Sub-catchment delineation for the Helsinki sewer network. (a) old sub-catchments in the HSY model; (b) new sub-catchments via SCALGO.*

Analysis

The Helsinki sewer model was simulated with new sub-catchments for a 3-month simulation from June 1, 2021, until October 1, 2021. First, general results were presented, and then four main events in which overflows happened were evaluated. Similar to the case areas, an increase in the number of sub-catchments did not affect the simulation time. Using the model settings and hardware described in Section 4.6, the simulation took 10 hours for both new and old sub-catchments. Minor changes were anticipated since hydrological calculation requires less calculation time, and the total number of sub-catchments is insignificant compared to the total number of model components (26,000). However, the size of the model result almost increased by 4%.

A comparison between simulated and measured values evaluated at the model's outlet is depicted in Figure 21. Measurement is located before the Viikinmäki WWTP. Similar patterns between the old and new sub-catchments show the new automated method can have a reliable result, as good as the previous method with manual work.

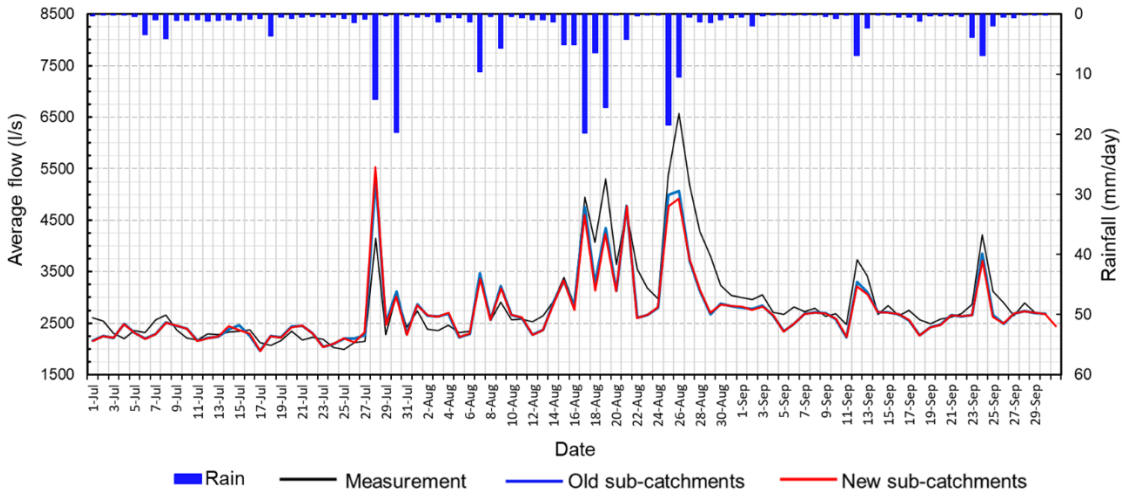


Figure 21. *Average inflow at Viikinmäki's WWTP.*

CSOs

The results at the Viikinmäki WWTP, where tunnels attenuate the effect of runoff, could only show general information about the performance of the sub-catchments. Therefore, to compare the effect of sub-catchments on flow in the main lines and CSOs, four events with the most CSOs were selected from the same quarter in 2021. The flows in the two main pumping stations in the Helsinki combined sewer area are shown in Figure 22. After updating the sub-catchments, the hydraulic results at pumping stations remained consistent, achieving the same or better results than the old, time-consuming manual process.

Furthermore, three main overflow locations were selected to evaluate the magnitude of overflows. Figure 23 compares the simulated values of the overflows for new and old sub-catchments. A significant decrease in overflow volume was observed with the new sub-catchments. The results from overflow locations and pumping stations show that the peaks in the old sub-catchments were steep, whereas the peaks in the new sub-catchments were reduced, resulting in a lower overflow volume.

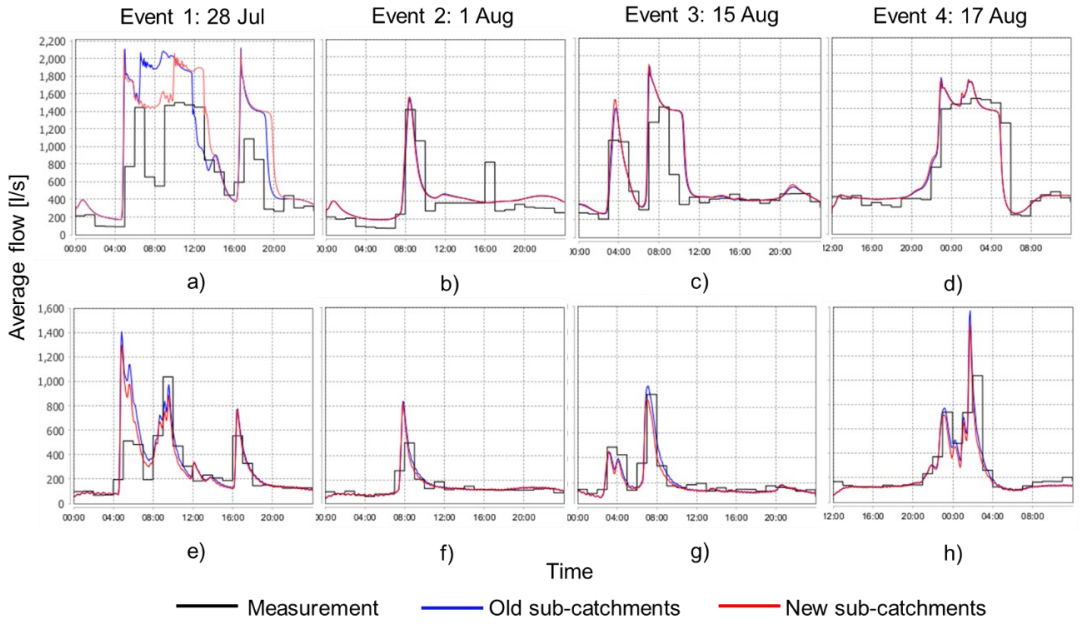


Figure 22. *Comparison of measured (1-hour average) vs. simulated flows at Mäntymäki pumping station: (a); (b); (c); and (d) and at Siltavuori pumping station: (e); (f); (g); and (h).*

As a result of using automated sub-catchment delineation, there has been a significant increase in the speed of model development. The method has been found to be useful in developing hydrological models ranging from small to city-scale. For complicated areas, like the Helsinki sewer system, it saves a lot of time while giving the expected hydraulic performance. The process still requires creating a "parallel model" for simplifications prior to importing the network to the SCALGO toolbox. Using advanced technologies and automated methods has great potential to improve the accuracy of stormwater, combined sewer, and online models in the future. Furthermore, updating the model's sub-catchments can be done very quickly and in a systematic manner, especially after the first round of implementation.

