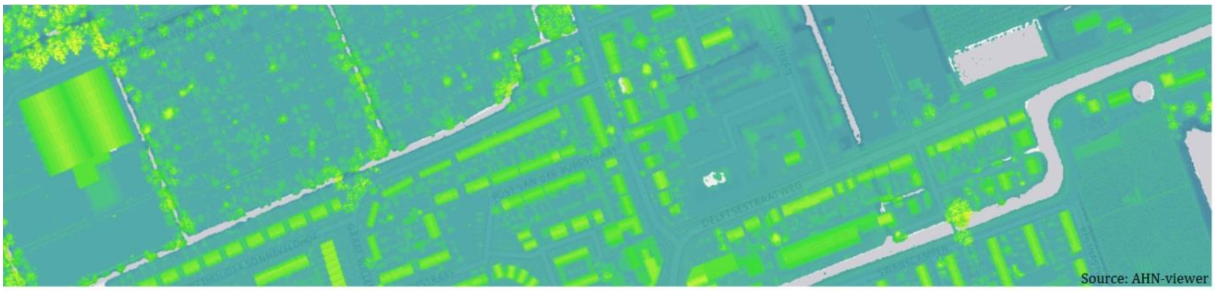
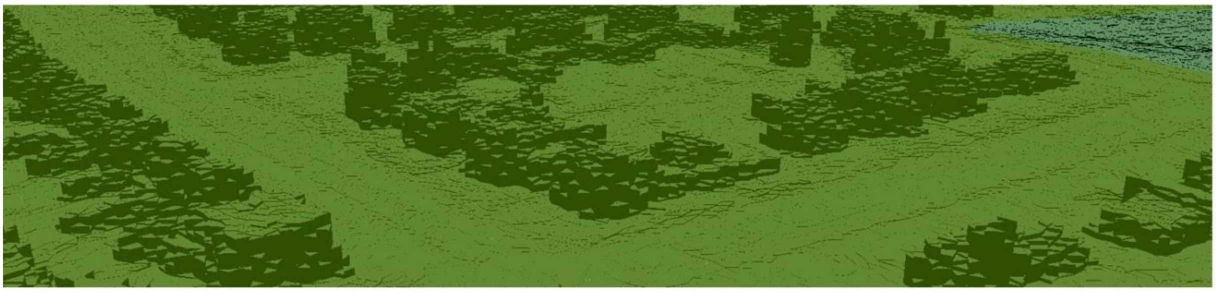
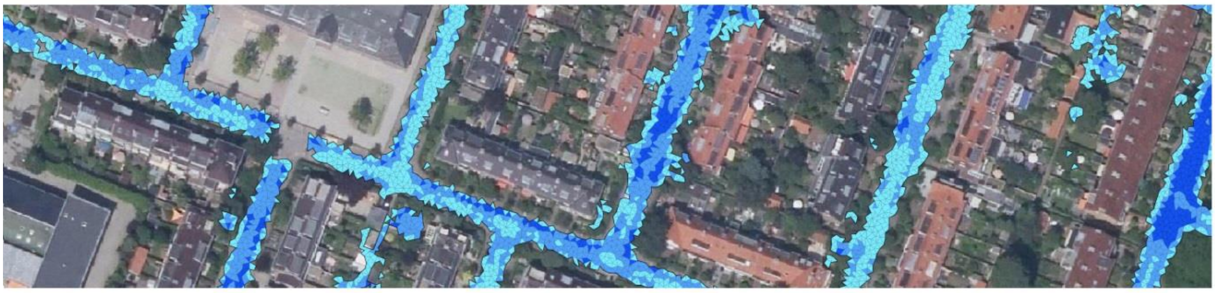
Assessing different modelling concepts for estimating pluvial flooding in urban areas



Source: AHN-viewer





Source: RTV Oost

MSc Thesis
Jacco Breedijk

Assessing different modelling concepts for estimating pluvial flooding in urban areas

By

Jacco Breedijk

As final fulfilment of the requirements for obtaining the degree of

Master of Science in Civil Engineering

at Delft University of Technology,

to be defended on Friday May 18, 2018 at 16:00

Graduation Committee:

Dr. ir. J.G. Langeveld	TU Delft,	Sanitary Engineering
Dr. ir. J.A.E ten Veldhuis	TU Delft,	Water Resources
ir. W.M.J. Luxemburg	TU Delft,	Water Resources
ir. G.J.R. Henckens	Royal HaskoningDHV	

Abstract

Our climate is changing. In the Netherlands, higher rainfall intensities and longer periods of drought are expected in the coming years. Urban environments have to be adapted in order to maintain (or improve) the living standards that we have set for today. Models of the urban drainage system play a vital role in the identification of flood-prone areas, participation of stakeholders and the design of effective measures. This research focuses on the modelling of the urban drainage system to assess pluvial flooding.

In the modelling of the urban drainage system, a large amount of choices has to be made by a modeller. There are multiple methods of representing the urban drainage system (modelling concepts) and various modelling software packages for the creation of such a model. Over the last years, the British Environment Agency and the Dutch STOWA have created benchmarks where a large number of modelling packages were compared by using predefined test cases (Environment Agency, 2013) (STOWA, 2017). It was concluded that differences between modelling packages are rather small. A similar comparison between modelling concepts is not available for all regularly used modelling concepts and for Dutch modelling practice. This research assesses different modelling concepts for estimating pluvial flooding in urban areas. The focus hereby lays on the Dutch practice of modelling. The main objective of this research is to compare the modelling concepts most often used in the Netherlands. In order to reach the objective, the following steps are required:

- Analyse the modelling concepts most often used in the Netherlands and determine aspects of modelling urban drainage that require further research.
- Create a set of models based on different modelling concepts.
- Select a transparent, objective and reproducible method for comparing the most frequently used modelling concepts.
- Apply the method on multiple test cases and compare and the results.

Among investigation of the modelling concepts most often used, four main modelling concepts were selected: 1D, 1D/2D, 1D/2D+ and 2D. Aside from a comparison of the four main modelling concepts, four additional aspects were determined to be possible causes of differences in modelling results:

- The inclusion of pervious surface area in models with semi-distributed runoff.
- The representation of roof surfaces with fully distributed runoff.
- The influence of paved yards on the amount of street flooding.
- The influence of infiltration zones on the results of a 2D model.

In total, eight different models were created, based on different modelling concepts. Four models are based on the main modelling concepts and four additional models are created to investigate each of the additional aspects. Three test cases were selected: Ulvenhout (Municipality of Breda), Tuindorp (Municipality of Utrecht) and Loenen (Municipality of Apeldoorn). The test cases were selected for their diverse characteristics in terms of slope and degree of imperviousness. For all test cases, a calibrated 1D model was available. All models in this research are based on the provided 1D model. Other models are created by adding or removing elements from the original 1D model without re-calibration of input parameters. The

performance of the models, based on different modelling concepts, is evaluated based on two tests: a comparison of the models with observed locations of flooding and no flooding (test 1) and a comparison with in-sewer monitoring results (test 2).

The addition of pervious surface area in models with semi-distributed runoff influences the computed amount of sewer inflow and street flooding. The 1D/2D_P model was created in which pervious area was added to the sub catchments. In test 1 and for test case Tuindorp, the model scores 96% in the recognition of observed locations of flooding using default test parameters. The 'standard' 1D/2D model (coded 1D/2D_St) scores only 8%. In the comparison with monitoring results in test 2, the difference in RMSE during peak water levels varied between 3 and 18 cm for the different test cases and storm events. However, the 1D/2D_P model overestimates the amount of street flooding and usually scores worse. The set of parameters often used in the Netherlands (based on Leidraad Riolering C2100), is concluded to provide unrealistically low parameters for infiltration and surface storage of pervious surface types. Proper validation of infiltration parameters is therefore advised.

Two models were compared for the assessment of the representation of roof surfaces with fully distributed runoff: a model where all runoff is modelled fully distributed (1D/2D+_Y) and a model where roof surfaces are modelled semi-distributed, creating a hybrid runoff model (1D/2D+_HY). The 1D/2D+_Y model underestimates the amount of sewer inflow. The bias ratio (β) between the model and measurements varied between 0.762 and 0.943 for the tested storm events. The 1D/2D+_HY performed better (0.904 - 1.007). Furthermore, differences of up to 12 cm in the RMSE for a storm event were found. In general, the 1D/2D+_Y model underestimates sewer inflow and results in more accumulation of water around houses, as roofs are not represented correctly. Hybrid runoff modelling cancels out part of the negative effect.

The runoff from paved yards also influences the amount of street flooding. Again, two models were compared. The 1D/2D+_H model assumed all yards are pervious. In the 1D/2D+_HY model, yards are further classified into impervious, semi-pervious and pervious parts. In test 1 and for test case Ulvenhout, the 1D/2D+_H model scored better in the recognition of observed non-flooding (71%) than observed flooding (39%) with default test parameters. The 1D/2D+_HY model scored 89% for both the recognition of flooding and non-flooding. This indicates that the 1D/2D+_H model underestimates the amount of street flooding or the chosen test parameters. Runoff from paved yards also influences sewer inflow: difference in RMSE between the models ranged from 0.1 to 2.4 cm. An important note is that proper validation of infiltration parameters has an even larger influence on the model performance, as the addition of paved yards did not always result in a better score for the tests. Furthermore, the addition of extra parameters and processes does not guarantee a better result. Over parametrisation could make calibration difficult.

As the last of the additional aspects, the influence of infiltration zones in 2D models was tested. The method of testing and the amount of test data were deemed insufficient to draw conclusions, as differences are most likely caused by the chosen parameters for discharge and infiltration.

Concerning general differences between modelling concepts, it was shown that a fully distributed runoff model provides benefits in the modelling of surface runoff and initial peaks in sewer water levels. However, fully distributed models tend to overestimate flooding around houses and underestimate sewer inflow. By modelling hybrid runoff, part of the negative effect is cancelled. Extending a 1D model to a 1D/2D model by adding an elevation model provides potential benefits for the modelling of storm events that result in significant street flooding. However, an improvement is not guaranteed and re-calibration of input parameters might be necessary.

The two tests proved effective in the assessment of modelling concepts and specific modelling aspects. A large benefit is that multiple types of information can be used. However, a large amount of accurate information, both in space and time, is needed. The amount of data that was available for this research supported conclusions on differences between modelling concepts. More data is needed to further quantify the differences that were found. Potentially, drones or satellites would be a reliable source of information. However, foliage, privacy regulations and the timing make widespread implementation difficult.

Acknowledgements

The document you are reading right now is my MSc Thesis. By completing this Thesis, an eighteen year journey through the educational system has almost come to an end (or climax, depending on the point of view).

From a very young age, I was fascinated by the weather. Most of my class mates in kindergarten wanted to become a police man, professional soccer player or veterinarian. I wanted to study and present the weather. For a surprisingly long time, that ambition remained. Geography was my favourite course during high school and meteorology was a big interest. After high school, I decided to study Applied Earth Sciences at the TU Delft. After a minor in Water Management, I decided to complete the transition and get a master degree in Water Management. Now, three years after making that choice, I am almost finished with that master degree. Although the end result does not entirely match the initial guess, 5-year-old me was still pretty close.

It is almost unimaginable what I learned and experienced over the past few years. Fieldwork in Luxemburg, a study trip to Ethiopia, a multidisciplinary project in Brazil and two internships are just some of the study related experiences. The experiences I had gave me a good view of the possibilities after my master degree. All the experiences, lectures and good moments outside of the classroom were not possible without some great people.

First, I would like to thank my graduation committee for their support. Wim, we have known each other for a long time and we have done about everything from giving presentations at in-house events to drinking whiskey. I appreciate the talks we had and counselling you gave. Marie-Claire, after an unfortunate change in my committee, you agreed to support me through the last two months. I want to thank you for your trust and valuable suggestions. Jeroen, after a lecture on sewer-WWTP interactions I approached you as I was looking for an internship. From that moment, you supervised me through two internships at Breda and Royal HaskoningDHV. Your interesting lectures definitely boosted my interest in urban drainage and I would like to thank you for your support during both good and bad times. Guy, I learned a lot from you in the nine months that we know each other. Thank you for the coffee we had, the tips and tricks of modelling urban drainage I learned and the personal advice you gave me. It is a pity that you are leaving Royal HaskoningDHV. It would be cool to do a project together some day.

Also outside of the committee, a lot of people supported me. I would like to thank Marco van Bijnen and the municipalities of Breda and Apeldoorn for their smooth cooperation. I would like to thank my colleagues at Royal HaskoningDHV for their help and the good talks we had. I am looking forward to work with you after a short break. I want to thank my house mates for the good times we had, from parties and playing board games to talking about goals and ambitions. Finally, my girlfriend, friends and family were of great support. A very pleasant observation is that throughout nine months of working long hours, often having a not-so-good mood, the bond with my family and girlfriend, Milou, only strengthened. Perhaps I consider that to be an even more pleasant thought than the irony of finalizing eighteen years of high level education with a specialization in the most unattractive structures in this world: sewer systems.

Jacco Breedijk
Delfgauw, May 2018

Contents

1. Introduction

Due to the effects of climate change, the global climate is expected to undergo some significant changes over the coming decades. In the Netherlands, higher rainfall intensities and longer periods of drought are expected (KNMI, 2014). To be able to cope with the more extreme weather conditions that are expected, urban environments and their surroundings have to be adapted accordingly.

1.1 Problem description

In the Netherlands, the governmental *Delta Program* urges governmental organizations like municipalities and water boards to increase their effort in climate adaptation. According to the *Delta Program*, climate adaptation concerns four main subjects: drought, heat stress, fluvial flooding and pluvial flooding. All governmental organizations are urged to highlight vulnerabilities in their climate resilience before 2020 by means of (rough) calculations. In 2050, climate resilient urban design should be the standard (Deltaprogramma, 2017). This poses a major task for the governmental organizations and asks for measures to create a more robust urban drainage system in the case of pluvial flooding.

Simulation models play a vital role in the process of creating a more robust urban drainage system. Simple models of urban drainage focus on topographic depressions and (usually) rainfall runoff for pointing out flood-prone areas. Only little data is needed and the models can be built up quickly. These models can help start up an initial dialogue by identifying vulnerable areas and for gaining support amongst stakeholders in order to take widely supported measures. More detailed models include a geometric representation of the sewer system, rainfall runoff, infiltration and in some cases surface water or groundwater. These models require more input data and proper calibration. Such models can be used for detailed spatial planning and to design measures against the consequences of heavy rainfall.

In the creation of such models, a large variety in both modelling concepts and modelling packages is available. A modelling concept is defined as a method of representing the urban drainage system. In the modelling of urban drainage, many aspects of the urban water system need to be included. Models are often built up by multiple partial components which interact with each other. For instance, the output of the rainfall runoff model forms the input for the hydrodynamic sewer model. A modeller has to make many decisions in the formation of a model. The combination of the choices made concerning the processes that are included in the model, their mathematic representation and relations between partial components are referred to in this Thesis as modelling concepts. A modelling package is the software in which the modelling concept is implemented. The variety in modelling concepts and packages provides engineers and policy makers with some difficult dilemmas: Which modelling concept should be applied, and what do I need to be aware of when applying a certain concept? Which modelling package supports the chosen modelling concept?

1.2 Comparing modelling packages

To help decision makers in selecting the right modelling package, benchmarks were composed with the most commonly used modelling packages. In such benchmarks, modelling packages are compared with each other by assessing performance indicators (formulas, run time, user interface) or by comparing the results for different test cases. In 2009, the British Environment Agency evaluated 2D inundation modelling packages against predefined performance indicators and recommended a set of simple benchmark test cases to be used in future research (Environment Agency, 2009). In that same year, a large number of modelling packages was benchmarked by the British Environment Agency. In 2013, this benchmark was renewed (Environment Agency, 2013). In the renewed benchmark, it was concluded that the benchmarked modelling packages that solve the full Shallow Water Equations are appropriate to support decision making involving flood risk management. In 2017, the Dutch STOWA made a similar benchmark that involved different modelling packages often used in the Netherlands (STOWA, 2017). The benchmark consisted of two parts: a questionnaire and a comparison by using seven test cases. It was concluded that the results of the benchmark give no immediate cause to prefer a certain modelling package. The only exception was modelling package HEC-RAS, which struggled to model flow of very shallow waters[1]. The test cases that were used in both the benchmark of the British Environment Agency and the STOWA are simple. For example: an inclined plane, a V-notch or flow through a single tube under a head difference. The test cases are not comparable to a complex urban environment. Still, as the benchmarks concluded that differences between modelling packages are rather small, it is assumed that conclusions based on one modelling package are generalizable. This research focuses on modelling concepts.

1.3 Comparing modelling concepts

The choice for the most suitable modelling concept is influenced by many factors, such as the areal characteristics of the selected case, the availability of data, the computational capacity and the ambition of the modeller (or budget in a commercial environment). The choice for a modelling concept is therefore far from straight-forward.

An overview of the most frequently used modelling concepts is given in (Henonin, et al., 2013). The article describes the differences between various classes of modelling concepts and what they are capable of in terms of flood prediction. Aside from the overview, comparisons between models that use different modelling concepts are found in literature (Leandro, et al., 2009) (Vojinovic & Tutulic, 2009) (Freni, et al., 2010) (Stichting RIONED, 2014) (Pina, et al., 2016). The implications of these comparisons, together with an overview of the various modelling concepts, are discussed in this thesis. The available literature is the starting point for creating a method for a full comparison of modelling concepts, with a focus on modelling practice in the Netherlands.

1.4 Research scope and aims

The main objective of this Thesis is to compare the modelling concepts most often used in the Netherlands. In order to do this, the following steps need to be taken:

- Analyse the modelling concepts most often used in the Netherlands and determine aspects of modelling urban drainage that require further research.
- Create a set of models based on different modelling concepts.
- Select a transparent, objective and reproducible method for comparing the most frequently used modelling concepts.
- Apply the method on multiple test cases and compare the results.

The focus of the comparison lays on the abilities of a model, based on a certain modelling concept, to calculate (in-sewer) water levels and the extent of street flooding for a single storm

[1] According to a recent email conversation, measures have been taken to solve the problem (Sanchez, 2018)

event. The data needed to assess these abilities will be discussed. The modelling of surface water or groundwater is not taken into account. It is assumed that rain falls equally distributed over the catchments in space. Only flooding of the urban drainage system due to heavy rainfall is assessed. Differences in computation time are mentioned, but not compared as part of the assessment.

1.5 Report structure

Chapter 2 presents the components of a modern day sewer system. The different types of sewer systems are shortly explained. Urban flooding and the factors that influence it are discussed. The scope and context of the research are further explained. In chapter 3, an overview is given of the main classes of modelling concepts that are often used in modelling urban drainage. In chapter 4, available literature is presented on comparisons between modelling concepts. Aspects that need further research are determined. Finally, eight models, based on different modelling concepts, are presented. In chapter 5, the three test cases that are used in the research are presented. Appropriate storms for the comparison are selected. Chapter 6 presents the method for assessing the modelling concepts. This method consists of two separate tests. Chapter 7 presents the results of the two tests. Chapter 8 forms the discussion. Both the differences between modelling results and the quality of the tests are discussed. In chapter 9, the conclusions are summarized and recommendations are given for further research.