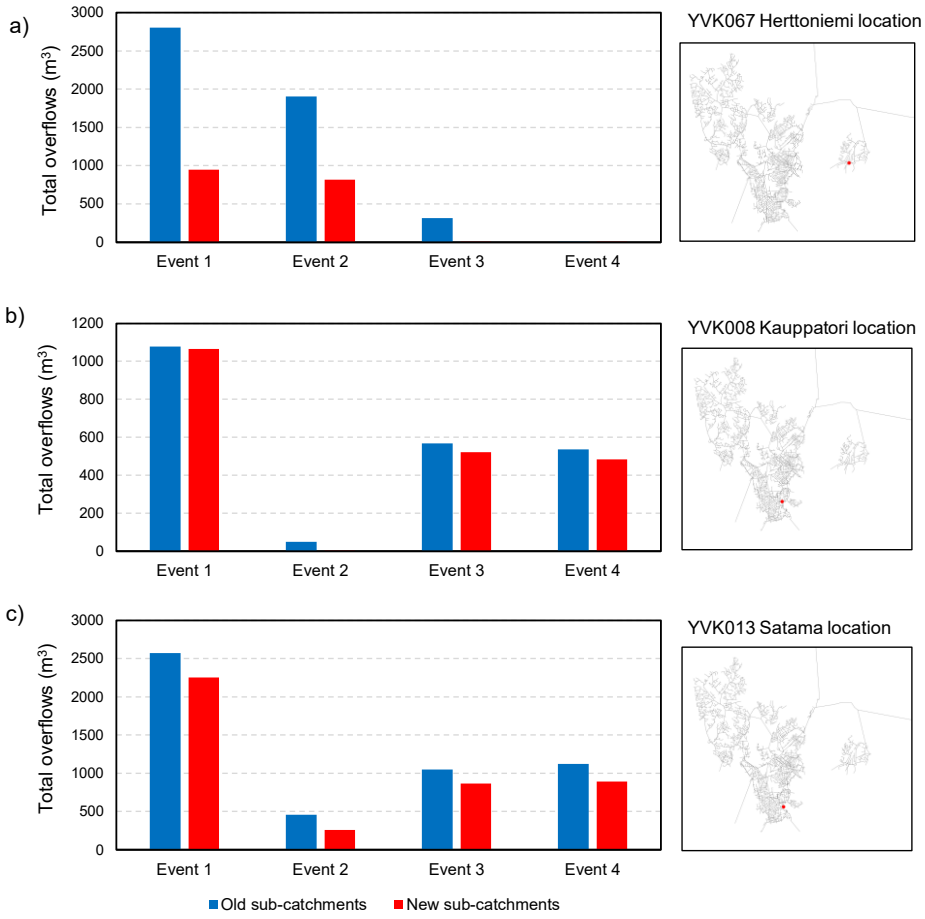


Figure 23. *Comparison of simulated total overflows in each rain event with old and new sub-catchments at three overflow locations (a) YVK067 Herttoniemi; (b) YVK008 Kauppatori; (c) YVK013 Satama.*

6. CONCLUSIONS

The aim of the thesis was to assess the automated delineation and parametrization of SWMM sub-catchments. Determining the model's practical level of details and spatial resolution, the automated methods' manual work reduction, and related processing time were the main challenges addressed in this work. The process was approached in two main steps. First, the application of the four catchment definition methods was investigated in four case areas. Then, based on the comparison of methods, one automated method was used to test its capabilities in the city-scale network. From these two steps, the following key conclusions can be drawn:

In four case areas that ranged from 22 ha to 55 ha, three automated catchment definitions were implemented and compared to the old method. Consequently, SCALGO was found to satisfy the demands that were investigated in the first step and selected to be used in a city-scale model.

Further practical results for stormwater modeling aspects were found in SCALGO scenarios. Firstly, the impacts of including inlets in the hydraulic results were found to be minor and can only be used in critical local places in extensive models. Secondly, using a minimum sub-catchment area of 5000 m² was deemed a practical and efficient choice for the case areas as the runoff volumes and peak flows remained consistent with that of smaller catchments. Finally, a new high-resolution imperviousness layer (LaserVesi) could be used to make a good estimate of the TIA in all case areas except Munkkiniemi.

In the second step, the SCALGO catchment toolbox was found successful in generating sub-catchments for the Helsinki city-scale CS model with a complex network. This was validated in the hydraulic model where results remained consistent after replacing old sub-catchments with the new sub-catchments. As a result, the method has been found to be useful in developing hydrological models ranging from small to city-scale. Even for complicated areas, like the Helsinki sewer system, it saves a lot of time while giving the expected hydraulic performance. The process still requires creating a "parallel model" for simplifications prior to importing the network into SCALGO Toolbox. Updating catchments can be done very quickly and in a systematic manner, especially after the first round of catchment definition.

In conclusion, the thesis has successfully served as a foundational analysis of the use of advanced technologies and openly available data in automated catchment definition. The results demonstrate an approach for improving Fluidit Storm accuracy and speed of

building SWMM stormwater or combined sewer models in the future. Based on the findings of this study, further research could be conducted to study the effect of the spatial distribution of infiltration parameters and precipitation to improve the spatial–temporal variability of rainfall-runoff modeling.

REFERENCES

Aaltonen, J., Hohti, H., Jylhä, K., Karvonen, T., Kilpeläinen, T., Koistinen, J., Kotro, J., Kuitunen, T., Ollila, M., Parvio, A., Pulkkinen, S., Silander, J., Tiihonen, T., Tuomenvirta, H., & Vajda, A. (2008). *Rankkasateet ja taajamatulvat (RATU), Suomen ympäristö*.

Ahm, M., Thorndahl, S., Nielsen, J. E., & Rasmussen, M. R. (2016). Estimation of combined sewer overflow discharge: a software sensor approach based on local water level measurements. *Water Science and Technology*, *74*(11), 2683–2696. https://doi.org/10.2166/wst.2016.361

Al Aukidy, M., & Verlicchi, P. (2017). Contributions of combined sewer overflows and treated effluents to the bacterial load released into a coastal area. *Science of The Total Environment*, *607–608*, 483–496. https://doi.org/10.1016/j.scitotenv.2017.07.050

Almeida Silva, P. (2019). *Urban sanitary sewer modelling in cold climate*. EuroAquae - Brandenburg Technical University.

Anees, M. T., Abu Bakar, A. F. Bin, San, L. H., Abdullah, K., Nordin, M. N. M., Rahman, N. N. N. A., Ishak, M. I. S., & Kadir, M. O. A. (2020). *Flood Vulnerability, Risk, and Susceptibility Assessment* (pp. 1–27). https://doi.org/10.4018/978-1-5225-9771-1.ch001

Arnaud, P., Lavabre, J., Fouchier, C., Diss, S., & Javelle, P. (2011). Sensitivity of hydrological models to uncertainty in rainfall input. *Hydrological Sciences Journal*, *56*(3), 397–410. https://doi.org/10.1080/02626667.2011.563742

Arnell, V. (1980). *Description and validation of the CTH-urban runoff model*. Chalmers University of Technology.

Autixier, L., Mailhot, A., Bolduc, S., Madoux-Humery, A.-S., Galarneau, M., Prévost, M., & Dorner, S. (2014). Evaluating rain gardens as a method to reduce the impact of sewer overflows in sources of drinking water. *Science of The Total Environment*, *499*, 238–247. https://doi.org/10.1016/j.scitotenv.2014.08.030

Bach, P., McCarthy, D., Urich, C., Sitzenfrei, R., Kleidorfer, M., Rauch, W., & Deletic, A. (2013). A planning algorithm for quantifying decentralised water management opportunities in urban environments. *Water Science and Technology*, *68*(8), 1857–1865. https://doi.org/10.2166/wst.2013.437

Barbosa, A. E., Fernandes, J. N., & David, L. M. (2012). Key issues for sustainable urban stormwater management. *Water Research*, *46*(20), 6787–6798. https://doi.org/10.1016/j.watres.2012.05.029

Bareš, V., Jirák, J., & Pollert, J. (2008). Spatial and temporal variation of turbulence characteristics in combined sewer flow. *Flow Measurement and Instrumentation*, *19*(3–4), 145–154. https://doi.org/10.1016/j.flowmeasinst.2007.06.002

Bedient, P., Huber, W., & Vieux. (2008). *Hydrology and floodplain analysis*. Upper Saddle River, NJ: Prentice Hall.

Bernard, J. M., & Tuttle, R. W. (1998). Stream Corridor Restoration: Principles, Processes, and Practices. *Engineering Approaches to Ecosystem Restoration*, 320–325. https://doi.org/10.1061/40382(1998)55

Berne, A., Delrieu, G., Creutin, J.-D., & Obled, C. (2004). Temporal and spatial resolution of rainfall measurements required for urban hydrology. *Journal of Hydrology*, *299*(3–4), 166–179. https://doi.org/10.1016/j.jhydrol.2004.08.002

Bren, L. (2015). *Forest hydrology and catchment management*. Dordrecht: Springer Netherlands.

Bruni, G., Reinoso, R., van de Giesen, N. C., Clemens, F. H. L. R., & ten Veldhuis, J. A. E. (2015). On the sensitivity of urban hydrodynamic modelling to rainfall spatial and temporal resolution. *Hydrology and Earth System Sciences*, *19*(2), 691–709. https://doi.org/10.5194/hess-19-691-2015

Chang, Q., Kazama, S., Touge, Y., & Aita, S. (2019). The effects of spatial discretization on performances and parameters of urban hydrological model. *Water Science and Technology*, *80*(3), 517–528. https://doi.org/10.2166/wst.2019.296

Chang, T., Wang, C., Chen, A., & Djordjević, S. (2018). The effect of inclusion of inlets in dual drainage modelling. *Journal of Hydrology*, *559*, 541–555. https://doi.org/10.1016/j.jhydrol.2018.01.066

Chen, Y., Zhou, H., Zhang, H., Du, G., & Zhou, J. (2015). Urban flood risk warning under rapid urbanization. *Environmental Research*, *139*, 3–10. https://doi.org/10.1016/j.envres.2015.02.028

Chen, Z., & Wang, J. (2010). Land use and land cover change detection using satellite remote sensing techniques in the mountainous Three Gorges Area, China. *International Journal of Remote Sensing*, *31*(6), 1519–1542. https://doi.org/10.1080/01431160903475381

Choi, K., & Ball, J. E. (2002). Parameter estimation for urban runoff modelling. *Urban Water*, *4*(1), 31–41. https://doi.org/10.1016/S1462-0758(01)00072-3

Cifelli, R., & Chandrasekar, V. (2010). *Dual-polarization radar rainfall estimation* (pp. 105–125). https://doi.org/10.1029/2010GM000930

De Feo, G., Antoniou, G., Fardin, H., El-Gohary, F., Zheng, X., Reklaityte, I., Butler, D., Yannopoulos, S., & Angelakis, A. (2014). The Historical Development of Sewers Worldwide. *Sustainability*, *6*(6), 3936–3974. https://doi.org/10.3390/su6063936

Dell, T., Razzaghmanesh, M., Sharvelle, S., & Arabi, M. (2021). Development and Application of a SWMM-Based Simulation Model for Municipal Scale Hydrologic Assessments. *Water*, *13*(12), 1644. https://doi.org/10.3390/w13121644

Dietz, M. E. (2007). Low Impact Development Practices: A Review of Current Research and Recommendations for Future Directions. *Water, Air, and Soil Pollution*, *186*(1–4), 351–363. https://doi.org/10.1007/s11270-007-9484-z

Dingman, S. L. (2015). Physical hydrology. In *Progress in Physical Geography* (Vol. 13, Issue 1). https://doi.org/10.1177/030913338901300106

Doney, B. (1999). *Electromagnetic flow measurement in partially filled pipes* (32-34. 146(11) (ed.)). Water Engineering & Management.

Dongquan, Z., Jining, C., Haozheng, W., Qingyuan, T., Shangbing, C., & Zheng, S. (2009). GIS-based urban rainfall-runoff modeling using an automatic catchment-discretization approach: a case study in Macau. *Environmental Earth Sciences*, *59*(2), 465–472. https://doi.org/10.1007/s12665-009-0045-1

Dou, Y., & Kuang, W. (2020). A comparative analysis of urban impervious surface and green space and their dynamics among 318 different size cities in China in the past 25 years. *Science of The Total Environment*, *706*, 135828. https://doi.org/10.1016/j.scitotenv.2019.135828

Dow, C. L., & DeWalle, D. R. (2000). Trends in evaporation and Bowen Ratio on urbanizing watersheds in eastern United States. *Water Resources Research*, *36*(7), 1835–1843. https://doi.org/10.1029/2000WR900062

Ebrahimian, A., Wilson, B. N., & Gulliver, J. S. (2016). Improved methods to estimate the effective impervious area in urban catchments using rainfall-runoff data. *Journal of Hydrology*, *536*, 109–118. https://doi.org/10.1016/j.jhydrol.2016.02.023

EEA. (20177). *Copernicus Land Monitoring Service - Urban Atlas*. https://www.eea.europa.eu/data-and-maps/data/copernicus-land-monitoring-service-urban-atlas

Fletcher, T. D., Andrieu, H., & Hamel, P. (2013). Understanding, management and modelling of urban hydrology and its consequences for receiving waters: A state of the art. *Advances in Water Resources*, *51*, 261–279. https://doi.org/10.1016/j.advwatres.2012.09.001