The subject of urban drainage often implies the use of technical jargon and abbreviations from multiple disciplines (such as remote sensing, statistics, hydrology or hydraulics). As an aid for the various terms, a glossary is added at the end of the report with short explanations and Dutch translations. Figure 1 shows the report structure.

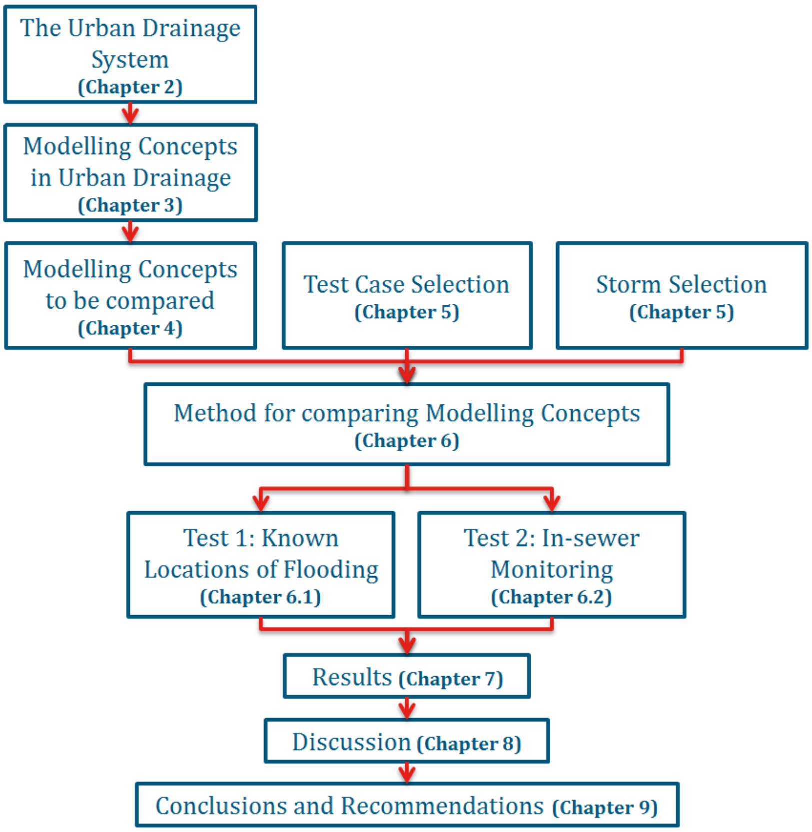


Figure 1 - Schematic overview of the report structure

2. The Urban Drainage System

Wherever humans interfere with the natural landscape, the water system is affected. Drinking water is needed, waste water needs to be discharged and rainwater infiltration is disrupted. In ancient civilizations, human waste water was often used to fertilize the surrounding land. With ever increasing sizes of settlements, the amount of waste water that needed to be discharged surpassed the need of the surrounding land. Because of the excess of waste water, the first sewers were constructed. A brief history on sewers is given in Appendix H. This chapter explains the components of a modern day sewer system (§ 2.1). The different types of sewer systems are shortly explained. Urban flooding and the factors that influence it are discussed (§ 2.2). The scope and context of the research are further explained (§ 2.3).

2.1 Sewer systems

In most sewer systems, waste water and storm water are collected by the same piping system. Such a system is called a combined sewer system. The water flows towards the waste water treatment plant (WWTP) under gravity. If the landscape has a small surface slope, the water flows towards a pump sump from where it is pumped towards the WWTP. This is often the case in the Netherlands. If waste water and storm water are collected and discharged separately, the system is called a separated sewer system. Figure 2 illustrates the difference in piping system in and around a house.

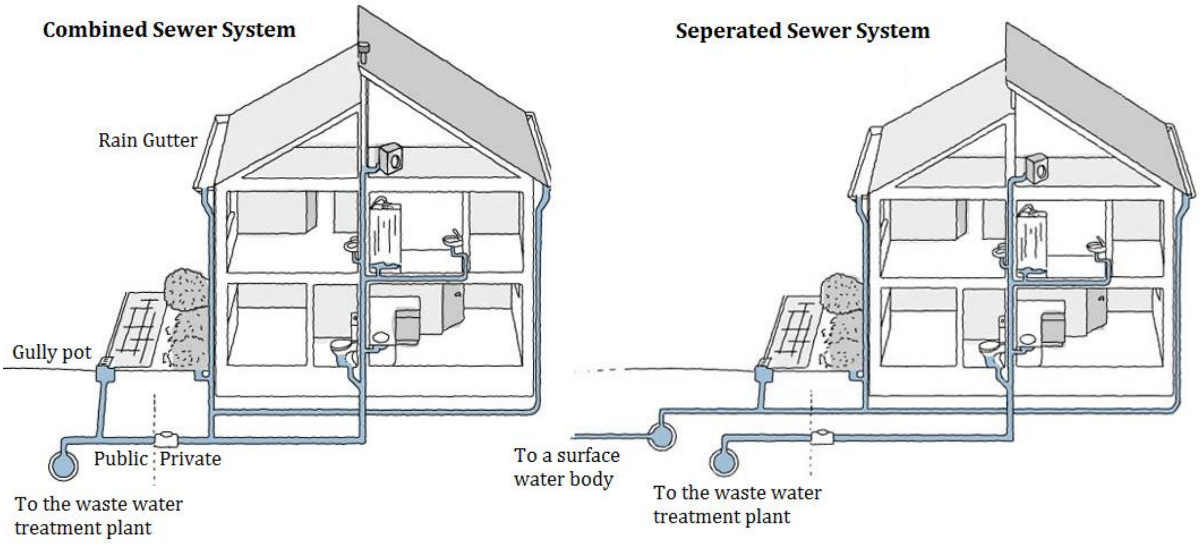


Figure 2 – A combined and a separate sewer system (based on (Gemeente Enschede, 2018))

2.1.1 Combined sewer systems

During dry weather conditions, only waste water is transported through the piping system. Under wet weather conditions, rain water enters the sewer system through roof drains (rain gutters), street drains (gully pots) or underground drains. For small storm events, all of the rain water and waste water are transported towards the WWTP. In order to prevent street flooding for larger storm events, (combined) sewer overflows are built. The structures form an escape route out of the sewer system and into the receiving water body in case the sewer almost overflows onto the streets. The structures are likely to overflow a few times per year. In such a case, the discharged waste water is significantly diluted with rain water. Still, combined sewer overflows often have a negative impact on the water quality of the receiving water body. For storms with a large return period, the sewer overflows onto the streets. In the Netherlands, most sewer systems are designed to prevent surface flooding for a one-hour storm event with a return period of two years (Dutch Design Storm 8 or Bui 08). Figure 3 shows the dynamics of a combined sewer system under dry and under wet conditions.

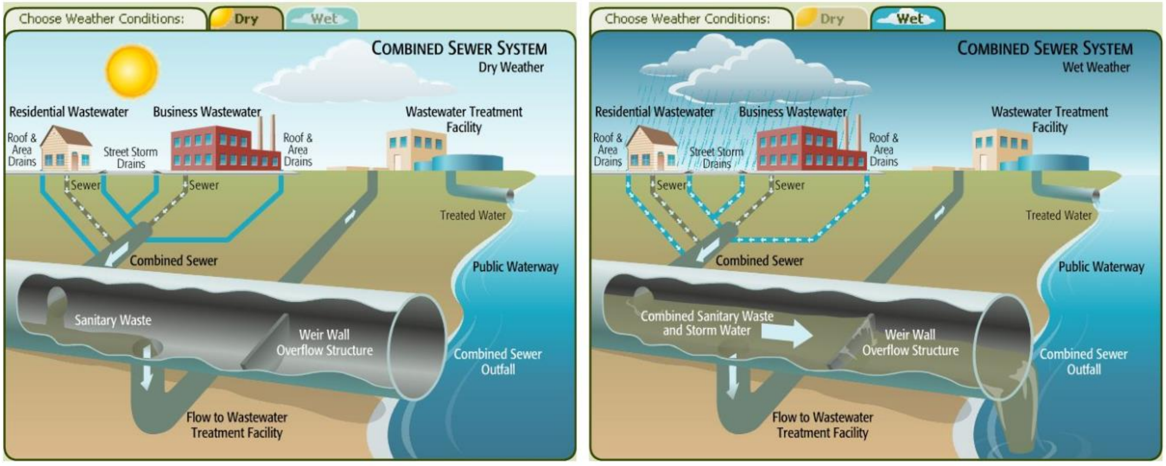


Figure 3 – Dynamics of a combined sewer system under dry and under wet conditions (Henderson Water Utility, 2018)

2.1.2 Separated sewer systems

Increased attention for the quality of the receiving water body helped to boost the creation of separated sewer systems. In a separated sewer system, waste water and storm water are collected separately. The waste water is transported to the WWTP; storm water is discharged directly to nearby surface water without further treatment.

By separating the collection of waste water and storm water, WWTP's face a more constant inflow and therefore reach a higher removal efficiency of pollutants. Furthermore, the separation of waste water and storm water prevents the pollution load from combined sewer overflows, as the storm water is directly discharged to the receiving water body without pollution from waste water. Figure 4 shows the dynamics a seperated sewer system under dry and under wet conditions.

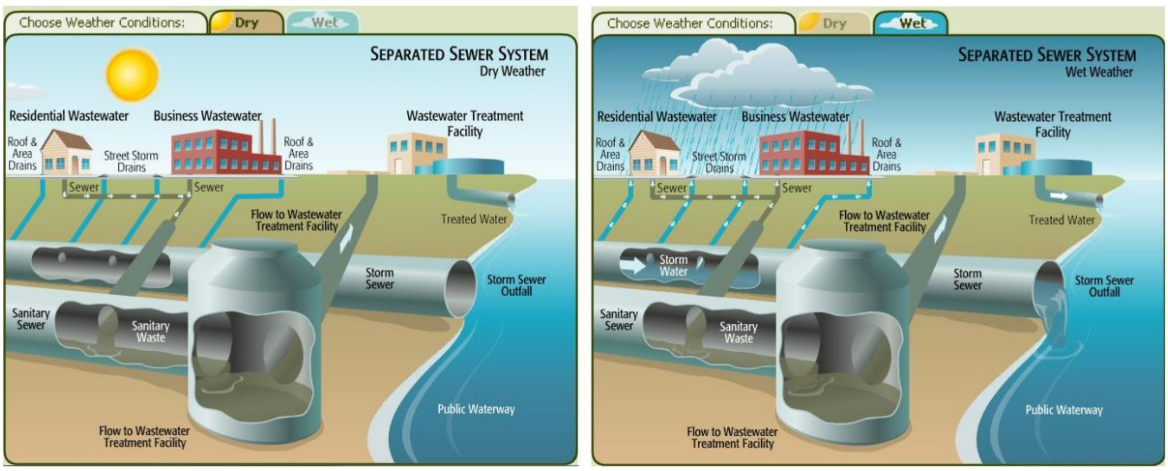


Figure 4 – Dynamics of a separated sewer system under dry and under wet conditions (Henderson Water Utility, 2018)

In practice, replacing a combined sewer system with a seperated sewer system does not always result in an improvement of the water quality. Even the storm sewer in a separated sewer system contributes to a significant pollution load. Reasons for this are incidental discharges of waste water into gully pots, polluted sediment on street surfaces and faulty sewer connections.

The high pollution load of a seperated sewer system led to the creation of improved seperated sewer systems, where the storm water sewer is connected to the waste water sewer by a small pipe to prevent pollution of surface water during dry weather conditions. The different types of sewer systems and the improvement of water quality are not within the scope of this research. This research focuses on the nuisance caused by urban flooding and the modelling of the processes that influence urban flooding.

2.2 Urban flooding

The sewer system is designed to discharge waste water and storm water away from the buildings and streets. However, sewers have a limited discharge capacity and storage. In the Netherlands, the design standard for a combined sewer system is to have a storage capacity that is enough to store 7 mm of water and pump capacity of 0.7 mm/h. Furthermore, a storage tank with 2 mm of extra storage should be present. A separate sewer system should have a storage capacity of 4 mm and a pump capacity of 0.3 mm/h (Vereniging VPB, 2008). The mentioned pump capacity should be the overcapacity, on top of the capacity for dry weather flow.

When the storage capacity is filled and the sewer inflow surpasses the discharge capacity of the sewer system (including pumps and overflow weirs), water will accumulate on the surface. This situation is undesired, as large scale street flooding causes both direct and indirect damage. Direct flood damage is directly caused by the water. Furniture or floors may get damaged or children may get sick and need health care. Indirect damage is not directly caused by the water. Traffic and business interruption are examples of this. Damage estimation is not within the scope of this research. For literature on damage estimation, see (Merz, et al., 2004) (ten Veldhuis & Clemens, 2010) (Spekkers, 2015).

Urban flooding is often caused by too high rain intensity for the sewer system to cope with. However, other processes could also lead to urban flooding. Sewer related examples are the blockage of gullies with sediment, blockage of house connections, root intrusion, fatty deposits in pump sumps or pipe failure (for examples, see (Post, 2016) (Nieuwenhuis, et al., 2018)). Besides, flooding due to a high surface water table is possible. Over the coming years, the frequency of urban pluvial flooding is expected to increase in the Netherlands. There are two main reasons for this:

- In modern cities, most of the surface area is impervious. Impervious surface area is surface with a small infiltration capacity, such as concrete, asphalt or roofs. A higher degree of imperviousness implies that a larger fraction of the incoming rainfall cannot infiltrate into the subsurface. The water is instead turned into surface runoff, which enters the sewer system. A higher degree of imperviousness therefore implies a larger amount of storm water inflow into the sewer system. Figure 5 shows the effect of an increasing degree of imperviousness on the amount of surface runoff.
- Due to the effects of climate change, higher rainfall intensities are expected over the coming years. In 2050, the hourly rainfall intensity with a return period of one year is expected to increase between 5.5% and 25%, depending on the scenario. In 2085, an

increase between 8% and 45% is expected compared to the reference period (1981-2010) (KNMI, 2014).

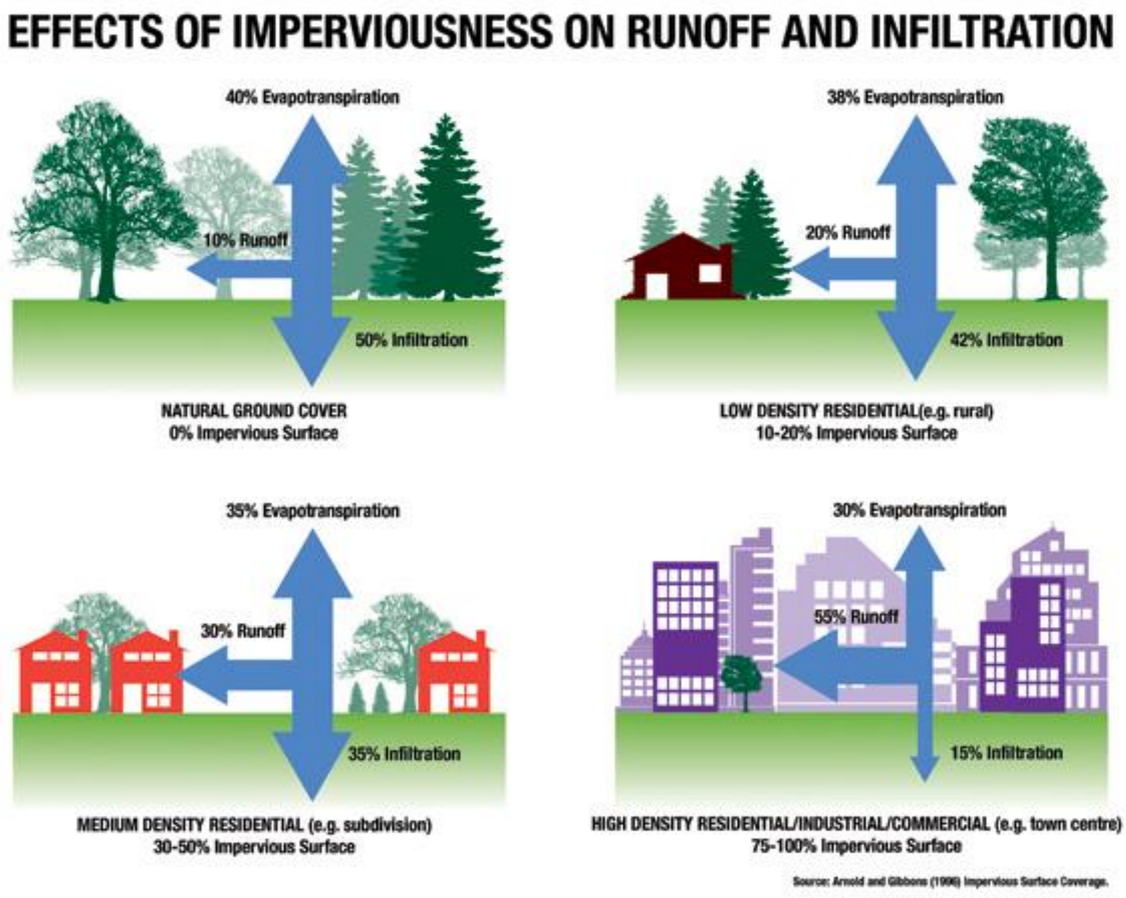


Figure 5 – The effect of imperviousness on the amount of surface runoff. More surface runoff usually implies more sewer inflow.

2.3 Urban flooding in this research

In this research, the modelling of the urban drainage system under heavy rainfall conditions is assessed. It is assumed that the sewers are cleaned and maintained properly. Surface water flooding is not taken into account. Many factors influence the amount of urban pluvial flooding, like the storage of the sewer system, discharge capacity, pipe diameters and the possibility for infiltration. A model of the urban drainage system should therefore include these components.

For a lot of cities around the world, it will be a challenging assignment to create a robust urban drainage system. Measures are needed for this. The design of measures starts with a model of the urban drainage system. The various modelling concepts are explained in Chapter 3.

3. Modelling Concepts in Urban Drainage

The modern urban water system deals with five types of water: rain water, drinking water, surface water, groundwater and waste water (Van de Ven, 2016). These types of water interact within the urban water system and form a complex balance. When urban drainage is modelled in order to assess the quantity of pluvial flooding, drinking water does not play a role. Waste water takes part of the sewer's storage, although volumes are rather small. Groundwater and surface water play a more significant role for longer storm events. Rainfall, rainfall runoff and the hydraulic performance of the sewer system are the key elements in modelling pluvial flooding.

An overview of the main classes of modelling concepts is presented (§ 3.1 - 3.5), mostly based on classifications used in (Henonin, et al., 2013) and a case study in Nijmegen (Stichting RIONED, 2014). Only hydrodynamic models are assessed, which means that the models have a time-varying input and output. The concept of a 1D model forms the basis of a 1D/1D, 1D/2D (and to a lesser extent 1D/2D+) model. An overview of the various classes of modelling concepts is given in § 3.6.