Fu, X., Goddard, H., Wang, X., & Hopton, M. E. (2019). Development of a scenario-based stormwater management planning support system for reducing combined sewer overflows (CSOs). *Journal of Environmental Management*, *236*, 571–580. https://doi.org/10.1016/j.jenvman.2018.12.089

Garofalo, G., Giordano, A., Piro, P., Spezzano, G., & Vinci, A. (2017). A distributed real-time approach for mitigating CSO and flooding in urban drainage systems. *Journal of Network and Computer Applications*, *78*, 30–42. https://doi.org/10.1016/j.jnca.2016.11.004

Ghosh, I., & Hellweger, F. L. (2012). Effects of Spatial Resolution in Urban Hydrologic Simulations. *Journal of Hydrologic Engineering*, *17*(1), 129–137. https://doi.org/10.1061/(ASCE)HE.1943-5584.0000405

Gironás, J., Niemann, J. D., Roesner, L. A., Rodriguez, F., & Andrieu, H. (2010). Evaluation of Methods for Representing Urban Terrain in Storm-Water Modeling. *Journal of Hydrologic Engineering*, *15*(1), 1–14. https://doi.org/10.1061/(ASCE)HE.1943-5584.0000142

Godley, A. (2002). Flow measurement in partially filled closed conduits. *Flow Measurement and Instrumentation*, *13*(5–6), 197–201. https://doi.org/10.1016/S0955-5986(02)00050-X

Gregow, E., Pessi, A., Mäkelä, A., & Saltikoff, E. (2017). Improving the precipitation accumulation analysis using lightning measurements and different integration periods. *Hydrology and Earth System Sciences*, *21*(1), 267–279. https://doi.org/10.5194/hess-21-267-2017

Guan, M., Sillanpää, N., & Koivusalo, H. (2015). Storm runoff response to rainfall pattern, magnitude and urbanization in a developing urban catchment. *Hydrological Processes*, n/a-n/a. https://doi.org/10.1002/hyp.10624

Heber Green, W., & Ampt, G. A. (1911). Studies on Soil Phyics. *The Journal of Agricultural Science*, *4*(1), 1–24. https://doi.org/10.1017/S0021859600001441

Horton, R. E. (1933). The Rôle of infiltration in the hydrologic cycle. *Transactions, American Geophysical Union*, *14*(1), 446. https://doi.org/10.1029/TR014i001p00446

Huang, F., Yu, Y., & Feng, T. (2019). Automatic extraction of urban impervious surfaces based on deep learning and multi-source remote sensing data. *Journal of Visual Communication and Image Representation*, *60*, 16–27. https://doi.org/10.1016/j.jvcir.2018.12.051

Humphrey, M. D., Istok, J. D., Lee, J. Y., Hevesi, J. A., & Flint, A. L. (1997). A New Method for Automated Dynamic Calibration of Tipping-Bucket Rain Gauges. *Journal of Atmospheric and Oceanic Technology*, *14*(6), 1513–1519. https://doi.org/10.1175/1520-0426(1997)014<1513:ANMFAD>2.0.CO;2

Hvitved-Jacobsen, T., Vollertsen, J., & Haaning Nielsen, A. (2010). *Urban and Highway Stormwater Pollution*. CRC Press. https://doi.org/10.1201/9781439826867

Jain, G. V., Agrawal, R., Bhanderi, R. J., Jayaprasad, P., Patel, J. N., Agnihotri, P. G., & Samtani, B. M. (2016). Estimation of sub-catchment area parameters for Storm Water Management Model (SWMM) using geo-informatics. *Geocarto International*, *31*(4), 462–476. https://doi.org/10.1080/10106049.2015.1054443

Jang, J., Chang, T., & Chen, W. (2018). Effect of inlet modelling on surface drainage in coupled urban flood simulation. *Journal of Hydrology*, *562*, 168–180. https://doi.org/10.1016/j.jhydrol.2018.05.010

Jankowfsky, S., Branger, F., Braud, I., Gironás, J., & Rodriguez, F. (2013). Comparison of catchment and network delineation approaches in complex suburban environments: application to the Chaudanne catchment, France. *Hydrological Processes*, *27*(25), 3747–3761. https://doi.org/10.1002/hyp.9506

Ji, S., & Qiuwen, Z. (2015). A GIS-based Subcatchments Division Approach for SWMM. *The Open Civil Engineering Journal*, *9*(1), 515–521. https://doi.org/10.2174/1874149501509010515

Kidd, C., Becker, A., Huffman, G. J., Muller, C. L., Joe, P., Skofronick-Jackson, G., & Kirschbaum, D. B. (2017). So, How Much of the Earth's Surface Is Covered by Rain Gauges? *Bulletin of the American Meteorological Society*, *98*(1), 69–78. https://doi.org/10.1175/BAMS-D-14-00283.1

Kokkonen, T., Warsta, L., Niemi, T. J., Taka, M., Sillanpää, N., Pusa, M., Kesäniemi, O., Salo, H., & Koivusalo, H. (2019). Impact of alternative land cover descriptions on urban hydrological model simulations. *Urban Water Journal*, *16*(2), 103–113. https://doi.org/10.1080/1573062X.2019.1634742

Kong, F., Ban, Y., Yin, H., James, P., & Dronova, I. (2017). Modeling stormwater management at the city district level in response to changes in land use and low impact development. *Environmental Modelling & Software*, *95*, 132–142.

https://doi.org/10.1016/j.envsoft.2017.06.021

Kostarelos, K., Khan, E., Callipo, N., & Velasquez, J. (2011). Erratum to: Field Study of Catch Basin Inserts for the Removal of Pollutants from Urban Runoff. *Water Resources Management*, *25*(4), 1251–1251. https://doi.org/10.1007/s11269-010-9703-z

Krajewski, W. F., Ciach, G. J., & Habib, E. (2003). An analysis of small-scale rainfall variability in different climatic regimes. *Hydrological Sciences Journal*, *48*(2), 151–162. https://doi.org/10.1623/hysj.48.2.151.44694

Krebs, G., Kokkonen, T., Setälä, H., & Koivusalo, H. (2016). Parameterization of a Hydrological Model for a Large, Ungauged Urban Catchment. *Water*, *8*(10), 443. https://doi.org/10.3390/w8100443

Krebs, G., Kokkonen, T., Valtanen, M., Setälä, H., & Koivusalo, H. (2014). Spatial resolution considerations for urban hydrological modelling. *Journal of Hydrology*, *512*, 482–497. https://doi.org/10.1016/j.jhydrol.2014.03.013

LaserVesi Project, S. (2022). *Mapping imperviousness for Greater Helsinki Region*. https://avoidatastr.blob.core.windows.net/avoindata/AvoinData/9_Kartat/Maanpeiterasterit/LaserVesi Documentation.pdf

Lee, K., Kim, H., Pak, G., Jang, S., Kim, L., Yoo, C., Yun, Z., & Yoon, J. (2010). Cost-effectiveness analysis of stormwater best management practices (BMPs) in urban watersheds. *Desalination and Water Treatment*, *19*(1–3), 92–96. https://doi.org/10.5004/dwt.2010.1900

Lehtinen, S. (2014). *Simulation of Stormwater Quality in an Urban Catchment Using the Stormwater Management Model (Swmm)*. Aalto University.