3.1 1D models

Most traditional urban storm water models are based on a 1D hydrodynamic model of the sewer system. A 1D model usually consists of two (partly) separate process descriptions: a hydrological model and a hydraulic model (Clemens, 2001). The hydrological model contains the transformation from rain into sewer inflow and the accompanying processes that influence that transformation. The hydraulic model contains a geometrical representation of the sewer system and the equations for fluid flow that describe the transport of waste water and storm water through the sewer system. The two partial models will now be discussed in more detail.

3.1.1 The hydrological model in a 1D model

When rain falls onto a surface, it does not necessarily flow towards the gullies and enters the sewer system. Initial losses due to wetting of the surface will occur if the surface is dry, depending on surface type, temperature and humidity (Clemens, 2001). Besides initial losses, every type of surface will have some retaining capacity due to small topographic depressions. This is called surface storage. The surface storage of pavement is low, while surface storage of grassland is high. Usually, the initial storage and surface storage are included in the model by using surface-depended constants.

Infiltration and evaporation are also key factors in the transformation process. Only when the rainfall intensity surpasses the rate of infiltration and evaporation, topographic depressions start filling up. For the infiltration, several models are available. A regularly used model is the Horton infiltration model (Equation 1) (Horton, 1940). At the start of a rain event, the unsaturated zone has maximum storage capacity left. At this moment, the infiltration capacity is at its maximum. The unsaturated zone is defined as the portion of the surface above the groundwater table (U.S. Geological Survey, 2018). The initial infiltration parameter is used for this situation (f_0). With a depleting storage in the unsaturated zone, the infiltration capacity decreases over time. When this storage capacity is depleted, f_c is the remaining capacity. The remaining capacity is equal to the rate of percolation. The rate of decay (k) influences the point in time where the soil storage capacity is depleted. Evaporation is represented by a (time-varying) constant, subtracted from the rainfall or assumed negligible.

$$f_t = f_c + (f_0 - f_c)e^{-kt} \qquad \text{(Equation 1)}$$

f_t = infiltration capacity at time t *(mm/hr)*
f_c = final infiltration capacity when the soil is saturated *(mm/hr)*
f_0 = initial infiltration capacity *(mm/hr)*
k = rate of decay *(1/hr)*
t = time of contact between water and soil *(hr)*

When (after initial losses) the surface storage is filled and the rainfall intensity surpasses the infiltration and evaporation, surface runoff occurs. In a 1D model, runoff is modelled semi-distributed (Pina, et al., 2016). This implies that a certain amount of surface area is assigned to each sewer manhole, a so-called sub catchment. If the rainfall is distributed evenly over the surface and 2D surface runoff towards gullies is modelled, the runoff is modelled fully-distributed. Figure 6 shows an example of a semi-distributed model (left) where the determination of sub catchments is based on Voronoi polygons. The surface area of every polygon is now divided into land use categories (for example, 40% pavement, 30% roofs, 30% yards, 0% parks). Each land use category has its own parameters for infiltration, surface storage and runoff delay. Runoff is now calculated for every manhole and for every predefined discrete time step. Waste water is added as a population equivalent per sub catchment and varying over time. The total calculated volume is the inflow for that specific manhole and forms the input for the hydraulic model.

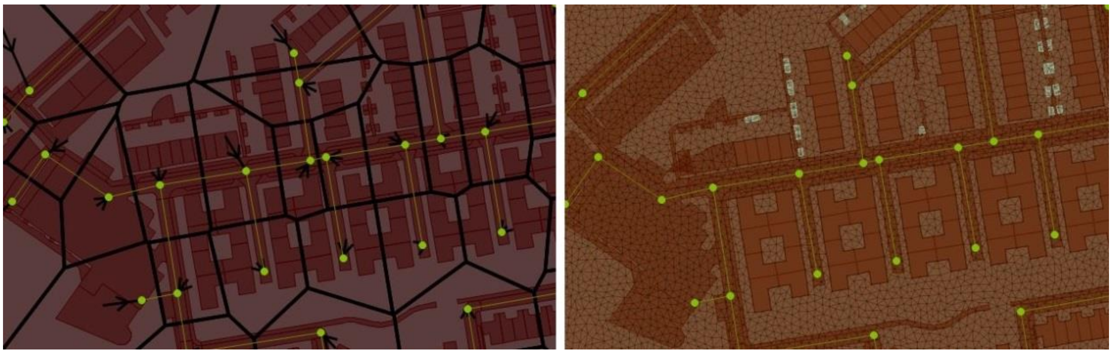


Figure 6 – A model with semi-distributed runoff (left) and fully distributed runoff (right). The sewer network is depicted in yellow. A model with semi-distributed runoff makes use of sub catchments to model sewer inflow.

3.1.2 The hydraulic model in a 1D model

Sewer inflow, as determined by the hydrological model, enters the sewer system directly at the locations of manholes. Gullies are not modelled in a 1D model. The fluid flow through the sewer system is modelled using the 1D De Saint-Venant equations (Saint-Venant, 1871). Most important assumptions in using the 1D-approach are hydrostatic pressure and a dominant flow velocity component in one direction (Clemens, 2001).

If the hydraulic head of the water in the sewer system is higher than the street level at a certain manhole, the manhole will overflow. Manholes are modelled open and act as a virtual link between the sewer system and the surface. A 1D model does not contain a surface model. The flood water is stored in virtual structures on top of the manholes, with no surface interaction between them (Figure 7). The water volumes are therefore correct, but the water depths have little meaning without post-treatment with a surface model. A 1D model has a very fast calculation time (usually seconds for a 1-hour storm event), but requires a well-maintained sewer database.

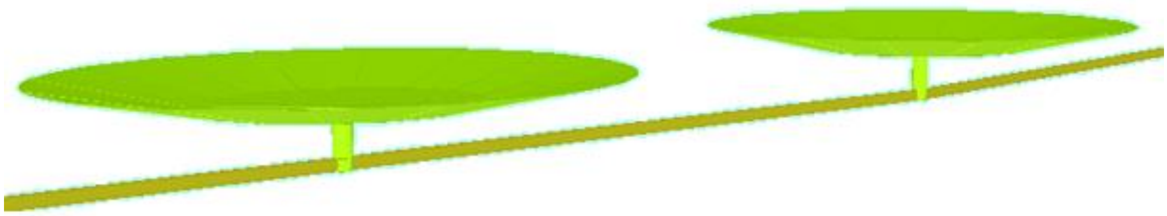


Figure 7- Virtual cones as used in a 1D model, retrieved from the 3D network viewer in InfoWorks ICM

3.2 1D/1D models

As the virtual flood structures do not represent the behaviour of water overflowing from a manhole onto the street surface, a 1D/1D model combines the 1D sewer model with a 1D surface network. The method is also called a dual drainage approach (Djordjevic, et al., 1999). The surface network is introduced for modelling the interaction between different manholes and to give an impression of the actual flooding that occurs. Manholes are the locations in the model where the two 1D models are linked with each other (see Figure 8).

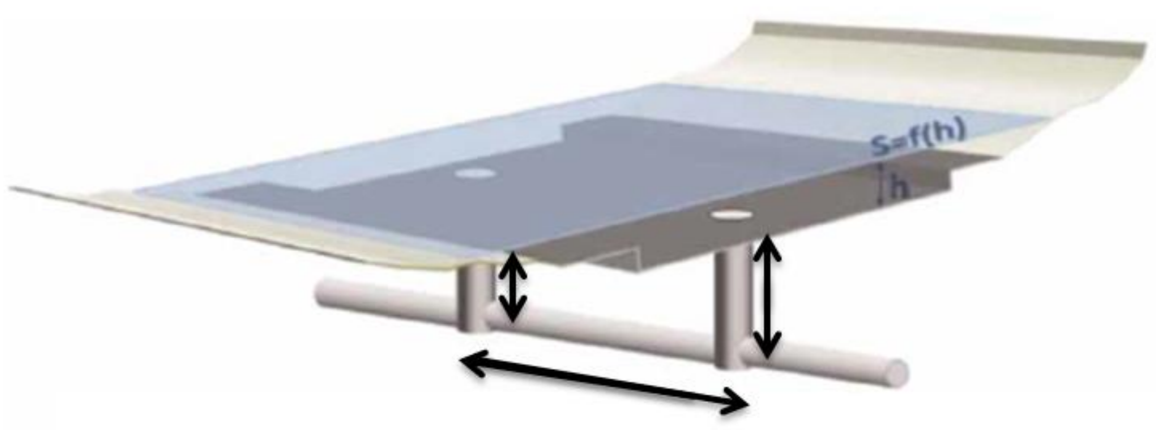


Figure 8 - 1D/1D model, where the drainage network and the street network are both represented in 1D (Henonin, et al., 2013)

3.3 1D/2D models

In a 1D/2D model, a 1D sewer model is coupled with a 2D surface model, which uses a digital elevation model (DEM). This allows for the modelling of overland flow. The hydrological model and hydraulic model mostly use the same basic principles as a 1D model. When a manhole overflows onto the surface, the flood water forms the input of the 2D surface model (see Figure 9). The 2D surface model makes use of a grid, where overflowed water from a manhole is converted into inflow at corresponding grid cells. The 2D Shallow Water Equations are used to calculate the surface flow. Extending a 1D model to a 1D/2D model provides a more realistic view of the surface processes when an intense rain event is modelled. When no manholes discharge in a certain storm event, the model performs equal to a 1D model. Because of the added elevation model and 2D surface flow, computation time of a 1D/2D model is significantly larger compared to a 1D model: running a model of a one hour storm takes in the order of minutes.

In (Adeogun, et al., 2015), the influence of modelling parameters on the performance of a 1D/2D model is discussed. Grid size, as a combination of the accuracy of the DEM and the grid size selected in the mode, was determined to be a key parameter in a 1D/2D model. Grid size greatly influences both computational time and the extent of the flooding. Besides grid size, surface roughness (represented by the Manning's number) was also analysed and pointed out as a significant influence. It was advised to carefully select computational grid size in order to balance computation time and model accuracy.

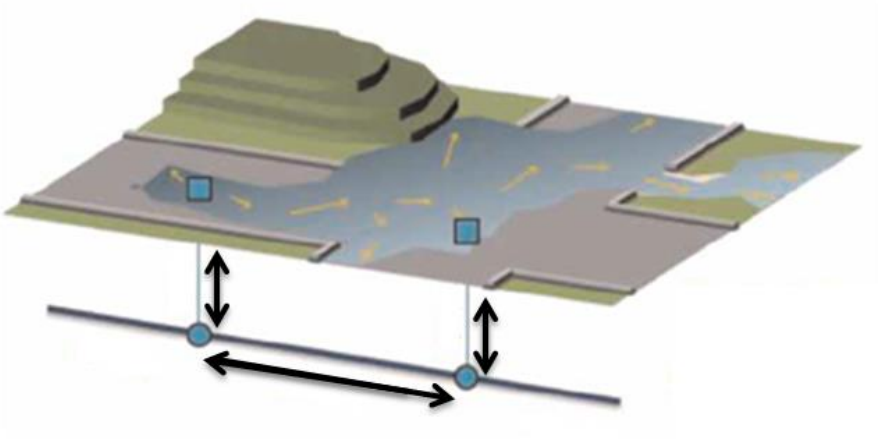


Figure 9 – 1D/2D model, where a 1D drainage network is coupled with a 2D surface network (Henonin, et al., 2013)

3.4 1D/2D+ models

1D/2D+ models do not represent the inflow of rainfall into the sewer system by making use of sub catchments (semi-distributed runoff). Rainfall is distributed equally over each grid cell of the 2D surface model and flows towards both gullies and manholes (fully distributed runoff). Figure 10 shows the difference in rainfall runoff between a 1D/2D model (left, semi-distributed) and a 1D/2D+ model (right, fully distributed).

In order to accurately replicate the overland flow towards the sewer system, gullies are modelled as entry points into the sewer system. Often, a fixed Q-h relation is used to model the relation between water depth and sewer inflow (Stichting RIONED, 2014). The land surface is now divided into infiltration zones with different surface types (green, paved etc.) and infiltration parameters. In contrast to the semi-distributed runoff, where sub catchments are made up of multiple surface types, every grid cell is assigned one surface type. A 1D/2D+ model has the largest amount of parameters, making calibration difficult and time consuming. Furthermore, using a fully distributed rainfall runoff model increases the computation time: a calculation time of one up to multiple hours is not uncommon for a single storm event.

3.5 2D models

2D models do not contain a hydraulic model of the sewer system. As sewers have a limited amount of storage and discharge capacity, the impact of the performance of the sewer system on the amount of street flooding decreases for more intense rainfall. This implies that the influence of local differences in sewer capacity or storage on the amount of street flooding decreases. The sewer is therefore not represented, but accounted for by simple assumptions, such as a standard discharge capacity that is subtracted from the incoming rainfall. 2D models contain a surface model with a DEM. Runoff is modelled fully distributed and sometimes infiltration is added to the grid cells. 2D models are often used when data is missing or when there is no (information on the) actual sewer network (Henonin, et al., 2013). Because of the simplicity and minor data requirements, the models are quickly built. Validation of the modelling results by comparing them with flood prone locations is seldom done. The models are often used as an initial estimate of vulnerable areas in a catchment, village or city.

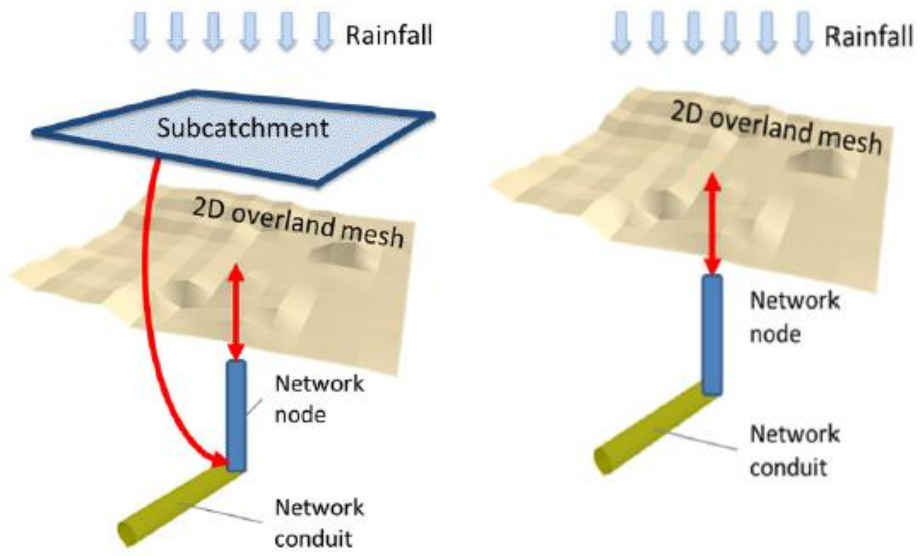


Figure 10 – Differnce between a 1D/2D model (left, semi-distributed rainfall runoff) and a 1D/2D+ model (right, fully distributed rainfall runoff) (Pina, et al., 2016)

3.6 Overview of modelling concepts

Table 1 gives an overview of the classes of modelling concepts that were discussed in this chapter. The two classifications for runoff (semi-distributed and fully distributed) are based on the classification in (Pina, et al., 2016) and are explained in Figure 10.

	1D	1D/1D	1D/2D	1D/2D+	2D
Sewer system representation	*Yes*	*Yes*	*Yes*	*Yes*	*No*
Elevation Model (DEM)	*No*	*No*	*Yes*	*Yes*	*Yes*
Runoff modelling	*Semi-distributed*	*Semi-distributed*	*Semi-distributed*	*Fully distributed*	*Fully distributed*
Infiltration model	*Horton, via sub catchments*	*Horton, via sub catchments*	*Horton, via sub catchments*	*Horton, via infiltration zones*	*Horton, via infiltration zones*
Calculation time[2]	*Very fast (sec-min)*	*Fast (min)*	*Moderate (min-1 hour)*	*Very slow (1 hour-hrs)*	*Slow (min-hrs)*
Flood map accuracy[3]	*None to low*	*Moderate*	*High*	*High*	*Moderate to high*

Table 1 - Overview of the main classes of modelling concepts as presented in this chapter

[2] In this research, only hydrodynamic models are assessed. Hydrostatic models have a shorter calculation time.
[3] The flood map accuracy is based on the overview shown in (Henonin, et al., 2013)

4. Comparing Modelling Concepts

This chapter first gives an overview of available literature on comparing modelling concepts (§ 4.1). Characteristics of modelling urban drainage in the Netherlands are given, together with possible differences with the test cases found in literature (§ 4.2). Additional aspects that require further research are pointed out based on literature and personal communication with modellers (§ 4.3). Different schematic setups are proposed in order to assess the additional aspects. Finally, a set of eight models, based on different modelling concepts, is presented.

4.1 Comparisons between modelling concepts in literature

In (Freni, et al., 2010), a 1D hydrodynamic sewer model was coupled with two different approaches modelling overland flow: a storage-weir approach and a dual drainage approach. The storage weir approach follows almost the same principles as a 1D model (§3.1). The only difference is that interaction between the virtual flood structures is made possible by using weir equations. In the dual drainage approach, a 1D street channel was added allowing a more realistic interaction between flooded manholes (1D/1D model). It was concluded that the inclusion of modelling overland flow between manholes results in less errors and outliers from the measurements, thus recommending the use of a 1D/1D over a 1D model.

In (Leandro, et al., 2009), the key factors in setting up a 1D/1D model are discussed, together with a detailed comparison with a 1D/2D model. The paper concluded that with proper calibration, a 1D/1D model shows consistencies with a 1D/2D model (grid size DEM: 2.0 m). However, applying a 1D/1D modelling concept causes difficulties, especially in the modelling of the spreading of flooding over the surface. The question is therefore asked whether it might be necessary to opt for a low resolution DEM in projects to lower the costs while still being able to model 2D overland flow. The choice for a 1D/1D model or a 1D/2D model is therefore strongly dependent on the availability (thus costs) of a DEM.

For a case study in Albergen, the Netherlands, five different models were compared (Stichting RIONED, 2014). The models were tested for their ability to identify six locations that were pointed out as vulnerable. It was concluded that the introduction of a 2D surface model significantly improves the modelling of overland flow. An important remark is that the 1D/1D and 1D/2D model also included surface water, something which is outside of the scope of this research. The report from Stichting RIONED also presented the results for a case study in Nijmegen, the Netherlands. Four 1D sewer models were combined with different surface models (1D, 1D/1D, 1D/2D and 1D/2D+) (Stichting RIONED, 2014). The most detailed model, a 1D/2D+ model, was determined to be the best match with the known locations of flooding.

(Vojinovic & Tutulic, 2009) compared the use of a 1D model with a 1D/2D model for the assessment of flood damage. The (available) resolution of the DEM, longer time to set up and increased computation time were pointed out as factors of influence on the choice for the type of model. For the chosen test case (with a steep topography), a 1D/2D approach was determined to be the better option. A comparison between a 1D/2D model and a 1D/2D+ model is presented in (Pina, et al., 2016). The article concluded that 1D/2D+ models tend to overestimate surface flooding, especially in residential areas. In comparison with a 1D/2D model, a 1D/2D+ model results in more flooding around houses and less street flooding. A lack of information on private connections and the representation of topographic depressions are pointed out as reasons for this. As 1D/2D+ models therefore overestimate the amount of rain water that doesn't reach the sewer system, sewer flow is often underestimated. However, with accurate data and proper calibration, it was concluded that 1D/2D+ models are able to represent urban drainage more realistically than 1D/2D models.

In general, the available comparisons conclude that investing time (thus money in a corporate environment) in a more complicated model provides benefits for the modelling results. Specifically, the addition of a surface model is a significant improvement of the quality of a model. However, more complicated models with more parameters require much data to calibrate, require a large amount of input data (DEM, sewer geometry, gully locations, land use) and need more time to run.

4.2 Dutch practice in modelling urban drainage

In the Netherlands, data availability for creating an urban drainage model is generally good and therefore not a limiting factor in the creation of models. Most municipalities have a database of their sewer network. As gullies are regularly cleaned, a database of gullies is also often available. *Kennisbank Stedelijk Water* provides design procedures, design storms and model parameters for creating a model (Stichting RIONED, 2017a). *Algemeen Hoogtebestand Nederland* (AHN) is a freely available DEM with a resolution of 0.5 m (Actueel Hoogtebestand Nederland, 2018). *Basisregistratie Grootschalige Topografie* (BGT) is a freely available database with information on land use (Kadaster, 2018). Additional information on these three sources is provided in Appendix A.

In (Leandro, et al., 2009), it was stated that the choice between a 1D/1D and a 1D/2D model is dependent on the cost and availability of a DEM. The quality of the DEM was also pointed out as a significant influence of the modelling results by other comparisons (Vojinovic & Tutulic, 2009) (Adeogun, et al., 2015). Because of the freely available DEM in the Netherlands, 1D/2D models are usually preferred over 1D/1D models. The combination of a very flat landscape, a high resolution DEM and a small chosen grid size allow for a detailed representation of the topography, including the smaller topographic depressions. The effect of incorrect representation of small topographic depressions and the resulting surface storage was pointed out as an influence on the modelling results of a 1D/2D+ model in (Pina, et al., 2016). It will be interesting to investigate whether the differences between the main classes of modelling concepts that were found are also found using Dutch test cases, and whether certain modelling concepts are more of less beneficial in the tested cases.

4.3 Aspects to be researched

The main objective of this research is to analyse the differences between the modelling concepts most often used in the Netherlands. In order to give a complete overview, it is decided to create four 'standard' models, based on the main classes of modelling concepts that are discussed in chapter 2. These standard models are a 1D, 1D/2D, 1D/2D+ and a 2D model. As in Dutch practice, 1D/1D is not applied very often (due to the flat landscape and readily available DEM), a model based on this concept is left out. Even within the main classes of modelling concepts, a lot of decisions need to be taken by the modeller, especially in the field of infiltration and surface runoff. Most of the articles discussed in this research therefore focus on the representation of the surface network and the modelling of runoff. Four additional aspects are pointed out as possible causes of differences between models. An additional model (see § 4.4) is created to analyse each of these aspects.

4.3.1 The inclusion of pervious surface area with semi-distributed runoff

In the Dutch practice of modelling with semi-distributed runoff (1D, 1D/2D), pervious surface area (grassland, parks) is not always taken into account and added to the sub catchments. It is assumed that runoff from pervious surface area is negligible. The combined area of the individual (impervious) land covers (roads, squares, buildings) is smaller than the total area of a sub catchment. For storms with a multiple-year return period, it is plausible that the rain intensity causes a noticeable amount of runoff from pervious surface area. Two setups are therefore tested: a 1D/2D model where only impervious surface area is taken into account and a 1D/2D model where also pervious surface area is added to sub catchments. Figure 11 gives a schematic overview of the two setups.

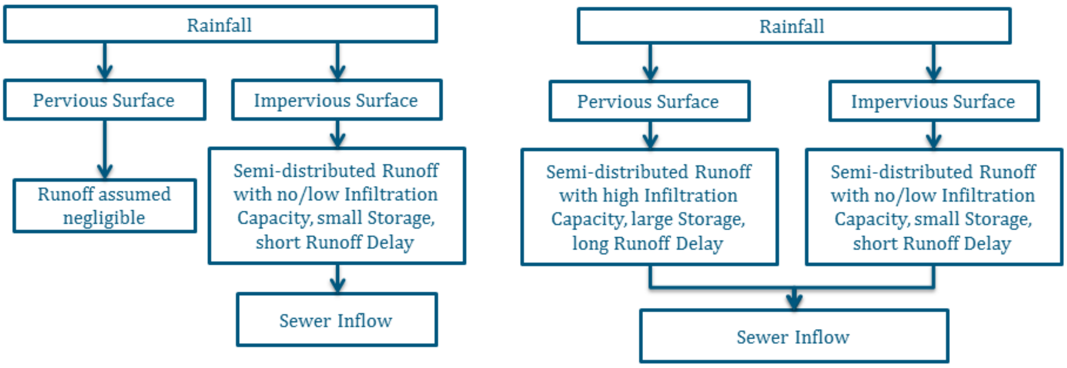


Figure 11 - Schematic overview of the two setups: on the left, it is assumed that no rainfall runs off from pervious surface types. On the right, also pervious surface area is taken into account.

4.3.2 The representation of roof surfaces in 1D/2D+ models

As the roofs of buildings are not always represented accurately in a DEM, they are often modelled as a flat elevated plane. When rain falls onto a building in such a model, it runs off and accumulates around it as rain gutters are not modelled. In a real situation, rain gutters collect the water from roof surfaces. The water is then drained directly to the sewer system by a vertical connection, to an impervious surface from where it flows towards gullies or to a pervious surface where the water can infiltrate. The different mechanism implies a different runoff delay and amount of rain water that enters the sewer system. In (Pina, et al., 2016), the representation of drainage from buildings was pointed out as a significant factor of influence of the modelling results of a 1D/2D+ model. Two different setups are tested: a setup where, as usual in a 1D/2D+ model, all surfaces are modelled with fully distributed runoff and a setup where houses are modelled with semi-distributed runoff. Figure 12 gives a schematic overview. The combination of two different ways of modelling runoff will be referred to as hybrid runoff.

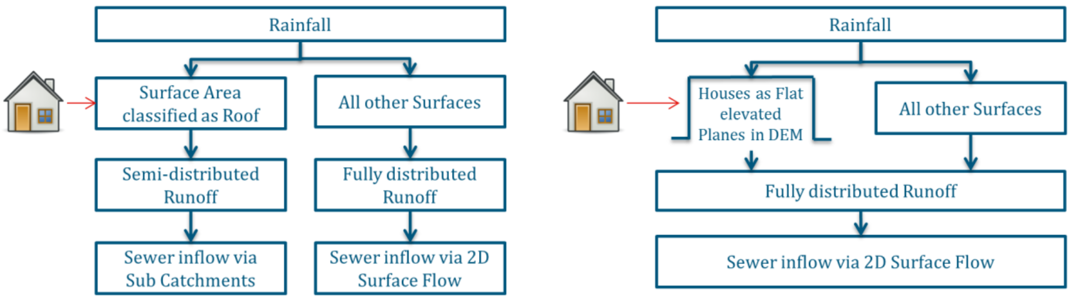


Figure 12 - Schematic overview of the two setups: on the left, runoff from houses is modelled semi-distributed. On the right, houses are modelled as flat elevated planes.

4.3.3 The influence of paved yards on the amount of street flooding

Usually, the classification of surface area is based on the maps provided by the Dutch BGT. The BGT does not differentiate between paved or unpaved front and back yards (and some other private terrain from companies, which from now on will be included in the term 'yards'). As no distinction is possible, the terrain is often not taken into account (semi-distributed runoff) or assumed pervious (fully distributed runoff). This implies that runoff from yards has no noticeable influence on the amount of street flooding. The aspect is tested by comparing two setups: a setup with a 1D/2D+ model where all yards are classified as pervious and a 1D/2D+ model where yards are sub classified into pervious, semi-pervious or impervious. Figure 13 gives a schematic overview of the setups.

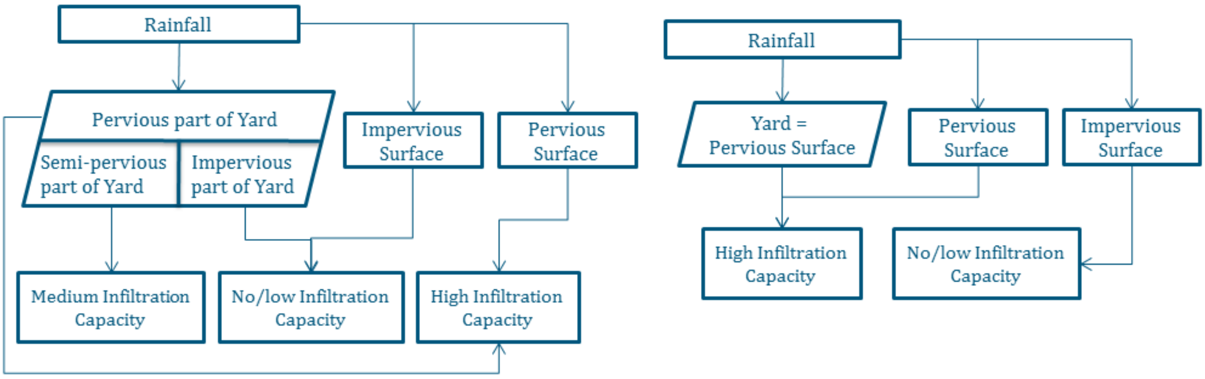


Figure 13 - Schematic overview of the two setups: on the left, yards are further classified into pervious, semi-pervious and impervious. On the right, all yards are assumed pervious.

4.3.4 The influence of infiltration zones on the results of a 2D model

In the simplest 2D models, only a DEM is used. As no geometric representation of the sewer system nor infiltration is present, water will accumulate in topographic depressions of the elevation model. Street flooding due to a locally underperforming sewer system or a lack of infiltration capacity will most likely not be detected. By using infiltration zones, areas with a high degree of imperviousness will result in a relatively higher amount of street flooding. The impervious surface area will be 'assigned' an infiltration capacity equal to a fixed discharge capacity to simulate sewer discharge. Figure 14 gives a schematic overview of the setups.

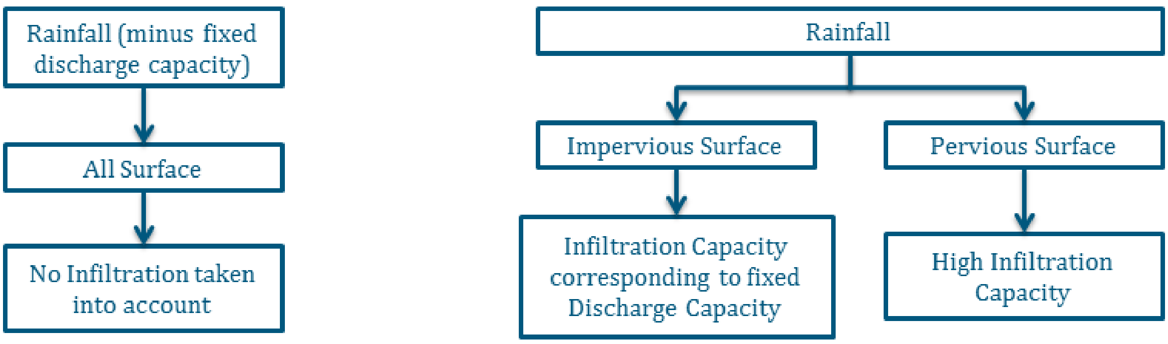


Figure 14 - Schematic overview of the two setups: on the left, no infiltration is taken into account. On the right, the surface area is divided into impervious and pervious surface area.

4.4 Models used in this research

The eight models created for this research are presented in this paragraph. The models are based different modelling concepts. Four main modelling concepts are distinguished: 1D, 1D/2D, 1D/2D+ and 2D. The four additional aspects from § 4.3 make up four extra models. Each model will be assigned an ID-tag for easier recognition. The ID-tag corresponding to the model is provided in the title of the paragraphs and in the overview in Table 2.

The models are all implemented in InfoWorks ICM (version 8.0.6). The starting point for these models is an existing 1D sewer model, as 1D sewer models are common in the Netherlands and the models of the test cases (see chapter 5) are all 1D. The other models are created by adding elements to the 1D model or removing them, without further re-calibration of parameters. In InfoWorks ICM, evaporation is subtracted from the incoming rainfall. Its value is set at zero, as evaporation for large storm events is assumed negligible in comparison with rainfall and infiltration. Where possible, default parameters of the software are used. This allows for an easier addition of modelling concepts in future research and for minimizing differences in parameters between models. An overview of the input parameters for InfoWorks ICM is provided in Appendix E. Soil-dependant parameters and parameters for gullies are discussed in this paragraph.

4.4.1 Standard 1D model (1D_St)

The 1D_St model makes use of a 1D modelling concept as explained in § 3.1. For the formation of sub catchments, guidelines as presented in the *Leidraad Riolering C2100* (Stichting RIONED, 2004) are used. In this guideline, twelve different categories for land use are presented. The guideline is used as a standard for most Dutch sewer models. In all three provided test cases, only four of the categories are used (as often in the Netherlands). Information on this guideline and the corresponding parameter values are found in Appendix B.

4.4.2 Standard 1D/2D model (1D/2D_St)

The 1D/2D_St model makes use of the concept explained in § 3.3 and uses the same land use categories as the 1D_St model. The AHN 2 is used as DEM. The AHN 2 has a resolution of 0.5 m. InfoWorks ICM makes use of a variable, triangular grid size. Two parameters are used for this: maximum triangle area and minimum element area. The values are chosen at 2.0 m^2 and 1.0 m^2 respectively for all models with a 2D surface grid. In this model, houses are excluded from the grid as the DEM does not properly represent houses. These holes in the DEM are called voids. This prevents unrealistic overland flow through houses, which will only occur when water depth exceeds the height of the front or the back door step. The value for the Manning's roughness of the surface is left at its default value: 0.0125.

4.4.3 1D/2D model including pervious area in sub catchments (1D/2D_P)

The 1D/2D_P model is created to investigate the aspect discussed in § 4.3.1: the inclusion of pervious surface area with semi-distributed runoff. To test this, the total surface area of each sub catchments is calculated. The combined area of the contributing impervious surface area is then subtracted from the total surface area. The remainder is added as pervious surface area (pervious, flat, stretched surface in Appendix B).

4.4.4 1D/2D+ model with hybrid runoff and pervious yards (1D/2D+_H)

The 1D/2D+_H model makes use of fully distributed runoff, except for roof surfaces, in order to investigate the aspect discussed in § 4.3.2: the representation of roof surfaces in 1D/2D+ models.

Roofs are modelled by making use of sub catchments. The roof areas that were determined for the sub catchments in the 1D model are used. The modelling of runoff is therefore a hybrid: mainly fully distributed with roof surface runoff modelled semi-distributed. For all test cases, the corresponding municipalities provided a database of gullies. Gullies are connected to the nearest manhole by a 125 mm conduit. The Q-h relation for inflow is based on numbers in (Martin, 2011). It is assumed that the cross-sectional slope is 1:50, the longitudinal slope is 1:100 and all gullies are of Type T (see the article for different types of gullies and definition of the types of slope).

Infiltration zones are created based on information from the BGT. Two types of surface are distinguished: impervious and pervious. Roads, squares and pavement are assumed impervious, all remaining surfaces are assumed pervious. This means that yards are all assumed fully pervious. This assumption is made to investigate the aspect discussed in § 4.3.3: the influence of paved yards on street flooding. As surface water modelling is not within the scope of this research, water courses are modelled as empty, impervious ditches. InfoWorks ICM requires four (Horton) input parameters. A surface storage constant is not required, as it is assumed that small topographic depressions are represented well enough to represent surface storage. Parameter values are based on research carried out by the US EPA (US Environmental Protection Agency, 1999) and are found in Appendix C.

4.4.5 1D/2D+ model with yard sub-classification (1D/2D+_Y)

The 1D/2D+_Y model uses the basic principles as explained in § 3.4. The 1D/2D+_Y model uses fully distributed runoff as houses are also included in the 2D surface grid as flat elevated planes. Infiltration zones are created based on the BGT. However, yards are further subdivided. By making use of satellite imagery, a distinction is made based on the NDVI-index of the surface. A four-band Triple sat image with a resolution of 80 cm is used for this (Netherlands Space Office, 2018). Appendix D goes into more detail on the classification process. In the end, yards are further divided into three categories: pervious, semi-pervious and impervious. For semi-pervious parts of a yard, both the Horton initial capacity (f_0) and final capacity (f_c) are halved. The rest of the surface area is classified the same way as the 1D/2D+_H model.

4.4.6 1D/2D+ model with hybrid runoff and yard sub-classification (1D/2D+_HY)

The 1D/2D+_HY model combines both the modelling of fast runoff from roof surfaces by making use of sub catchments and the subdivision of yards according to the NDVI index. By comparing the results of all 1D/2D+ based models, the aspects presented in § 4.3.2 and § 4.3.3 can be investigated.

4.4.7 Standard 2D model (2D_St)

The 2D_St model only makes use of a surface model. It follows the principles as discussed in § 3.5. The sewer system is compensated for by subtracting 20 mm/h from the incoming rainfall at each time step. Infiltration is not at all taken into account. As no infiltration, evaporation and sewer storage are included, the model will most likely overestimate the amount of street flooding, especially at the beginning of a storm event.

4.4.8 2D model including infiltration (2D_I)

The model replicates the sewer system by assuming all impervious surfaces have a constant infiltration capacity of 60 l/s/ha (21.6 mm/hr). This is a design standard often used for Dutch sewer systems (Stichting RIONED, 2004). Pervious surface area is assigned a constant infiltration capacity of 5 mm/h. As no sewer storage is modelled, it is expected that street flooding will occur (too) quickly in the first part of an event simulation. Besides this, rainwater from roof surfaces will fall into yards as houses are modelled as flat elevated planes (see § 4.3.2). The choice for infiltration capacity will influence how much these two processes influence the amount of street flooding.

4.4.9 Overview of models

Table 2 presents an overview of the eight models that were presented. Remarks regarding the table are given below.

#	ID	Runoff	Land use classes	Pervious surface	Land use source	Infiltration capacity source
1.	**1D_St**	*Semi-distributed*	*4*	*No*	*BGT*	*Leidraad C2100*
2.	**1D/2D_St**	*Semi-distributed*	*4*	*No*	*BGT*	*Leidraad C2100*
3.	**1D/2D_P**	*Semi-distributed*	*5*	*Yes*	*BGT*	*Leidraad C2100*
4.	**1D/2D+_H**	*Hybrid*	*2*	*Yes*	*BGT*	*Both*
5.	**1D/2D+_Y**	*Fully distributed*	*3*	*Yes*	*BGT + NDVI*	*US EPA*
6.	**1D/2D+_HY**	*Hybrid*	*3*	*Yes*	*BGT + NDVI*	*Both*
7.	**2D_St**	*Fully distributed*	*N/A*	*N/A*	*N/A*	*N/A*
8.	**2D_I**	*Fully distributed*	*2*	*Yes*	*BGT*	*Leidraad C2100*

Table 2 – Overview of the eight modelling concepts that are applied

Remarks regarding Table 2:

- '_St' stands for **St**andard model.
- '_P' stands for the inclusion of **P**ervious surface area in the sub catchments.
- '_H' stands for **H**ybrid runoff.
- '_Y' stands for **Y**ard sub-classification.
- '_I' stands for **I**nfiltration zones.
- In the column 'Pervious surface', a 'Yes' means that pervious surface area is taken into account of the formation of sub catchments or infiltration zones.
- More information on the sources for the infiltration capacity is found in Appendices B (Leidraad Riolering C2100) and C (US Environmental Protection Agency, 1999).