Lhomme, J., Bouvier, C., & Perrin, J.-L. (2004). Applying a GIS-based geomorphological routing model in urban catchments. *Journal of Hydrology*, *299*(3–4), 203–216. https://doi.org/10.1016/j.jhydrol.2004.08.006

Li, Z., Huang, Q., Duan, Y., Chen, W., & Zou, L. (2020). Research on Electromagnetic Flowmeter Based on Double-frequency Trapezoidal Wave Excitation. *Journal of Physics: Conference Series*, *1549*(5), 052086. https://doi.org/10.1088/1742-6596/1549/5/052086

Liu, T., Lawluvy, Y., Shi, Y., & Yap, P.-S. (2021). Low Impact Development (LID) Practices: A Review on Recent Developments, Challenges and Prospects. *Water, Air, & Soil Pollution*, *232*(9), 344. https://doi.org/10.1007/s11270-021-05262-5

Madrazo-Uribeetxebarria, E., Garmendia Antín, M., Almandoz Berrondo, J., & Andrés-Doménech, I. (2021). Sensitivity analysis of permeable pavement hydrological modelling in the Storm Water Management Model. *Journal of Hydrology*, *600*, 126525. https://doi.org/10.1016/j.jhydrol.2021.126525

Martin, C., Ruperd, Y., & Legret, M. (2007). Urban stormwater drainage management: The development of a multicriteria decision aid approach for best management practices. *European Journal of Operational Research*, *181*(1), 338–349. https://doi.org/10.1016/j.ejor.2006.06.019

Martin, P. H., LeBoeuf, E. J., Dobbins, J. P., Daniel, E. B., & Abkowitz, M. D. (2005). INTERFACING GIS WITH WATER RESOURCE MODELS: A STATE-OF-THE-

ART REVIEW. *Journal of the American Water Resources Association*, *41*(6), 1471–1487. https://doi.org/10.1111/j.1752-1688.2005.tb03813.x

Mbajiorgu, C. C. (1995). Watershed resources management (WRM) model 1. Model description. *Computers and Electronics in Agriculture*, *13*(3), 195–216. https://doi.org/10.1016/0168-1699(95)00011-R

Mirza, M. (2003). Climate change and extreme weather events: can developing countries adapt? *Climate Policy*, *3*(3), 233–248. https://doi.org/10.3763/cpol.2003.0330

Mnih, V. (2013). *Machine learning for aerial image labeling* [University of Toronto (Canada)]. https://www.proquest.com/openview/215c1c8c5424a6facf8885f4ac00c575/1?pq-origsite=gscholar&cbl=18750

Montserrat, A., Bosch, L., Kiser, M. A., Poch, M., & Corominas, L. (2015). Using data from monitoring combined sewer overflows to assess, improve, and maintain combined sewer systems. *Science of The Total Environment*, *505*, 1053–1061. https://doi.org/10.1016/j.scitotenv.2014.10.087

Moulin, L., Gaume, E., & Obled, C. (2009). Uncertainties on mean areal precipitation: assessment and impact on streamflow simulations. *Hydrology and Earth System Sciences*, *13*(2), 99–114. https://doi.org/10.5194/hess-13-99-2009

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

Niemi, T. (2017). *Improved Precipitation Information for Hydrological Problem Solving - Focus on Open Data and Simulation*. Aalto University.

Niemi, T., Kokkonen, T., Sillanpää, N., Setälä, H., & Koivusalo, H. (2019). Automated Urban Rainfall–Runoff Model Generation with Detailed Land Cover and Flow Routing. *Journal of Hydrologic Engineering*, *24*(5), 04019011. https://doi.org/10.1061/(ASCE)HE.1943-5584.0001784

Nowogoński, I., Ogiołda, E., & Musielak, M. (2019). Estimation of the Hydraulic Width of the Subcatchment Depending on the Degree of Detail of the Drainage System Model. *Civil and Environmental Engineering Reports*, *29*(4), 128–140. https://doi.org/10.2478/ceer-2019-0049

O'Loughlin, G., Huber, W., & Chocat, B. (1996). Rainfall-runoff processes and modelling. *Journal of Hydraulic Research*, *34*(6), 733–751. https://doi.org/10.1080/00221689609498447

Ochoa-Rodriguez, S., Wang, L.-P., Gires, A., Pina, R. D., Reinoso-Rondinel, R., Bruni, G., Ichiba, A., Gaitan, S., Cristiano, E., van Assel, J., Kroll, S., Murlà-Tuyls, D., Tisserand, B., Schertzer, D., Tchiguirinskaia, I., Onof, C., Willems, P., & ten Veldhuis, M.-C. (2015). Impact of spatial and temporal resolution of rainfall inputs on urban hydrodynamic modelling outputs: A multi-catchment investigation. *Journal of Hydrology*, *531*, 389–407. https://doi.org/10.1016/j.jhydrol.2015.05.035

Pan, S., Guan, H., Chen, Y., Yu, Y., Nunes Gonçalves, W., Marcato Junior, J., & Li, J. (2020). Land-cover classification of multispectral LiDAR data using CNN with optimized hyper-parameters. *ISPRS Journal of Photogrammetry and Remote*

Sensing, *166*, 241–254. https://doi.org/10.1016/j.isprsjprs.2020.05.022

Peleg, N., Ben-Asher, M., & Morin, E. (2013). Radar subpixel-scale rainfall variability and uncertainty: lessons learned from observations of a dense rain-gauge network. *Hydrology and Earth System Sciences*, *17*(6), 2195–2208. https://doi.org/10.5194/hess-17-2195-2013

Peterson, E. W., & Wicks, C. M. (2006). Assessing the importance of conduit geometry and physical parameters in karst systems using the storm water management model (SWMM). *Journal of Hydrology*, *329*(1–2), 294–305. https://doi.org/10.1016/j.jhydrol.2006.02.017

Petrucci, G., & Bonhomme, C. (2014). The dilemma of spatial representation for urban hydrology semi-distributed modelling: Trade-offs among complexity, calibration and geographical data. *Journal of Hydrology*, *517*, 997–1007. https://doi.org/10.1016/j.jhydrol.2014.06.019

Pirinen, P., Simola, H., Aalto, J., Kaukoranta, J., Karlsson, P., & Ruuhela, R. (2012). *Climatological statistics of Finland 1981–2010, Reports.* http://hdl.handle.net/10138/35880

Podobnikar, T., Stancic, Z., & Oštir, K. (2000). Data integration for the DTM production. *International Cooperation and Technology Transfer, Proceedings of the Workshop*.

Poff, N. L. (2002). Ecological response to and management of increased flooding caused by climate change. *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, *360*(1796), 1497–1510. https://doi.org/10.1098/rsta.2002.1012

QGIS Development Team. (2022). *QGIS geographic information system*. Open Source Geospatial Foundation Project. http://qgis.osgeo.org

Quevauviller, P., Thomas, O., & Van der Beken, A. (2007). *Wastewater quality monitoring and treatment*. John Wiley & Sons.

Radinja, M., Comas, J., Corominas, L., & Atanasova, N. (2019). Assessing stormwater control measures using modelling and a multi-criteria approach. *Journal of Environmental Management*, *243*, 257–268. https://doi.org/10.1016/j.jenvman.2019.04.102

Ray, L. K. (2018). Limitation of automatic watershed delineation tools in coastal region. *Annals of GIS*, *24*(4), 261–274. https://doi.org/10.1080/19475683.2018.1526212

Ren, R., Wang, H., Sun, X., & Quan, H. (2022). Design and Implementation of an Ultrasonic Flowmeter Based on the Cross-Correlation Method. *Sensors*, *22*(19), 7470. https://doi.org/10.3390/s22197470

Rossman. (2016). *Storm Water Management Model Reference Manual Volume I – Hydrology*. U.S. Environmental Protection Agency.

Rossman. (2017). *Storm Water Management Model Reference Manual Volume II – Hydraulics*. U.S. Environmental Protection Agency.

Salih, A. A. M., & Hamid, A. A. (2017). Hydrological studies in the White Nile State in Sudan. *The Egyptian Journal of Remote Sensing and Space Science*, *20*, S31–S38. https://doi.org/10.1016/j.ejrs.2016.12.004

Salvadore, E., Bronders, J., & Batelaan, O. (2015). Hydrological modelling of urbanized catchments: A review and future directions. *Journal of Hydrology*, *529*, 62–81. https://doi.org/10.1016/j.jhydrol.2015.06.028

Sanz-Ramos, M., Bladé, E., González-Escalona, F., Olivares, G., & Aragón-Hernández, J. L. (2021). Interpreting the Manning Roughness Coefficient in Overland Flow Simulations with Coupled Hydrological-Hydraulic Distributed Models. *Water*, *13*(23), 3433. https://doi.org/10.3390/w13233433

Sanzana, P., Gironás, J., Braud, I., Branger, F., Rodriguez, F., Vargas, X., Hitschfeld, N., Muñoz, J. F., Vicuña, S., Mejía, A., & Jankowfsky, S. (2017). A GIS-based urban and peri-urban landscape representation toolbox for hydrological distributed modeling. *Environmental Modelling & Software*, *91*, 168–185. https://doi.org/10.1016/j.envsoft.2017.01.022

Saunders, W. (2000). Preparation of DEMs for use in environmental modeling analysis. *Hydrologic and Hydraulic Modeling Support*, 29–51.

Schilling, W. (1991). Rainfall data for urban hydrology: what do we need? *Atmospheric Research*, *27*(1–3), 5–21. https://doi.org/10.1016/0169-8095(91)90003-F

Semadeni-Davies, A., Hernebring, C., Svensson, G., & Gustafsson, L.-G. (2008). The impacts of climate change and urbanisation on drainage in Helsingborg, Sweden: Combined sewer system. *Journal of Hydrology*, *350*(1–2), 100–113. https://doi.org/10.1016/j.jhydrol.2007.05.028

Seo, D.-J., Seed, A., & Delrieu, G. (2010). *Radar and multisensor rainfall estimation for hydrologic applications* (pp. 79–104). https://doi.org/10.1029/2010GM000952

Shahed Behrouz, M., Zhu, Z., Matott, L. S., & Rabideau, A. J. (2020). A new tool for automatic calibration of the Storm Water Management Model (SWMM). *Journal of Hydrology*, *581*, 124436. https://doi.org/10.1016/j.jhydrol.2019.124436

Sheng, J., & Wilson, J. P. (2009). Watershed urbanization and changing flood behavior across the Los Angeles metropolitan region. *Natural Hazards*, *48*(1), 41–57. https://doi.org/10.1007/s11069-008-9241-7

Shuster, W. D., Bonta, J., Thurston, H., Warnemuende, E., & Smith, D. R. (2005). Impacts of impervious surface on watershed hydrology: A review. *Urban Water Journal*, *2*(4), 263–275. https://doi.org/10.1080/15730620500386529

Sillanpää, N. (2013). *Effects of suburban development on runoff generation and water quality* [Aalto University]. http://urn.fi/URN:ISBN:978-952-60-5374-5

Smith, J. A., Baeck, M. L., Zhang, Y., & Doswell, C. A. (2001). Extreme Rainfall and Flooding from Supercell Thunderstorms. *Journal of Hydrometeorology*, *2*(5), 469–489. https://doi.org/10.1175/1525-7541(2001)002<0469:ERAFFS>2.0.CO;2

Swathi, V., Srinivasa Raju, K., Varma, M. R. R., & Sai Veena, S. (2019). Automatic calibration of SWMM using NSGA-III and the effects of delineation scale on an urban catchment. *Journal of Hydroinformatics*, *21*(5), 781–797. https://doi.org/10.2166/hydro.2019.033

Talukdar, S., Singha, P., Mahato, S., Shahfahad, Pal, S., Liou, Y.-A., & Rahman, A. (2020). Land-Use Land-Cover Classification by Machine Learning Classifiers for Satellite Observations—A Review. *Remote Sensing*, *12*(7), 1135.

https://doi.org/10.3390/rs12071135

Thorndahl, S., Einfalt, T., Willems, P., Nielsen, J. E., ten Veldhuis, M.-C., Arnbjerg-Nielsen, K., Rasmussen, M. R., & Molnar, P. (2017). Weather radar rainfall data in urban hydrology. *Hydrology and Earth System Sciences*, *21*(3), 1359–1380. https://doi.org/10.5194/hess-21-1359-2017

Thorndahl, S., Schaarup-Jensen, K., & Rasmussen, M. R. (2015). On hydraulic and pollution effects of converting combined sewer catchments to separate sewer catchments. *Urban Water Journal*, *12*(2), 120–130. https://doi.org/10.1080/1573062X.2013.831915

Tibbetts, J. (2005). Combined Sewer Systems: Down, Dirty, and Out of Date. *Environmental Health Perspectives*, *113*(7). https://doi.org/10.1289/ehp.113-a464

UNISDR. (2015). *The human cost of weather related disasters*. https://www.unisdr.org/files/46796_cop21weatherdisastersreport2015.pdf

Viessman, W., & Lewis, G. . (2003). *Introduction to Hydrology* (Fifth Edit). Prentice-Hall, Upper Saddle River.