5. Case Description

In § 4.4, eight modelling setups were selected for the research. In this chapter, three test cases are presented (§ 5.2 - 5.4). A test case consists of a sewer system somewhere in the Netherlands. The models are initially tested with Dutch Design Storm 8 to reveal vulnerable locations. This storm represents a return period of two years and is a design standard for most Dutch sewer systems. More information on Dutch design storms is found in Appendix F. Aside from this, information from the corresponding municipalities is gathered on sewer monitoring and known locations of flooding, preferably for a certain (gauged) storm event.

5.1 Selecting test cases

For the evaluation of modelling concepts, three existing sewer models were selected: Ulvenhout, Tuindorp and Loenen. The models were selected because of their diverse characteristics in terms of topography and amount of impervious area. All three original models are 1D and correspond to the 1D_St model. All models are calibrated by setting up a monitoring campaign, using a similar approach to the calibration of sewer system the Hoven that is described in (Clemens, 2001). However, all models have been assigned the parameter set from the *Leidraad Riolering C2100* for the modelling of runoff (Stichting RIONED, 2004). This research assumes that all models are calibrated with a sufficient amount of data. An overview with information on the three cases is given in Table 3.

	Ulvenhout	Tuindorp	Loenen
Municipality	*Breda*	*Utrecht*	*Apeldoorn*
Year of monitoring campaign	*2002-2003*	*2014-2015*	*2001*
Total area sub catchments	*144 ha*	*140 ha*	*approx. 80 ha*
Impervious area	*59 ha (41%)*	*60 ha (43%)*	*21 ha (26%)*
Population (model)	*4,800*	*10,656*	*2,100*
Ground level/surface level	*2 – 6 m*	*1 – 2 m*	*17 – 29 m*
Type of sewer	*Mainly combined*	*Combined*	*Combined*
Average sewer gradient	*2.94 mm/m*	*2.24 mm/m*	*3.38 mm/m*
Dominant soil type[4]	*Sand (medium size)*	*Sand (medium size)*	*Sand (medium size)*
Main source(s) of information	(Langeveld, 2004)	(van Bijnen, et al., 2017)	(Henckens, et al., 2003) (Langeveld, 2004)

Table 3 – Basic information on the sewer systems used in this research

5.2 Ulvenhout

Ulvenhout is a village just south from the city of Breda. The village has a flat topography and a high groundwater table (DINOLoket, 2017), which is common in most parts of the Netherlands. The sewer system is maintained by the municipality of Breda. The sewer model was calibrated after a monitoring campaign in 2002 and 2003. Figure 15 gives an overview of the sewer system. The provided 1D_St model is tested with Dutch Design Storm 8 (Bui 08) and flooded manholes are depicted. The resulting water depths are a good measure for the occurrence of flooding, but have no further meaning for the extent of the flooding. Table 4 shows the case specific information that was available and is used for this research. Combined with pictures and videos shared on social media, it was possible to mark a number of streets where flooding occurred in 2016. These streets are shown in Figure 15. In the same figure, the modelling results of the 1D_St

[4] Based on drill cores from (DINOLoket, 2017)

model according to Bui 08 are shown. Two flood-prone areas are encircled: in the south (A1, Molenstraat) and in the centre of Ulvenhout (A2, Craenlaar/Annevillelaan). The vulnerable areas are both detected by the model and the video footage.

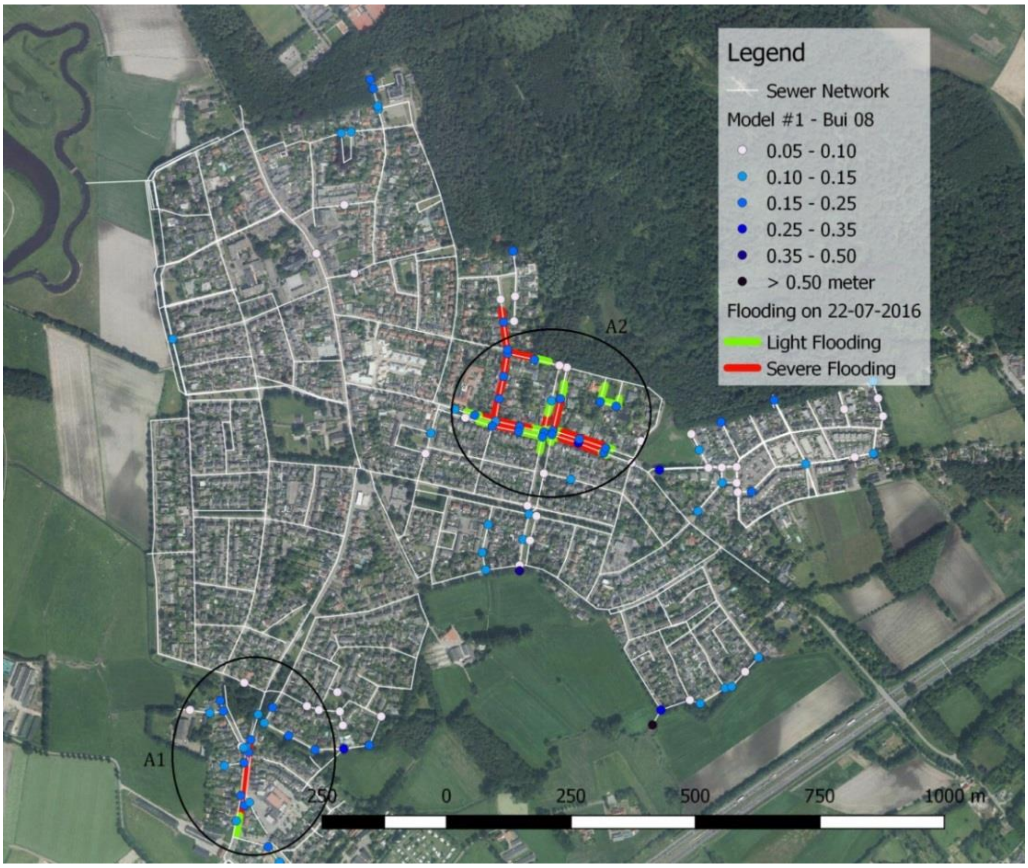


Figure 15 – Sewer system Ulvenhout. The flooded manholes according to the 1D_St model (model #1 in the legend) are shown, together with information on a storm event on 22-07-2016. Areas A1 and A2 are vulnerable areas detected by both the model and the footage from the 2016 storm event.

Event	Rainfall	Rainfall Duration	Peak intensity	Monitoring Data	Flood Data
July 2016	*51 mm*	*5 h*	*35.8 mm/h*	*A pump sump*	*Video footage, social media*
July 2017	*51 mm*	*13 h*	*27.9 mm/h*	*A pump sump*	*None*

Table 4 – Available case specific information for sewer system Ulvenhout

Remarks regarding Table 4:

- Water level is monitored at multiple locations in the sewer system, mostly at weirs and pumps. At four locations, the data is automatically sent to the municipal database. However, only one location is deemed usable for this research. Two locations are heavily influenced by groundwater (R.J.A.M. Franken 2018, personal communication, 29 January), while the InfoWorks ICM model does not simulate groundwater. The other location is at a sewer storage tank, where the modelled and measured water levels differ throughout the entire time-series. This is possibly explained by modifications being implemented in the model over the last years without re-calibration (V. de Bont 2018, personal communication, 23 February).
- During a flood event in July 2016, an employee from the municipality drove through the area where a lot of nuisance occurred and filmed the flooded streets.

5.3 Tuindorp

Tuindorp is a neighbourhood of the city of Utrecht. The neighbourhood has a very flat topography. The sewer model was calibrated by Marco van Bijnen as part of a research on sewer defects (van Bijnen, et al., 2017). A monitoring campaign with data from 2014 and 2015 was used to calibrate the provided model and this data was made available for this research. A map with flooded streets, corresponding to a storm event in 2013, was also provided. Figure 16 gives an overview of the system. Table 5 shows the case specific information that is used in this research. The vulnerable streets and modelling results according to the originally provided 1D_St model are depicted in Figure 16. The vulnerable area in the northern part of Tuindorp (B1, Albrecht Thaerlaan) is also recognized by the 1D_St model and with Dutch Design Storm 8, although the exact streets do not entirely match. The vulnerable street in the eastern part of Tuindorp (B2, Gerretsonlaan) is not recognized by the originally provided 1D_St model.

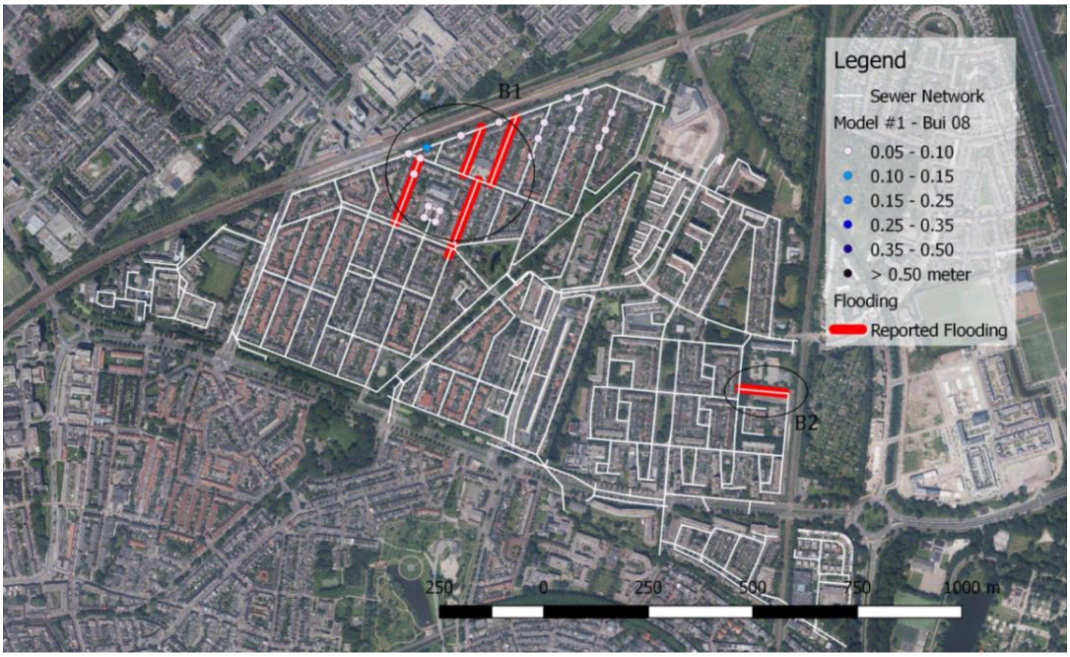


Figure 16 - Sewer system Tuindorp. The flooded manholes according to the 1D_St model (model #1 in the legend) are shown. The streets in red are known to be vulnerable according to call data and conversations with inhabitants.

Event	Rainfall	Rainfall Duration	Peak intensity	Monitoring Data	Flood Data
Nov 2013	*26 mm*	*2 h*	*43.2 mm/h*	-	*Map of flooded streets*
Sep 2014	*24 mm*	*5 h*	*58.8 mm/h*	*14 manholes*	*None*
Sep 2015	*33 mm*	*11 h*	*25.2 mm/h*	*14 manholes*	*None*

Table 5 – Available case specific information for sewer system Tuindorp

Remarks regarding Table 5:

- A map with flood prone streets was composed based on call data and some conversations with inhabitants after a storm event in November 2013 (M. van Bijnen 2017, personal communication, 27 November 2017).
- Data from the monitoring campaign that was used to calibrate the model is used. At fifteen locations, monitoring data was available for both storm events. At one location, the initial water level differed too significant and large differences between model and measurements were observed. Therefore, fourteen locations were used.

5.4 Loenen

Loenen is a village that is part of the municipality of Apeldoorn. The village has a mildly sloping topography and a deep groundwater table (DINOLoket, 2017), which is only common around the Veluwe, the Utrechtse Heuvelrug and in the south of the Netherlands. The sewer system was calibrated based on a monitoring campaign from 2001. Figure 17 gives an overview of the sewer system. The provided 1D_St model is tested with Dutch Design Storm 8 and the flooded manholes are depicted.

At the municipality, only one location of frequent flooding is known (J. Veurink 2017, personal communication, 14 December). The location (Hoofdweg 60) is added in red. The location is not recognized by the 1D_St InfoWorks ICM model using Dutch Design Storm 8. A possible reason for this could be the chosen time step of 60 seconds (J. Veurink 2018, personal communication, 9 May). A WOLK-model was created on behalf of the municipality. In Appendix G, the vulnerable areas according to this model are shown, together with a brief description of the software package. The results of this model are not used in this research, as this research does not compare models with results from other urban drainage models.

The municipality was not able to provide in-sewer monitoring data. Therefore, no comparison of in-sewer monitoring with modelled results is possible.

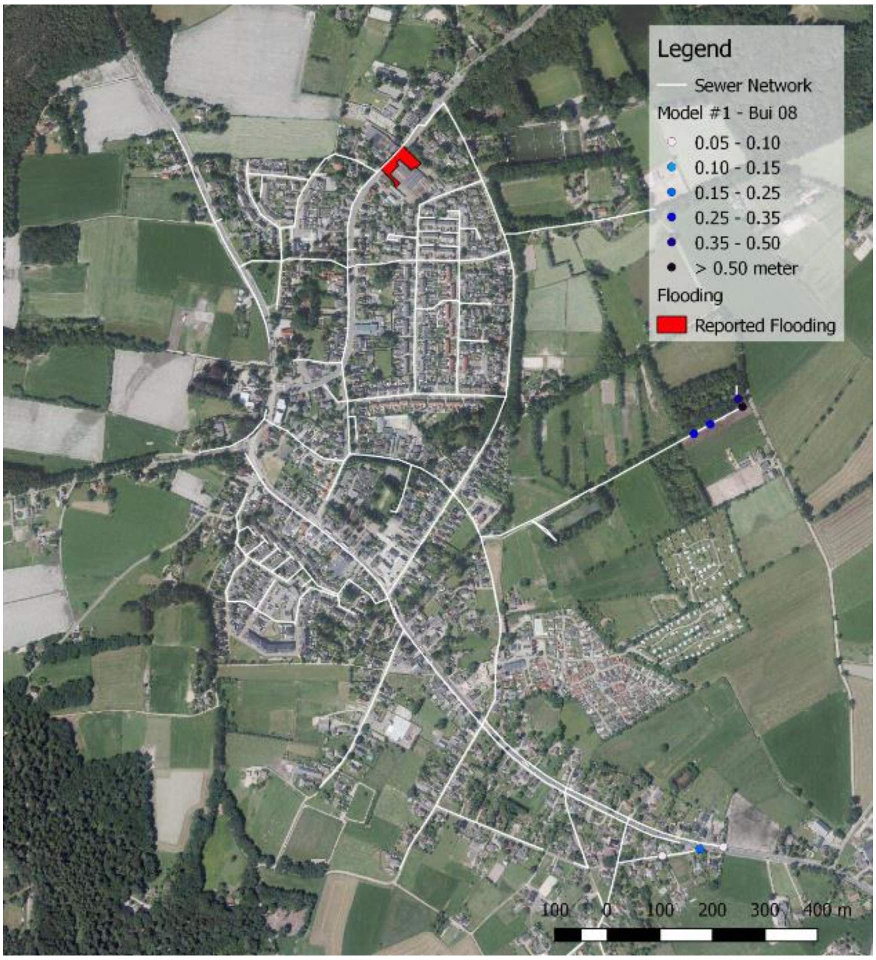


Figure 17 - Sewer system Loenen. The flooded manholes according to the 1D_St model (model #1 in the legend) are shown, together with the only location of frequent flooding known at the municipality.

6. Method for comparing Modelling Concepts

This chapter presents a method to assess the differences between the eight different models that were created. Usually, two types of data are available to compare a model with observations: in-sewer monitoring and data on flooding. The data has different sources and serves different purposes in assessing the performance of models. The method therefore consists of two parts: a comparison of known locations of flooding with modelling results (§ 6.1) and a comparison of modelling results with in-sewer monitoring data (§ 6.2). As defined in the research scope, the method should be transparent, objective and reproducible. In such a way, the method can then be used in future research for adding more modelling concepts, using different test cases or improving the method used.

6.1 Test 1: Known locations of flooding

The first test consists of a comparison of modelling results with locations that are prone to flooding, preferably linked to a certain storm event. The aim of this test is to assess the ability of the models to match the locations that are prone to flooding in the actual situation. The results of the test are expected to provide information on the general tendency of a model to overestimate or underestimate the amount of street flooding and the influence of the various differences in the modelling concepts on the recognition of flood prone areas. The maximum street flooding is assessed for a certain storm event. This is a built-in function in InfoWorks ICM. There is a difference in the assessment of the 1D_St model compared to the other models, as the 1D_St model does not contain a surface model (see § 3.1 for explanation on 1D models). § 6.1.3 describes the steps that are taken and the rules that are applied for models with and without surface model.

6.1.1 Sources of information for observed flooding

Flood prone locations can be based on various sources of information. Sources can include municipal call data, knowledge from sewer managers, social media, aerial images or footage taken by the municipality. All of this data has a different reliability and spatial accuracy. It is therefore important to funnel the data into a concept in which the observation data is comparable with the modelling results. Since damage estimation is not within the scope of this research, the relevant part of information is the extent (area) of the flooding. For the extent of surface flooding, information sources can be divided into three main categories:

- Point data: Specific locations in a catchment, such as addresses, manholes or gullies.
- Line data: Elongated locations, such as streets, tunnels or highways.
- Polygon data: Flooded planes, such as squares, soccer fields or parking lots.

6.1.2 Comparing modelling results with observations

By making use of a grid, all three types of data for observed flooding (point data, line data and polygon data) can be added in the same figure. Both the observed flooding and the modelled flooding can be added in a grid and compared with each other.

When information is available on flooded locations according to a specific storm event, it is also valuable to know locations where no flooding occurred. Almost never, each flooded location in a whole village or neighbourhood is known for a certain flood event. It is therefore hard to judge the ability of a model to recognize only the flooded locations, as other flooded areas may not be recorded or observed. When in a certain area and for a certain storm event, the extent of the flooding is well documented (both flooded and non-flooded locations are known), it can be assessed whether the model overestimates or underestimates the amount of surface flooding, without the need for a 100% coverage of the catchment. Based on this principle, the proposed method in test 1 is an analysis of the model's ability to point out flooded locations without overestimating or underestimating the amount of surface flooding. The terms that are used in the comparison of model and observations are shown in Table 6.

	Observation: Flooding	Observation: No Flooding
Model: Flooding	*True Positive (correct)*	*False Positive (type I error)*
Model: No Flooding	*False Negative (type II error)*	*True Negative (correct)*

Table 6 – Terms used in grid comparison, based on jargon used in hypothesis testing

The information that is provided does not always include non-flooded areas and the information is not always linked to a specific flood event. For instance, municipal call data does not include non-flooded locations, since this gives no reason to call. Inhabitants or sewer managers may provide information on flood-prone locations, but they cannot always point out a specific day on which nuisance occurred.

In the case that information on flooded locations is not linked to a specific flood event, it is not possible to judge the model's ability to accurately point out flood prone areas without underestimation or overestimation. Dutch Design Storm 9 (Bui 09), a storm with a return period of five years (see Appendix F), is used in these test cases. As this storm event is more intense than the design standard for Dutch sewer systems (Bui 08), street flooding will most likely occur in vulnerable areas according to the model. By comparing the flooded locations according to Bui 09 with the known information on flood prone areas, it can be assessed whether the model has the ability to point out the same flood prone areas. It is however not possible to assess locations where no flooding occurred (false positive or true negative), as the extent of the flooding is not comparable in any way to the observed flooding. The assessment is therefore only based on locations where flooding did occur (true positive or false negative). The accuracy, determined by the percentage of true positives, is likely to be lower compared to a case where information is linked to a specific flood event.

In the case that information is linked to a specific flood event but no information is available on non-flooded areas, it is also not possible to assess overestimation and underestimation. In that case, only flooded locations are assessed (true positive or false negative). Figure 18 shows the different options in the form of a decision tree.

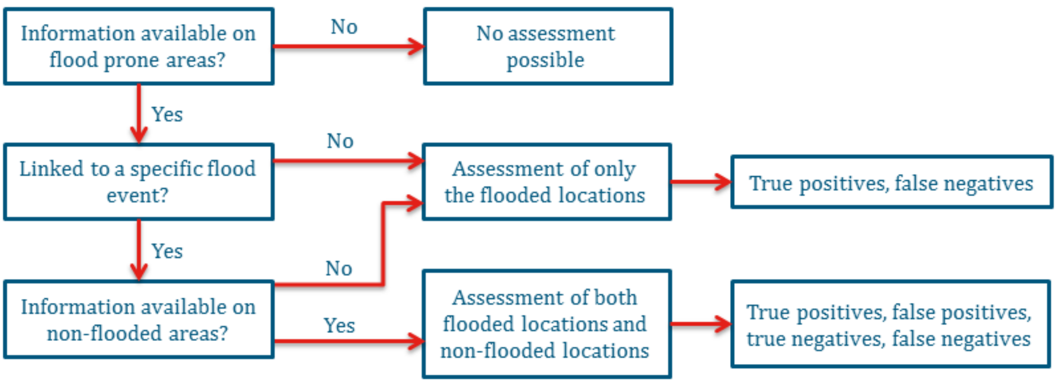


Figure 18 – Overview of the process of assessing known locations of flooding

For the test case Loenen, the information on flood prone areas is not linked to a specific flood event. Bui 09 is used to assess flooding. For test case Tuindorp, the information on flooding is linked to a specific flood event. However, no information on non-flooded areas was available. Therefore, (also) only flooded locations are assessed. In the case of Ulvenhout, case-specific information is available and both flooded and non-flooded locations are known.

6.1.3 Steps of test 1

Six steps are taken for the method in test 1. There is a difference in the assessment between the 1D_St model and all other models. The results of the 1D_St model are in the form of water volumes or depths in an artificial cone, while the other models all contain a surface model and therefore topographically realistic flood depths. Step 4 does not apply to the 1D_St model. Step 5 distinguishes between the two types of modelling results.

1. **Known locations of flooding (and no flooding) are plotted in a GIS map.**

2. **A square grid is placed over the area.**

A square grid is easily produced in a GIS program and has a constant size, area and shape. The grid size selected should be in the same order of magnitude as the accuracy of the available flood data. A too large grid size causes loss of data, as multiple streets or houses may be present in one grid cell while only one was flooded during a storm event. A too small grid size causes inaccuracy, as the accuracy of the extent of the flooding may be smaller than the grid size. Grid cells near the edge of a flooded plane or street may be assumed flooded, while the information only contained a rough estimate.

A default grid size of 30 x 30 m is chosen in the assessment. However, as the hypothesis is that the grid size will influence the results, four grid sizes are tested in the sensitivity analysis: 20 x 20 m, 30 x 30 m, 40 x 40 m, 50 x 50 m.

3. **The grid cells are marked 'flooded' or 'not flooded' according to known locations of flooding.**

The grid cells are now marked according to the flood map (step 1). It is assumed that data has 100% spatial accuracy and all water depths are treated equal. The values for the boundary conditions are based on (minor) user experience and are not based on actual scientific literature as there is no documented literature on a quantitative comparison of this form. Also, a full sensitivity analysis for all input parameters is not within the scope and timeframe of this research. The following set of rules applies:

- Point data: a grid cell in which point data falls is marked as 'flooded' or 'not flooded'.
- Line data: in the case of line data, 50% of the width of a grid cell should overlap a location of flooding or no flooding. For instance, if a street with a length of 20 meters experienced flooding during a storm event and the street falls entirely within a 40 meter grid cell, that grid cell is marked 'flooded'. In the same situation, but with a street length of 15 meters, the grid cell is not marked.
- Polygon data: polygon planes should cover at least 10% of the area of that grid cell. For instance, if a grass field of 300 m^2 that experienced no flooding falls within a 50 meter grid cell (2500 m^2), the grid cell is marked 'not flooded'.

4. A filter is applied to the modelling results to account for the difference between water on the streets and flooded back yards (Only models with surface model).

Models with fully distributed runoff (1D/2D+ and 2D) tend to show more flooding in backyards compared to models with semi-distributed runoff (1D and 1D/2D). A reason for this is that gullies in fire breaks are often not modelled. Also, the representation of roof surfaces influences this (Pina, et al., 2016). As the (easily overestimated) flooded back yards should not be recognized as flooded streets, the data is filtered. A polygon with main roads is used as a mask to only select the flooding in this area. The polygon includes the possibly known flood polygons (squares, soccer fields, parking lots) and the main roads plus a four meter barrier for sidewalks and front yards.

5. Grid cells are marked 'flooded' or 'not flooded' according to modelling results. False negatives and false positives are counted.

The grid cells were marked flooded or not flooded according to the observations in step 3. In this step, the grid cells are marked flooded or not flooded according to the modelling results. The marks are then compared:

- If a grid cell is marked 'not flooded' based on both the modelling results and the flood data, the result is *true negative.*
- If a grid cell is marked 'flooded' based on both the modelling results and the flood data, the result is *true positive.*
- If, based on the flood data, a grid cell is marked 'flooded' while it is marked 'not flooded' on basis of the modelling results, the result is *false negative.*
- If, based on the flood data, a grid cell is marked 'not flooded' while it is marked 'flooded' on basis of the modelling results, the result is *false positive.*

As there is a difference in the modelling results of the 1D_St model and the other models with a surface map, two different sets of rules applies for marking the grid cells. The difference will be taken into account in the discussion of test results.

- **Rules for models with a surface model (1D/2D, 1D/2D+ and 2D)**

If at least an area A_f with at least a flood depth of h_f is present in a grid cell, the grid cell is marked 'flooded'. If less than area A_f with at least a flood depth of h_f is present within a grid cell, the grid cell is marked 'not flooded'.

The completeness of the test is based on the the available information (are areas of non-flooding included in the flood data? Is the data linked to a specific storm event?). The hypothesis is that the thresholds that are used for depth and area (A_f, h_f) have a significant influence on the differences between the scores of the eight models. Too low thresholds cause all models to score high, potentially decreasing differences in scores between models. Too high thresholds cause low scores, also potentially decreasing differences between models. Furthermore, 'the right' thresholds could differ for different storm events and areal characteristics.

As a default, the thresholds A_f and h_f are chosen at 5 m^2 and 10 cm respectively for test cases Ulvenhout and Loenen. As the 2013 storm event in Tuindorp caused less rainfall, the thresholds A_f and h_f are chosen at 5 m^2 and 5 cm respectively. In total, four combinations of thresholds are tested in the sensitivity analysis for all test cases:

- Setup 1: A_f (areal threshold) = 5 m^2, h_f (depth threshold) = 5 cm.
- Setup 2: A_f (areal threshold) = 5 m^2, h_f (depth threshold) = 10 cm.
- Setup 3: A_f (areal threshold) = 1 m^2, h_f (depth threshold) = 10 cm.
- Setup 4: A_f (areal threshold) = 5 m^2, h_f (depth threshold) = 10 cm for assessing flooded locations. A_f = 1 m^2, h_f = 10 cm for assessing non-flooded locations. This means that in the assessment of flooding, at least 5 m^2 of area (with ≥ 10 cm water depth) should be present in a 'flooded' grid cell for a true positive judgement. In the assessment of non-flooding, 1 m^2 of flooded area (≥ 10 cm water depth) in a 'non-flooded' grid cell results in a false positive.

Figure 19 shows an example of the method for test 1 for modelling results of models with a surface map.

Figure 19 – An overview of the method that is used for test 1 and for models with a surface map. On the left, the flooded and non-flooded grid cells according to the observations are shown (step 3). The barrier named 'area of interest' is described in step 4. The right figure shows which observations match the modelling results according to the rules defined. In this case, setup 4 is chosen for the values of A_f and h_f.

- **Rules for the 1D_St model**

As manholes are usually only present on streets and squares, not every grid cell will contain a manhole. Furthermore, manholes are often distanced 40-50 meters from each other in the selected test cases. If a grid size of less than 50 meters is chosen, not every grid cell contains a manhole, even when the grid cell covers a street. In order to mark every grid cell, results need to be interpolated.

The results are interpolated by assigning a flood depth to each grid cell, equal to the nearest manhole. If the nearest manhole has a flood depth more than h_f in its cone, the grid cell is marked 'flooded' according to the modelling results. If the nearest manhole contains less water, the grid cell is marked 'non-flooded' according to the modelling results. If two manholes are present in one grid cell, the manhole with the largest h_f counts. The rule of using the value from the closest manhole will be referred to as the 'nearest manhole' rule.

As a default, the thresholds for h_f are chosen the same as in the method for models with a surface model: 10 cm or Ulvenhout and Loenen, 5 cm for Tuindorp. It should be kept in thought that as the flood cones do not represent an actual flood depth, the scores are not exactly comparable.

As an alternative rule, the 'surrounding cells' rule is proposed for testing in the sensitivity analysis. If no manhole is present in a certain grid cell, the eight surrounding cells are assessed. If one of the eight surrounding cells contains a manhole with a h_f larger than the threshold, the grid cell is marked flooded. Aside from the applied rule, two different depth thresholds are tested. In total, four setups are tested in the sensitivity analysis:

- Setup 1: h_f (depth threshold) = 5 cm, rule = 'nearest manhole'.
- Setup 2: h_f (depth threshold) = 10 cm, rule = 'nearest manhole'.
- Setup 3: h_f (depth threshold) = 5 cm, rule = 'surrounding cells'.
- Setup 4: h_f (depth threshold) = 10 cm, rule = 'surrounding cells'.

Figure 20 explains the differences between the two methods and shows an example of the rule for a group of grid cells that were not flooded according to observations.

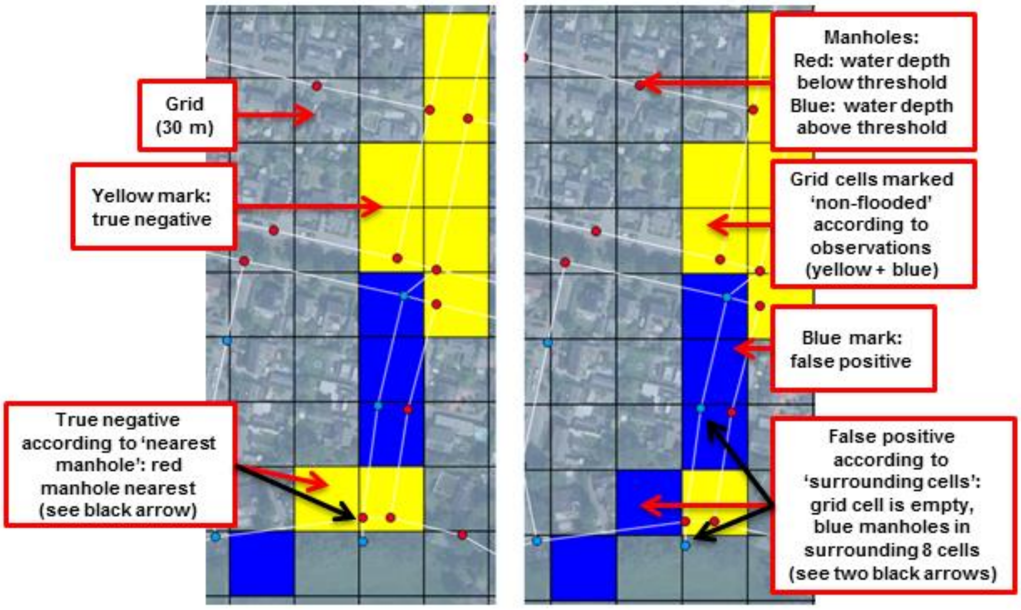


Figure 20 – An overview of the method that is used for test 1 and for the 1D_St model. A map is shown with the grid cells marked 'non-flooded' according to the observations. On the left, the 'nearest manhole' rule is used. On the right, the 'surrounding cells' rule is used. The difference in rules leads to a difference in the amount of true negatives: one manhole is marked differently.

6. Scores are calculated

The scores are normalized to the amount of marked grid cells (true positive, true negative, false positive and false negative). A total score is composed (Equation 2), together with partial scores for true negatives and true positives. For example: according to the observed locations of flooding, 30 grid cells were marked flooded and 40 were marked as not flooded (step 3). According to the modelling results, 24 out of the 30 grid cells were confirmed to be flooded (step 6, true positive) and 30 out of the 40 cells were 'correctly' recognized as not flooded (true negative). The total test score is now (24 + 30) * 100% / (30 + 40) = 77%. The partial scores for true positive and true negative are 80% and 75% respectively.

$$Test\ Score = \frac{n_{true\ positives} + n_{true\ negatives}}{n_{marked\ grid\ cells}} * 100\% \qquad \text{(Equation 2)}$$

If a good accuracy of the observed flooding is assumed, a difference in the scores for true positives (flooded grid cells) and true negatives (non-flooded grid cells) is a sign that the model overestimates or underestimates the amount of street flooding. For example, the score for true positives is 90% and the score for true negatives is 30%. This means that almost all the flooded grid cells are recognized correctly. However, also 70% of the non-flooded grid cells are recognized as flooded. The model therefore overestimates the amount of street flooding.

6.2 Test 2: In-sewer monitoring results

The second test compares modelling results with in-sewer monitoring data. The aim of this test is to study the ability of modelling concepts to represent the sewer dynamics during a storm event. The performance of the models during a storm event is expected to grant information on the influence of measures in infiltration and runoff on sewer dynamics and the tendency of a model to overestimate or underestimate sewer inflow. Models based on a 2D concept are not tested as they do not contain a geometric representation of the sewer system (see § 3.5). To quantify the agreement between modelled and measured water levels, three indicators are proposed: the Root Mean Squared Error (RMSE) (Martens & Magni, 2001), the Nash-Sutcliffe Efficiency (NSE) (Nash & Sutcliffe, 1970) and the Kling-Gupta Efficiency (KGE) (Gupta, et al., 2009). The indicators and their components are shown in equations 3-8:

$$RMSE = \sqrt{\frac{\sum_{t=1}^{n}(h_{s,t} - h_{o,t})^2}{n}} \qquad \text{(Equation 3)}$$

$$NSE = 1 - \frac{\sum_{t=1}^{n}(h_{s,t} - h_{o,t})^2}{\sum_{t=1}^{n}(h_{o},t - \mu_{o})^2} \qquad \text{(Equation 4)}$$

$$KGE = 1 - \sqrt{(r-1)^2 + (\alpha - 1)^2 + (\beta - 1)^2} \qquad \text{(Equation 5)}$$

$$r = \frac{Cov(h_s, h_o)}{\sigma_s * \sigma_o} \qquad \alpha = \frac{\sigma_s}{\sigma_o} \qquad \beta = \frac{\mu_s}{\mu_o} \qquad \text{(Equation 6, 7, 8)}$$

$h_{o,t}$ = observed water level at time t	*(m)*
$h_{s,t}$ = simulated water level at time t	*(m)*
n = number of time steps	*(-)*
μ = mean water level	*(m)*
r = correlation coefficient	*(-)*
α = variability ratio	*(-)*
β = bias ratio	*(-)*
$Cov(h_s, h_0)$ = covariance	*(m^2)*
σ = standard deviation	*(m)*

The RMSE is a measure for the difference between modelling results and monitoring data. Larger differences between model and measurement weigh more heavily on the RMSE compared to smaller differences because of the squared difference. The unit of the RMSE is a unit of length: meters or centimeters. It is therefore a good measure for the accuracy of a model. In an ideal case, the RMSE is zero. Besides the RMSE of the total storm event, each event is divided into four stages, roughly representing dry weather flow (DWF), filling up of the system, peak water level and emptying of the system. By doing this, differences in specific parts of the storm event can be studied.

The NSE and KGE are regularly used indexes in hydrological modelling. For both indexes, their ideal value is at unity. The NSE is the ratio of the difference between the modelled and observed values (numerator) and the difference between the observed value and the mean of the observations (denominator). If the NSE is larger than zero, it means that the model is a better estimator of water level (in this specific case) than the mean of the observations (a constant value).

The KGE is built up by three components: r, α and β. Coefficient r quantifies the correlation between the observed and modelled water levels. In an ideal case, the value is at -1 or +1. In that case, there is a perfect linear relationship between the observations and modelling results. A value of zero means that there is no linear correlation at all. α is the ratio between the variability of the observed and modelled water levels. If the value is higher than unity, it means that the modelled time series shows a larger statistical spreading in water levels than the observed time series. Larger deviations from the mean water level of the whole time series weigh more heavily on the variability ratio, as the standard deviation uses the quadratic difference between the mean and data points in the time series. β is the bias ratio, a measure of whether the model generally overestimates or underestimates the water levels. All data points have the same weight. As the total KGE value is one minus the length of the Euclidean vector (3D vector built from the 1D vectors made up by the three components), its value is most influenced by the value of the three partial components (r, α and β) that is furthest from unity. An extensive comparison between the NSE and KGE is presented in (Gupta, et al., 2009).

The functions are very similar to the ones used in the STOWA benchmark for modelling packages (STOWA, 2017). Two differences are made; the RMSE instead of the dimensionless Normalized RMSE (NRMSE) is used and the KGE proposed in the original literature is used (Gupta, et al., 2009) instead of the modified Kling-Gupta Efficiency (KGE') (Kling, et al., 2012). In the KGE', both the denominator and nominator in the α-term (Equation 7) are divided by μ_s and μ_o respectively to account for possible cross-correlation between α and β when precipitation data is biased. However, the value of μ_s/μ_o increases enormously when μ_s and μ_o are close to zero. As mean water levels are often close to zero in a flat, low-laying country like the Netherlands, this scenario will be true quite often. Therefore, α and β values can be very high (or even negative) in some manholes, dominating the KGE'. To counteract this, all water levels are corrected by adding the average depth of all manhole bottoms and the 'original' KGE is used. The bias ratio in rain measuring equipment is rather small (β = 0.994 compared with the municipal average for the 2017 storm in Ulvenhout, β = 0.937 for the two gauges in Tuindorp for the 2014 storm event). It is therefore assumed that bias in measuring equipment is small.