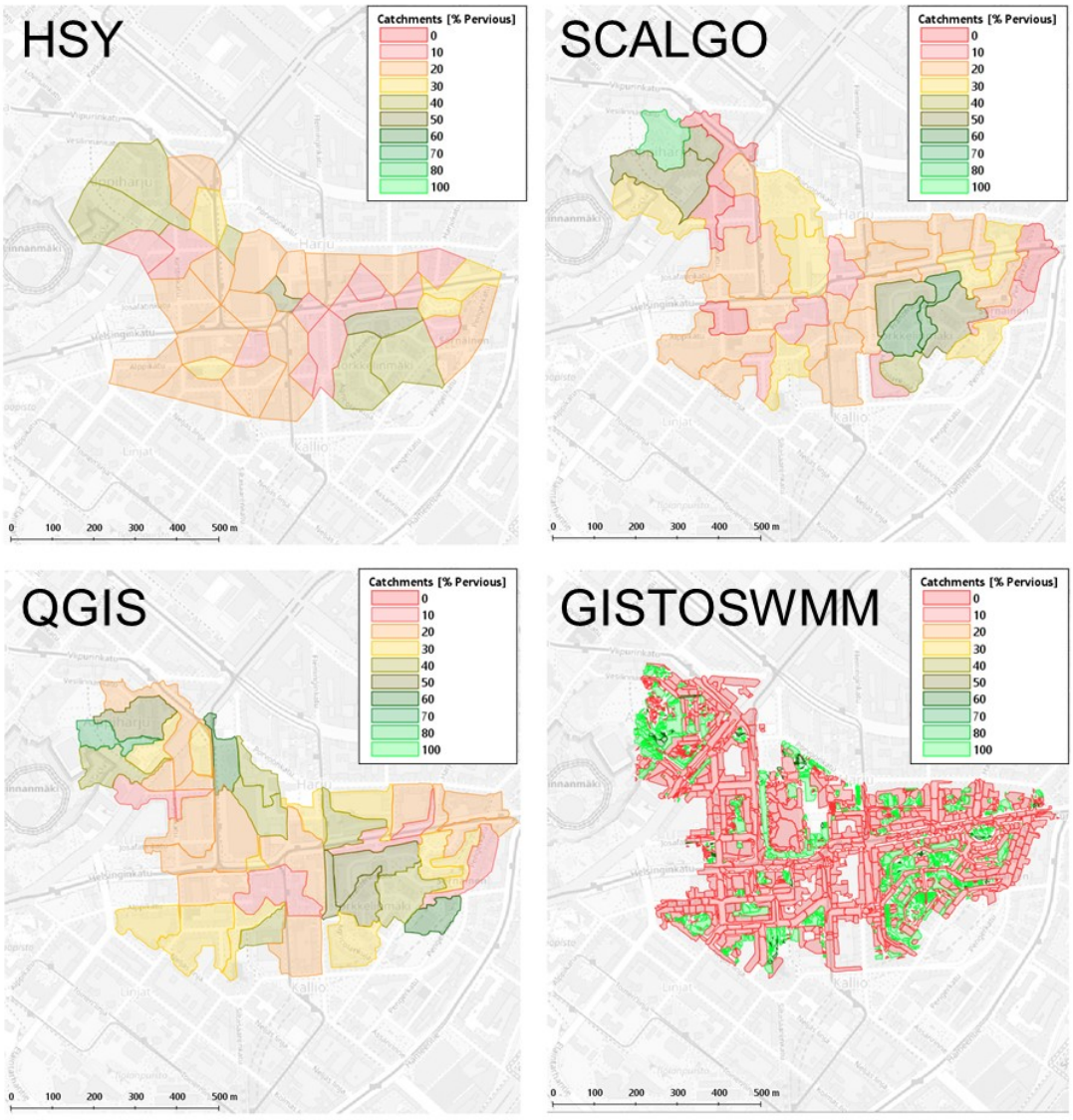
Warsta, L., Niemi, T. J., Taka, M., Krebs, G., Haahti, K., Koivusalo, H., & Kokkonen, T. (2017). Development and application of an automated subcatchment generator for SWMM using open data. *Urban Water Journal*, *14*(9), 954–963. https://doi.org/10.1080/1573062X.2017.1325496

Xu, Z., Xiong, L., Li, H., Xu, J., Cai, X., Chen, K., & Wu, J. (2019). Runoff simulation of two typical urban green land types with the Stormwater Management Model (SWMM): sensitivity analysis and calibration of runoff parameters. *Environmental Monitoring and Assessment*, *191*(6), 343. https://doi.org/10.1007/s10661-019-7445-9