7. Results

This chapter gives an overview of the results of the two tests. The results of test 1, using the default test parameter settings, are presented (§ 7.1). Test parameters for test 1 are varied. The results of sixteen combinations of different test parameters are presented for each individual test case (§ 7.2). Finally, the results of test 2 are shown (§ 7.3). The results are discussed in Chapter 8.

In order to support the test results, three types of flood maps are added in the appendices. Appendix I presents flood maps of each individual model. Appendix J shows flood contour maps, where the extent of flooding of all models is plotted in the same figure. Finally, Appendix K depicts flood difference maps, where the differences between the models corresponding to the aspects determined in § 4.3 are shown. As infiltration capacity is mentioned multiple times as an important input for the models, Appendix N explores the differences in test results by varying two of the Horton infiltration parameters.

7.1 Test 1 - Default Test Parameters

Test 1 assesses the performance of a model based on its ability to correctly represent the extent of street flooding. A grid was introduced to plot all observed flooding (or no flooding) in the same format to be able to compare it with modelling results. Figure 21 shows the results for the assessment of flooded grid cells (true positives) and for all test cases. For test case Ulvenhout, both flooded (true positives) and non-flooded (true negatives) grid cells are assessed. These results are depicted in Figure 22.

The originally provided 1D_St model scores well in the recognition of both flooded and non-flooded grid cells for test case Ulvenhout. The combined score is 86%. However, scores are low for Tuindorp (4%) and Loenen (0%). The results of the 1D/2D_St model follow the same pattern. This was to be expected as the hydrological model is exactly the same. The 1D/2D_P model, where pervious surface is added to the sub catchments, scores higher in the recognition of flooded grid cells for the Tuindorp and Ulvenhout test case. As additional runoff from pervious surfaces is added to the sub catchments, more street flooding was to be expected. Although the 1D/2D_P model scores 100% in the recognition of flooded grid cells for test case Ulvenhout, the model scores significantly less in the recognition of non-flooded grid cells (57%). The difference indicates overestimation of street flooding.

There is a considerable difference between the results of the three models based on a 1D/2D+ concept. The 1D/2D+_H model scores lower in the recognition of flooded grid cells than the other models. The model does not recognize the flooded location in Loenen and the model scores lower for the test cases Tuindorp (71% compared to 75% and 83%) and Ulvenhout (39% compared to 86% and 89%). For the Ulvenhout test case, a 97% is reached for the recognition of non-flooded grid cells. The difference in the scores for true positives and true negatives implies that the model underestimates the amount of street flooding for the chosen test parameters. The other two 1D/2D+ models score similar to each other. The 1D/2D+_HY model scores slightly higher for test case Ulvenhout and the 1D/2D+_Y scores higher for test case Tuindorp. The scores for true positives and true negatives are balanced, indicating that the models do not overestimate or underestimate flooding with the chosen test parameters.

In all test cases, the 2D_I model scores near-perfect in the recognition of flooded grid cells. The scores for true positives range from 96% to 100%. For test case Ulvenhout, the model scores lower in the recognition of non-flooded grid cells (63%), indicating overestimation of street flooding. The 2D_St model scores lower for test case Ulvenhout and Tuindorp in the recognition of flooded grid cells. The combined score for test case Ulvenhout is equal, both score 80%.

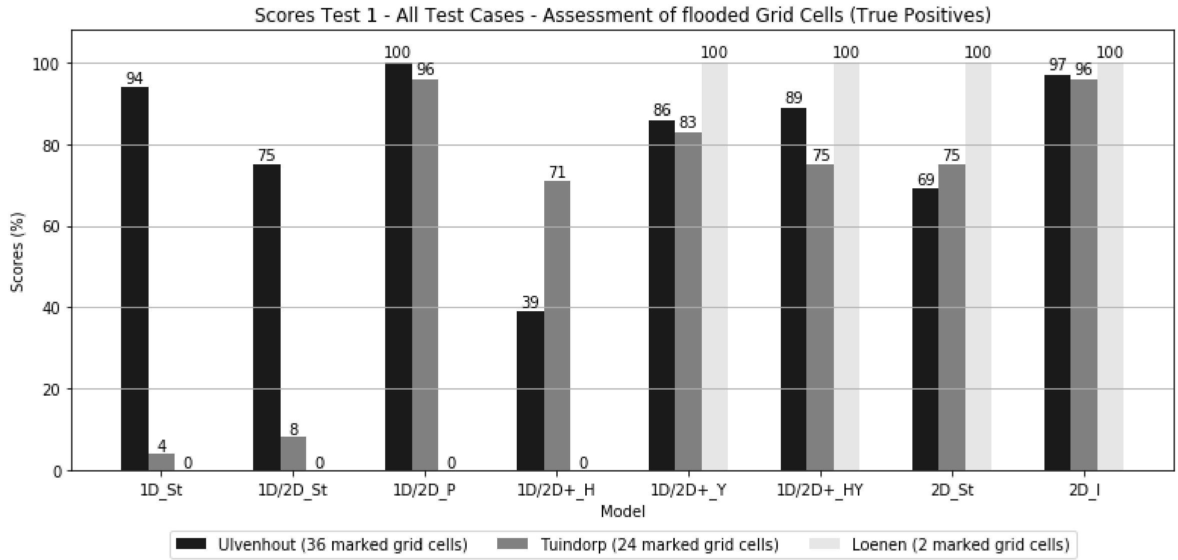


Figure 21 – Results of test 1 using the default settings for test parameters. Only the flooded grid cells are assessed (true positives), as this was tested in every test case. For test case Loenen, considerably less data was available compared to the other cases.

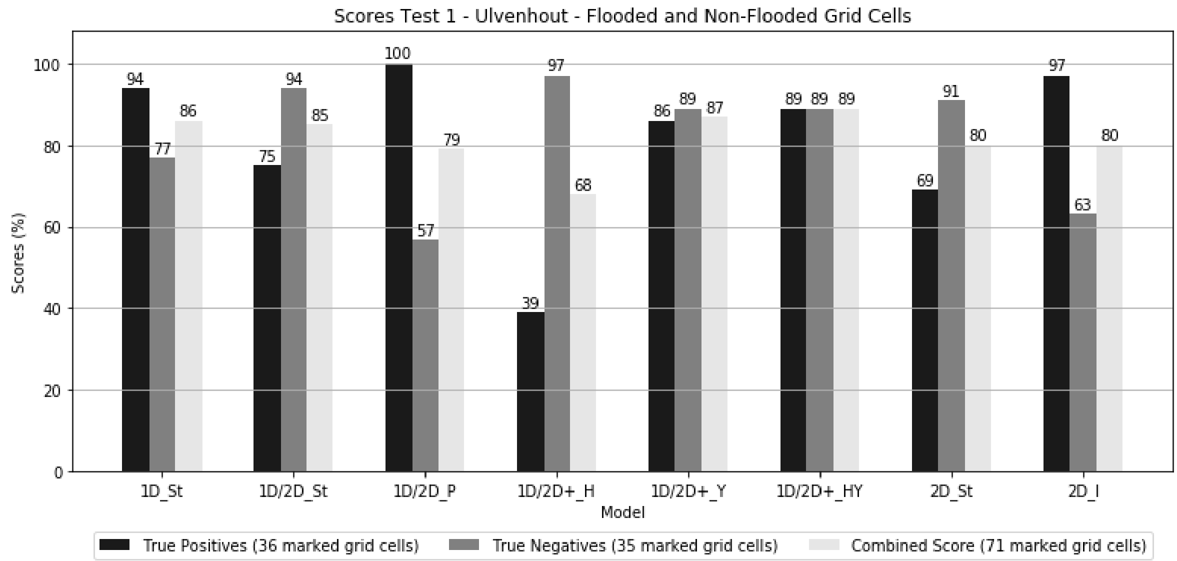


Figure 22 – Results of test 1 for test case Ulvenhout, using the default settings for test parameters. As information was available on both flooded and non-flooded locations, both true positives and true negatives are assessed.

7.2 Test 1 - Sensitivity of Test Parameters

This paragraph contains the results of test 1 for different test parameters. For the assessment of flooded and non-flooded grid cells, multiple combinations of test parameters are used. h_f is the depth threshold for flooding, A_f is the areal threshold for flooding, a certain grid size is used and the different types of data (point data, line data and polygon data) also contain parameters. In §

6.1.3, a default grid size of 30 m was selected. For the models with a 2D surface model, the areal threshold was set at 5 m^2 and depth threshold at 5 or 10 cm depending on the test case. For the 1D_St model, the same depth thresholds were chosen, together with the remark that water depths in models based on a 1D concept are not topographically correct. The results for the sensitivity of parameters for grid size, minimum depth and minimum area are shown in this paragraph. Four grid sizes were tested: 20 x 20 m, 30 x 30 m, 40 x 40 m and 50 x 50 m. Aside from the grid size, four combinations of test parameters for depth (h_f) and area (A_f) are selected for the models with a 2D surface model. For the 1D_St model, two different depth thresholds and two different methods of assessing flooding are tested. In total, the four grid sizes and four setups add up to sixteen configurations of test parameters, both for the 1D_St model and the urban drainage models with a surface model. The results are represented in tables (Table 7 – Table 16) and summarized in box plots (Figure 23 – Figure 27). Results are discussed per test case.

7.2.1 Test case Ulvenhout

The results for the combined score of all sixteen configurations are shown in Table 7 and Table 8. The corresponding box plot is shown in Figure 23. Considering this combined average score, the 1D/2D_St scores highest with an average score of 81%. This means that the model has an 81% chance of scoring either a true positive or a true negative, considering the average of all sixteen configurations and in case of the 2016 storm event. The 1D/2D+_Y and 1D/2D+_HY model follow closely with an average score of 77%. The 1D_St model scores 73% on average. All other models score below 70% for the combined average score of the sixteen configurations. Except for a lower score (both absolute and relative) of the 1D_St model and a higher ranking of the 1D/2D_St model , the relative scores of the other models considering all sixteen configurations coincide with the findings based on the default test parameters. The 1D/2D+_Y and 1D/2D+_HY model show a larger range in scores compared to the other models. For the 1D/2D+_Y model, scores range between 52% and 91% for the sixteen test parameter configurations. The two models score relatively high for a depth threshold (h_f) of 10 cm (ranked in the top-3 in each parameter configuration) and low for a depth threshold (h_f) of 5 cm.

Table 9, Table 10 and Figure 24 show the results in the case that only the recognition of flooded grid cells is considered (true positives). The varying of the parameter for areal threshold of non-flooded grid cells ($A_{f,2}$) has no influence on the score. This was to be expected, as non-flooded grid cells were not assessed in this case. The 1D/2D_P scores a perfect 100% in all configurations, except for the four configurations with a 20 m grid size. The 1D/2D+_Y, 1D/2D+_HY and 2D_I model all score above 90% on average. The 1D/2D+_H model scores lowest with an average score of 55% and has the largest range in scores. The ranks of the seven models with a surface model, based on all sixteen parameter configurations, exactly matches the ranks based on the default test parameters. The 1D_St model scores high (93% - 100%) in all configurations. In general, higher scores are reached for a water depth threshold (h_f) of 5 cm than for a water depth threshold of 10 cm, as a lower threshold causes more grid cells being confirmed as flooded.

Table 11, Table 12 and Figure 25 show the results in the case that only the recognition of non-flooded grid cells is considered (true negatives). As expected, the varying of parameter $A_{f,1}$ (areal threshold for flooded grid cells) has no influence on the score. The range in scores for the sixteen configurations is larger compared to the assessment of flooded grid cells. The 1D/2D_St model scores highest (83% average), the 1D/2D+_H model follows closely (77% average). In contrast to the assessment of true positives, the 1D/2D_P model scores lowest (38% average). As already explained in § 6.1.3, such a large difference implies that the 1D/2D_P model overestimates the amount of street flooding. All models, except the 1D/2D_St and 1D/2D+_H model, score higher in the recognition of flooded grid cells than the recognition of non-flooded grid cells.

Another difference compared to the assessment of flooded grid cells (true positives) is that models generally score higher for a water depth threshold of 10 cm. For that reason, the average score of the sixteen configurations is lower than the score for the default test parameters. A higher depth threshold results in more grid cells being marked 'non-flooded' according to modelling results. The chance of a true negative therefore increases. At the same time, more 'non-flooded' grid cells results in less grid cells being marked 'flooded'. The chance of a true negative increases, but the chance of a true positive decreases. As the trends are in contrast, most of the effect is cancelled out when the combined average is assessed. For this reason, the spreading in results in the combined score is smaller than the spreading of results in the partial scores.

7.2.2 Test case Tuindorp

The results for the Tuindorp case are shown in Table 13, Table 14 and Figure 26. In the case of Tuindorp, only flooded grid cells are assessed. The 1D/2D_P model scores highest with an average score of 64%. All models score low, especially for the depth threshold (h_f) of 10 cm. For a depth threshold (h_f) of 10 cm and areal threshold for flooding ($A_{f,1}$) of 5 m^2, none of the models, apart from the 1D/2D_P model, scores higher than 8%. For a depth threshold (h_f) of 5 cm, most models score high. Only the 1D_St and 1D/2D_St model score low for the lower depth threshold. The large difference in scores for different depth thresholds implies that the rain event did not cause much street flooding with a depth of at least 10 cm. However, street flooding with at least 5 cm of water depth is present.

7.2.3 Test case Loenen

The results for the Loenen test case are shown in Table 15, Table 16 and Figure 27. Only one flooded forecourt is assessed and Dutch Design Storm 9 (Bui 09) is used. The 1D_St and 1D/2D_St model do not recognize the location in any of the sixteen configurations. The outcome is not illogical, as Figure 17 already showed that the original 1D_St model did not recognize flooding at the provided location in case of Dutch Design Storm 8 (Bui 08). The models based on a 2D concept recognize the flooded location in each parameter configuration. The flood maps of Loenen provide information on the differences between models, but a valid comparison with observed flooding is not possible based on only one flood prone location.

Ulvenhout – 2016 storm event – Combined Score – Model 1D_St				
Grid (m)	**h_f: 5 cm closest manhole**	**h_f: 10 cm closest manhole**	**h_f: 5 cm surrounding cells**	**h_f: 10 cm surrounding cells**
20	*66%*	*86%*	*62%*	*80%*
30	*69%*	*86%*	*65%*	*85%*
40	*68%*	*84%*	*64%*	*78%*
50	*64%*	*76%*	*62%*	*76%*
Score (average)		**73%**		
Score (median)		**73%**		

Table 7 – Results for test case Ulvenhout and model 1D_St. Combined score of true positives and true negatives. h_f stands for water depth threshold in a grid cell. For explanation on 'closest manhole' and 'surrounding cells', see § 6.1.3.

Configuration				Ulvenhout – 2016 storm event – Combined Score – All except 1D_St													
h_f (cm)	**$A_{f,1}$** (m^2)	**$A_{f,2}$** (m^2)	**Grid** (m)	**1D/2D_St** %	**+ Rank**	**1D/2D_P** %	**+ Rank**	**1D/2D+_H** %	**+ Rank**	**1D/2D+_Y** %	**+ Rank**	**1D/2D+_HY** %	**+ Rank**	**2D_St** %	**+ Rank**	**2D_I** %	**+ Rank**
5	**5**	**5**	**20**	*87%*	*1*	*62%*	*7*	*81%*	*2*	*70%*	*4*	*73%*	*3*	*66%*	*6*	*69%*	*5*
10	**5**	**5**	**20**	*78%*	*5*	*83%*	*4*	*62%*	*7*	*86%*	*2*	*91%*	*1*	*78%*	*6*	*86%*	*3*
10	**5**	**1**	**20**	*77%*	*5*	*78%*	*4*	*61%*	*7*	*85%*	*2*	*86%*	*1*	*68%*	*6*	*79%*	*3*
10	**1**	**1**	**20**	*82%*	*3*	*78%*	*5*	*66%*	*7*	*91%*	*1*	*88%*	*2*	*73%*	*6*	*80%*	*4*
5	**5**	**5**	**30**	*85%*	*1*	*62%*	*4*	*73%*	*2*	*59%*	*6*	*63%*	*3*	*61%*	*5*	*58%*	*7*
10	**5**	**5**	**30**	*85%*	*3*	*79%*	*6*	*68%*	*7*	*87%*	*2*	*89%*	*1*	*80%*	*5*	*80%*	*5*
10	**5**	**1**	**30**	*85%*	*1*	*68%*	*6*	*66%*	*7*	*83%*	*2*	*79%*	*3*	*68%*	*6*	*73%*	*4*
10	**1**	**1**	**30**	*89%*	*1*	*68%*	*7*	*70%*	*6*	*85%*	*2*	*82%*	*3*	*72%*	*5*	*73%*	*4*
5	**5**	**5**	**40**	*82%*	*1*	*58%*	*3*	*60%*	*2*	*52%*	*7*	*56%*	*4*	*54%*	*5*	*52%*	*7*
10	**5**	**5**	**40**	*82%*	*3*	*74%*	*6*	*66%*	*7*	*88%*	*2*	*92%*	*1*	*78%*	*4*	*76%*	*5*
10	**5**	**1**	**40**	*80%*	*2*	*68%*	*5*	*64%*	*6*	*80%*	*2*	*78%*	*3*	*58%*	*7*	*68%*	*5*
10	**1**	**1**	**40**	*82%*	*2*	*68%*	*6*	*72%*	*4*	*82%*	*2*	*78%*	*3*	*66%*	*7*	*68%*	*6*
5	**5**	**5**	**50**	*76%*	*1*	*57%*	*2*	*55%*	*5*	*55%*	*5*	*55%*	*5*	*52%*	*7*	*52%*	*7*
10	**5**	**5**	**50**	*79%*	*3*	*71%*	*5*	*64%*	*7*	*81%*	*2*	*81%*	*2*	*69%*	*6*	*71%*	*5*
10	**5**	**1**	**50**	*76%*	*1*	*62%*	*5*	*62%*	*5*	*74%*	*3*	*74%*	*3*	*57%*	*7*	*57%*	*7*
10	**1**	**1**	**50**	*79%*	*1*	*62%*	*6*	*62%*	*6*	*74%*	*3*	*74%*	*3*	*62%*	*6*	*57%*	*7*
Score (average)				**81% (1)**		**69% (5)**		**66% (7)**		**77% (3)**		**77% (2)**		**66% (6)**		**69% (4)**	
Score (median)				**82% (1)**		**68% (5)**		**65% (7)**		**82% (2)**		**78% (3)**		**67% (6)**		**70% (4)**	

Table 8 – Results for test case Ulvenhout of seven models (all except 1D_St). Combined score of true positives and true negatives. h_f stands for water depth threshold in a grid cell. A_f stands for the areal threshold for flooding in a grid cell. The subscripts '1' and '2' divide between the value for flooded and non-flooded locations respectively. The rank function picks the lower rank when two models score the same.