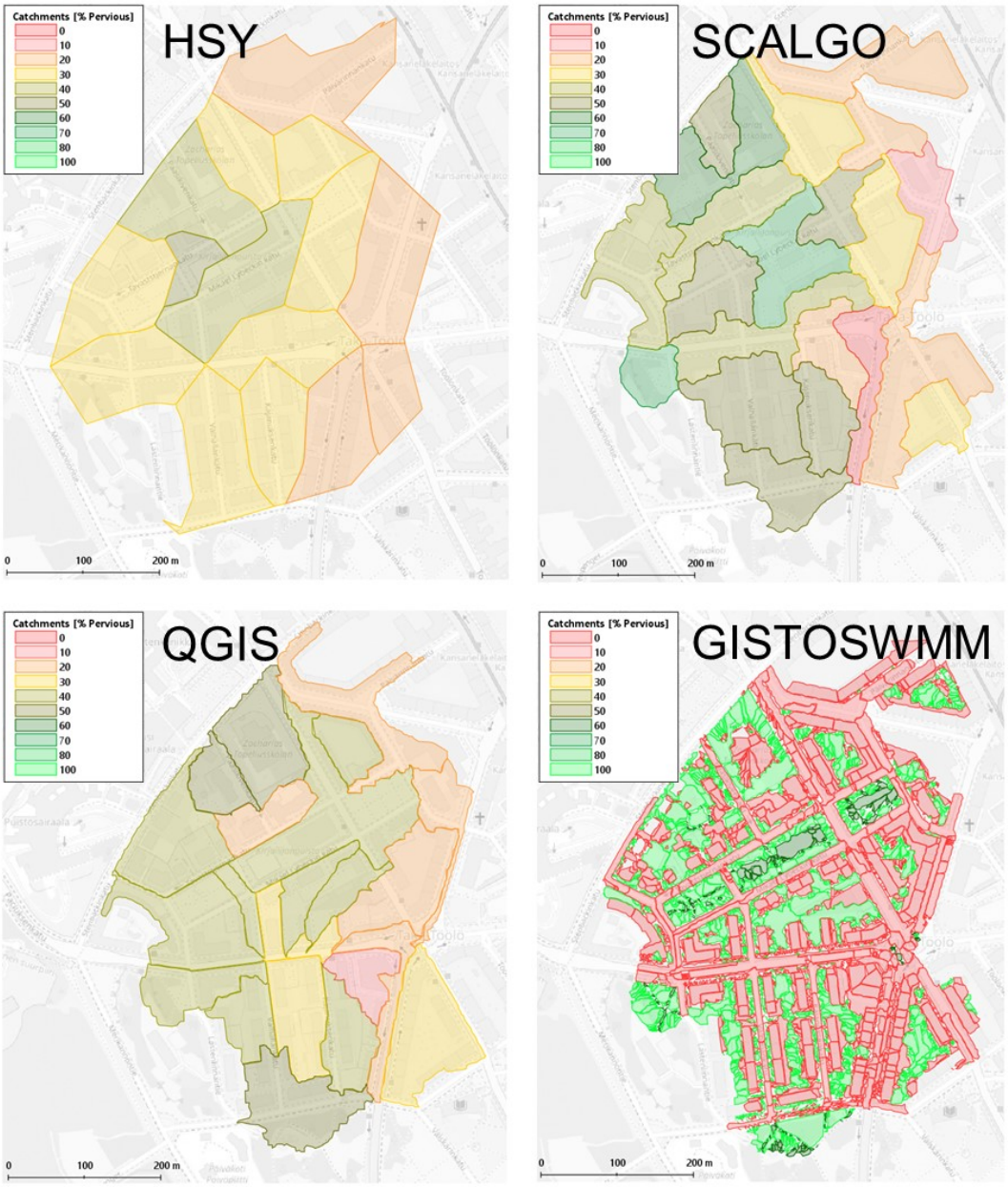
Zhang, H., Wan, L., Wang, T., Lin, Y., Lin, H., & Zheng, Z. (2019). Impervious Surface Estimation From Optical and Polarimetric SAR Data Using Small-Patched Deep Convolutional Networks: A Comparative Study. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *12*(7), 2374–2387. https://doi.org/10.1109/JSTARS.2019.2915277

Zhang, P., Ke, Y., Zhang, Z., Wang, M., Li, P., & Zhang, S. (2018). Urban Land Use and Land Cover Classification Using Novel Deep Learning Models Based on High Spatial Resolution Satellite Imagery. *Sensors*, *18*(11), 3717. https://doi.org/10.3390/s18113717

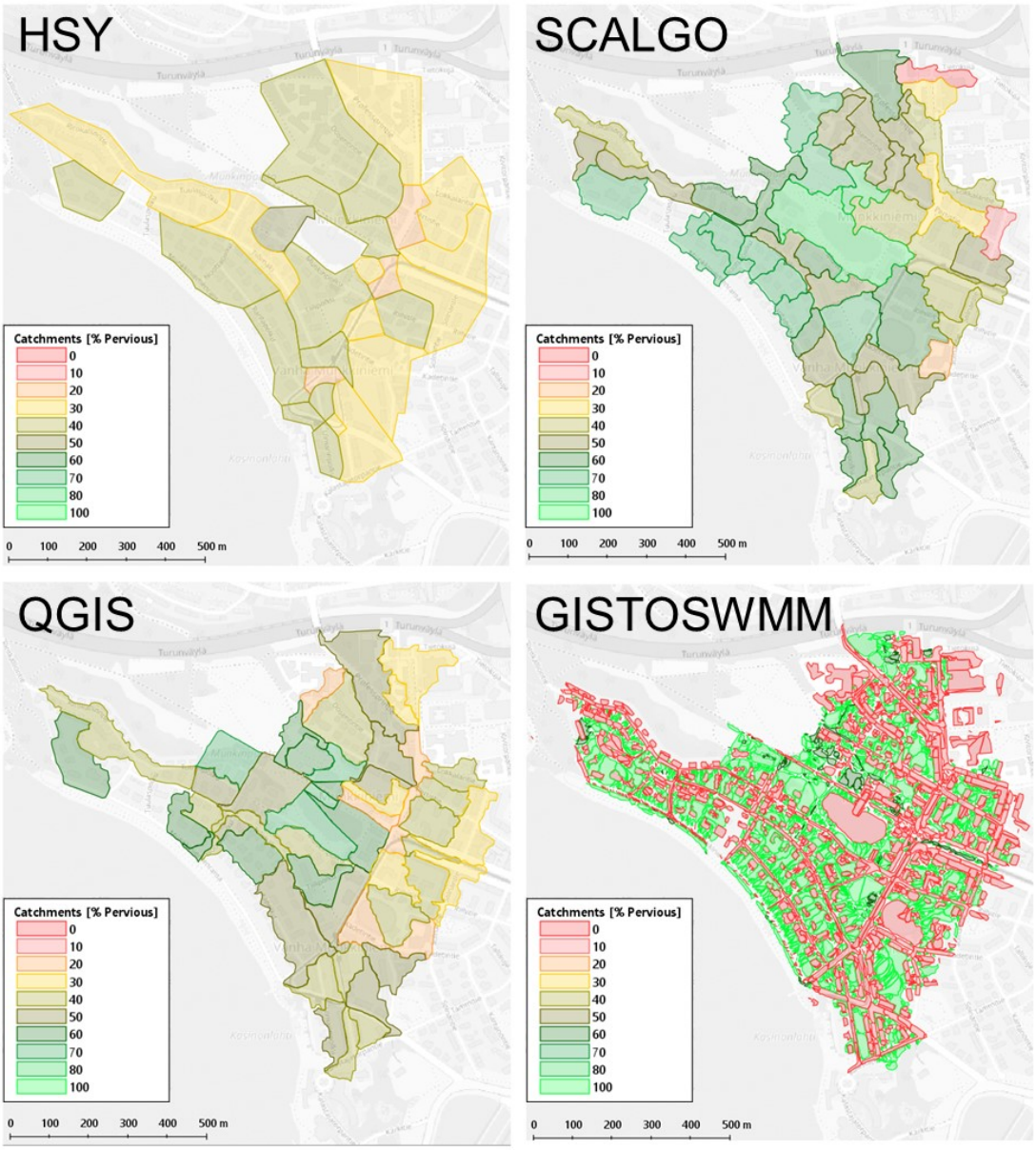
APPENDIX 1: BRAHENKENTTÄ SUB-CATCHMENTS



APPENDIX 2: TAIVALLAHTI SUB-CATCHMENTS



APPENDIX 3: MUNKKINIEMI SUB-CATCHMENTS



APPENDIX 4: HERTTONIEMI SUB-CATCHMENTS