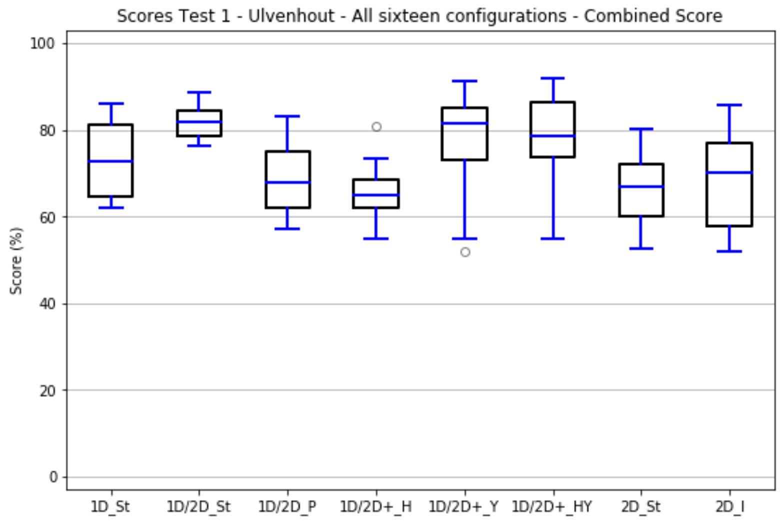


Figure 23 – Results of test 1 for test case Ulvenhout, composed out of the average for both true negatives and true positives

Ulvenhout – 2016 storm event – True Positives – Model 1D_St				
Grid (m)	h_f: 5 cm closest manhole	h_f: 10 cm closest manhole	h_f: 5 cm surrounding cells	h_f: 10 cm surrounding cells
20	*93%*	*93%*	*97%*	*97%*
30	*97%*	*94%*	*100%*	*97%*
40	*96%*	*96%*	*96%*	*96%*
50	*100%*	*95%*	*100%*	*100%*
Score (average)		97%		
Score (median)		96%		

Table 9 – Results for test case Ulvenhout of models 1D_St. True positive scroes. h_f stands for water depth threshold in a grid cell. For explanation on 'closest manhole' and 'surrounding cells', see § 6.1.3.

Configuration				Ulvenhout – 2016 storm event – True Positives – All except 1D_St													
h_f (cm)	$A_{f,1}$ (m²)	$A_{f,2}$ (m²)	Grid (m)	1D/2D_St % + Rank		1D/2D_P % + Rank		1D/2D+_H % + Rank		1D/2D+_Y % + Rank		1D/2D+_HY % + Rank		2D_St % + Rank		2D_I % + Rank	
5	**5**	**5**	**20**	*92%*	*7*	*97%*	*6*	*97%*	*6*	*98%*	*1*	*98%*	*1*	*95%*	*1*	*98%*	*1*
10	**5**	**5**	**20**	*59%*	*5*	*97%*	*1*	*26%*	*7*	*79%*	*4*	*89%*	*3*	*57%*	*6*	*92%*	*2*
10	**5**	**1**	**20**	*59%*	*5*	*97%*	*1*	*26%*	*7*	*79%*	*4*	*89%*	*3*	*57%*	*6*	*92%*	*2*
10	**1**	**1**	**20**	*69%*	*5*	*97%*	*1*	*36%*	*7*	*92%*	*4*	*93%*	*3*	*67%*	*6*	*93%*	*3*
5	**5**	**5**	**30**	*97%*	*7*	*100%*	*5*	*100%*	*5*	*100%*	*5*	*100%*	*5*	*97%*	*7*	*100%*	*5*
10	**5**	**5**	**30**	*75%*	*5*	*100%*	*1*	*39%*	*7*	*86%*	*4*	*89%*	*3*	*69%*	*6*	*97%*	*2*
10	**5**	**1**	**30**	*75%*	*5*	*100%*	*1*	*39%*	*7*	*86%*	*4*	*89%*	*3*	*69%*	*6*	*97%*	*2*
10	**1**	**1**	**30**	*83%*	*5*	*100%*	*1*	*47%*	*7*	*89%*	*4*	*94%*	*3*	*78%*	*6*	*97%*	*2*
5	**5**	**5**	**40**	*100%*	*6*	*100%*	*6*	*100%*	*6*	*100%*	*6*	*100%*	*6*	*96%*	*7*	*100%*	*6*
10	**5**	**5**	**40**	*72%*	*5*	*100%*	*1*	*40%*	*7*	*92%*	*4*	*96%*	*3*	*64%*	*6*	*96%*	*3*
10	**5**	**1**	**40**	*72%*	*5*	*100%*	*1*	*40%*	*7*	*92%*	*4*	*96%*	*3*	*64%*	*6*	*96%*	*3*
10	**1**	**1**	**40**	*76%*	*6*	*100%*	*1*	*56%*	*7*	*96%*	*4*	*96%*	*4*	*80%*	*5*	*96%*	*4*
5	**5**	**5**	**50**	*100%*	*7*	*100%*	*7*	*100%*	*7*	*100%*	*7*	*100%*	*7*	*100%*	*7*	*100%*	*7*
10	**5**	**5**	**50**	*81%*	*5*	*100%*	*1*	*43%*	*7*	*90%*	*4*	*90%*	*4*	*71%*	*6*	*95%*	*2*
10	**5**	**1**	**50**	*81%*	*5*	*100%*	*1*	*43%*	*7*	*90%*	*4*	*90%*	*4*	*71%*	*6*	*95%*	*2*
10	**1**	**1**	**50**	*86%*	*5*	*100%*	*1*	*43%*	*7*	*90%*	*4*	*90%*	*4*	*81%*	*6*	*95%*	*2*
Score (average)				80% (5)		99% (1)		55% (7)		91% (4)		94% (3)		76% (6)		96% (2)	
Score (median)				79% (5)		100% (1)		43% (7)		91% (4)		94% (3)		71% (6)		96% (2)	

Table 10 – Results for test case Ulvenhout of 7 models (all except 1D_St). True positive scores. h_f stands for water depth threshold in a grid cell. A_f stands for the areal threshold for flooding in a grid cell. The subscripts '1' and '2' divide between the value for flooded and non-flooded locations respectively. The rank function picks the lower rank when two models score the same.

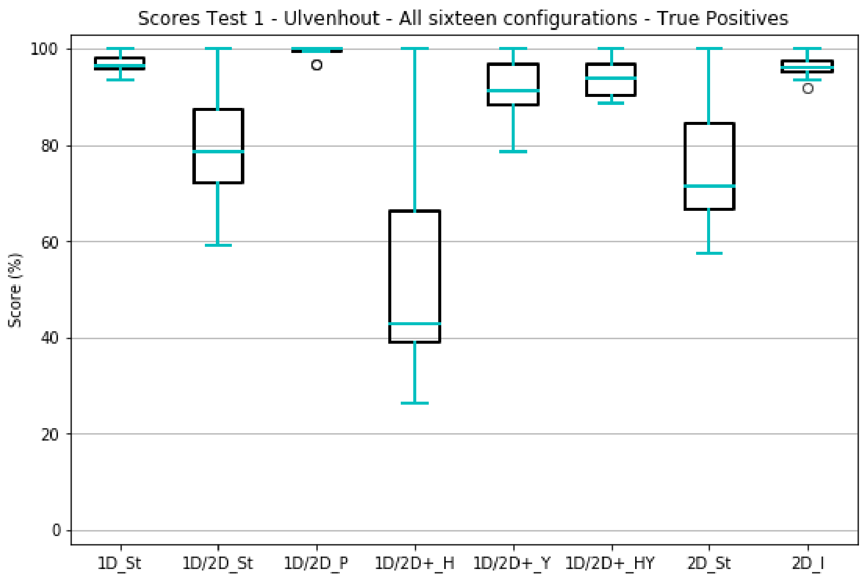


Figure 24 – Results of test 1 for test case Ulvenhout. The score based only on true positives is shown

Ulvenhout – 2016 storm event – True Negatives – Model 1D_St				
Grid (m)	h_f: 5 cm closest manhole	h_f: 10 cm closest manhole	h_f: 5 cm surrounding cells	h_f: 10 cm surrounding cells
20	*41%*	*78%*	*30%*	*64%*
30	*40%*	*77%*	*29%*	*71%*
40	*40%*	*72%*	*32%*	*60%*
50	*29%*	*57%*	*24%*	*52%*
Score (average)		**50%**		
Score (median)		**47%**		

Table 11 – Results for test case Ulvenhout of models 1D_St. True negative scroes. h_f stands for water depth threshold in a grid cell. For explanation on 'closest manhole' and 'surrounding cells', see § 6.1.3.

Configuration				Ulvenhout – 2016 storm event – True Negatives – All except 1D_St													
h_f (cm)	$A_{f,1}$ (m²)	$A_{f,2}$ (m²)	Grid (m)	1D/2D_St %	Rank	1D/2D_P %	Rank	1D/2D+_H %	Rank	1D/2D+_Y %	Rank	1D/2D+_HY %	Rank	2D_St %	Rank	2D_I %	Rank
5	**5**	**5**	**20**	*83%*	*1*	*28%*	*7*	*66%*	*2*	*44%*	*4*	*48%*	*3*	*38%*	*6*	*41%*	*5*
10	**5**	**5**	**20**	*97%*	*3*	*70%*	*7*	*97%*	*3*	*94%*	*5*	*94%*	*5*	*97%*	*3*	*80%*	*6*
10	**5**	**1**	**20**	*94%*	*2*	*59%*	*7*	*94%*	*2*	*91%*	*3*	*83%*	*4*	*78%*	*5*	*67%*	*6*
10	**1**	**1**	**20**	*94%*	*2*	*59%*	*7*	*94%*	*2*	*91%*	*3*	*83%*	*4*	*78%*	*5*	*67%*	*6*
5	**5**	**5**	**30**	*71%*	*2*	*23%*	*5*	*46%*	*3*	*17%*	*6*	*26%*	*4*	*23%*	*5*	*14%*	*7*
10	**5**	**5**	**30**	*94%*	*2*	*57%*	*7*	*97%*	*1*	*89%*	*5*	*89%*	*5*	*91%*	*3*	*63%*	*6*
10	**5**	**1**	**30**	*94%*	*2*	*34%*	*7*	*94%*	*2*	*80%*	*3*	*69%*	*4*	*66%*	*5*	*49%*	*6*
10	**1**	**1**	**30**	*94%*	*2*	*34%*	*7*	*94%*	*2*	*80%*	*3*	*69%*	*4*	*66%*	*5*	*49%*	*6*
5	**5**	**5**	**40**	*64%*	*1*	*16%*	*3*	*20%*	*2*	*4%*	*7*	*12%*	*5*	*12%*	*5*	*4%*	*7*
10	**5**	**5**	**40**	*92%*	*3*	*48%*	*7*	*92%*	*3*	*84%*	*5*	*88%*	*4*	*92%*	*3*	*56%*	*6*
10	**5**	**1**	**40**	*88%*	*2*	*36%*	*7*	*88%*	*2*	*68%*	*3*	*60%*	*4*	*52%*	*5*	*40%*	*6*
10	**1**	**1**	**40**	*88%*	*2*	*36%*	*7*	*88%*	*2*	*68%*	*3*	*60%*	*4*	*52%*	*5*	*40%*	*6*
5	**5**	**5**	**50**	*52%*	*1*	*14%*	*2*	*10%*	*5*	*10%*	*5*	*10%*	*5*	*5%*	*7*	*5%*	*7*
10	**5**	**5**	**50**	*76%*	*2*	*43%*	*7*	*86%*	*1*	*71%*	*4*	*71%*	*4*	*67%*	*5*	*48%*	*6*
10	**5**	**1**	**50**	*71%*	*2*	*24%*	*6*	*81%*	*1*	*57%*	*4*	*57%*	*4*	*43%*	*5*	*19%*	*7*
10	**1**	**1**	**50**	*71%*	*2*	*24%*	*6*	*81%*	*1*	*57%*	*4*	*57%*	*4*	*43%*	*5*	*19%*	*7*
Score (average)				**83% (1)**		**38% (6)**		**77% (2)**		**63% (3)**		**61% (4)**		**56% (5)**		**41% (7)**	
Score (median)				**88% (2)**		**35% (6)**		**88 % (2)**		**70% (3)**		**64% (4)**		**59% (5)**		**44% (7)**	

Table 12 – Results for test case Ulvenhout of 7 models (all except 1D_St). True negative scores. H_f stands for water depth threshold in a grid cell. A_f stands for the areal threshold for flooding in a grid cell. The subscripts '1' and '2' divide between the value for flooded and non-flooded locations respectively. The rank function picks the lower rank when two models score the same.

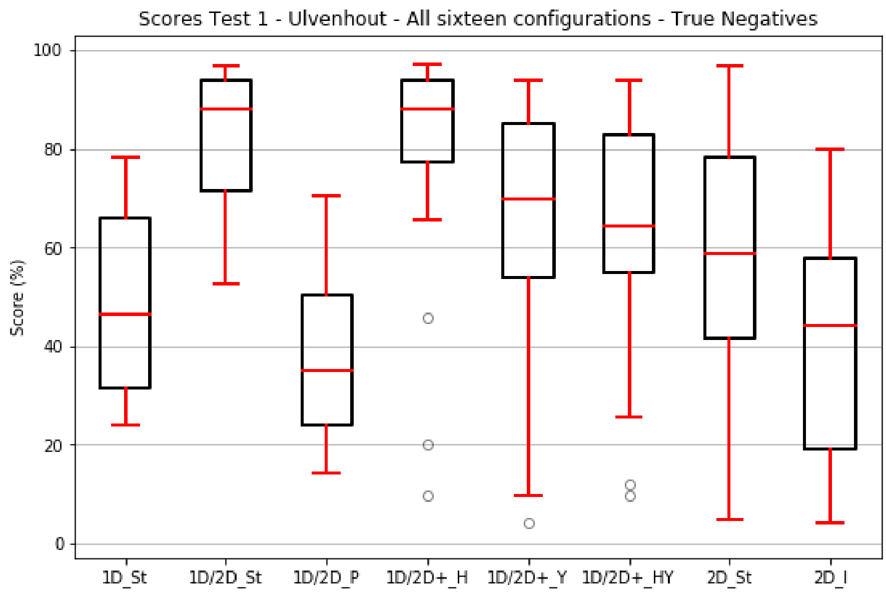


Figure 25 – Results of test 1 for test case Ulvenhout. The score based only on true negatives is shown

Tuindorp – 2013 storm event – True Positives – Model 1D_St				
Grid (m)	h_f: 5 cm closest manhole	h_f: 10 cm closest manhole	h_f: 5 cm surrounding cells	h_f: 10 cm surrounding cells
20	*3%*	*0%*	*3%*	*0%*
30	*4%*	*0%*	*4%*	*0%*
40	*0%*	*0%*	*0%*	*0%*
50	*0%*	*0%*	*0%*	*0%*
Score (average)		**1%**		
Score (median)		**0%**		

Table 13 – Results for test case Tuindorp of models 1D_St. True positive scores. h_f stands for water depth threshold in a grid cell. For explanation on 'closest manhole' and 'surrounding cells', see § 6.1.3.

Configuration				Tuindorp – 2013 storm event – True Positives – All except 1D_St													
h_f (cm)	$A_{f,1}$ (m²)	$A_{f,2}$ (m²)	Grid (m)	1D/2D_St %	Rank	1D/2D_P %	Rank	1D/2D+_H %	Rank	1D/2D+_Y %	Rank	1D/2D+_HY %	Rank	2D_St %	Rank	2D_I %	Rank
5	**5**	**5**	**20**	*9%*	*7*	*86%*	*2*	*57%*	*6*	*77%*	*3*	*69%*	*4*	*60%*	*5*	*94%*	*1*
10	**5**	**5**	**20**	*0%*	*7*	*37%*	*1*	*0%*	*7*	*3%*	*3*	*6%*	*2*	*0%*	*7*	*0%*	*7*
10	**5**	**1**	**20**	*0%*	*7*	*37%*	*1*	*0%*	*7*	*3%*	*3*	*6%*	*2*	*0%*	*7*	*0%*	*7*
10	**1**	**1**	**20**	*0%*	*7*	*54%*	*1*	*3%*	*6*	*23%*	*3*	*14%*	*4*	*3%*	*6*	*34%*	*2*
5	**5**	**5**	**30**	*8%*	*7*	*96%*	*2*	*71%*	*6*	*83%*	*3*	*75%*	*5*	*75%*	*5*	*96%*	*2*
10	**5**	**5**	**30**	*0%*	*7*	*46%*	*1*	*0%*	*7*	*4%*	*3*	*8%*	*2*	*0%*	*7*	*0%*	*7*
10	**5**	**1**	**30**	*0%*	*7*	*46%*	*1*	*0%*	*7*	*4%*	*3*	*8%*	*2*	*0%*	*7*	*0%*	*7*
10	**1**	**1**	**30**	*0%*	*7*	*63%*	*1*	*4%*	*6*	*25%*	*3*	*17%*	*4*	*4%*	*6*	*33%*	*2*
5	**5**	**5**	**40**	*5%*	*7*	*100%*	*2*	*86%*	*6*	*86%*	*6*	*86%*	*6*	*90%*	*3*	*100%*	*2*
10	**5**	**5**	**40**	*0%*	*7*	*52%*	*1*	*0%*	*7*	*5%*	*3*	*5%*	*3*	*0%*	*7*	*0%*	*7*
10	**5**	**1**	**40**	*0%*	*7*	*52%*	*1*	*0%*	*7*	*5%*	*3*	*5%*	*3*	*0%*	*7*	*0%*	*7*
10	**1**	**1**	**40**	*0%*	*7*	*71%*	*1*	*5%*	*6*	*33%*	*3*	*10%*	*5*	*10%*	*5*	*43%*	*2*
5	**5**	**5**	**50**	*8%*	*7*	*100%*	*3*	*92%*	*6*	*92%*	*6*	*92%*	*6*	*100%*	*3*	*100%*	*3*
10	**5**	**5**	**50**	*0%*	*7*	*54%*	*1*	*0%*	*7*	*8%*	*3*	*8%*	*3*	*0%*	*7*	*0%*	*7*
10	**5**	**1**	**50**	*0%*	*7*	*54%*	*1*	*0%*	*7*	*8%*	*3*	*8%*	*3*	*0%*	*7*	*0%*	*7*
10	**1**	**1**	**50**	*0%*	*7*	*77%*	*1*	*8%*	*6*	*54%*	*3*	*23%*	*4*	*15%*	*5*	*54%*	*3*
Score (average)				**2% (7)**		**64% (1)**		**20% (6)**		**32% (3)**		**27% (4)**		**22% (5)**		**35% (2)**	
Score (median)				**0% (7)**		**54% (1)**		**1% (6)**		**15% (3)**		**9% (4)**		**1% (6)**		**17% (2)**	

Table 14 – Results for test case Tuindorp of 7 models (all except 1D_St). True positive scores. h_f stands for water depth threshold in a grid cell. A_f stands for the areal threshold for flooding in a grid cell. The subscripts '1' and '2' divide between the value for flooded and non-flooded locations respectively. The rank function picks the lower rank when two models score the same.

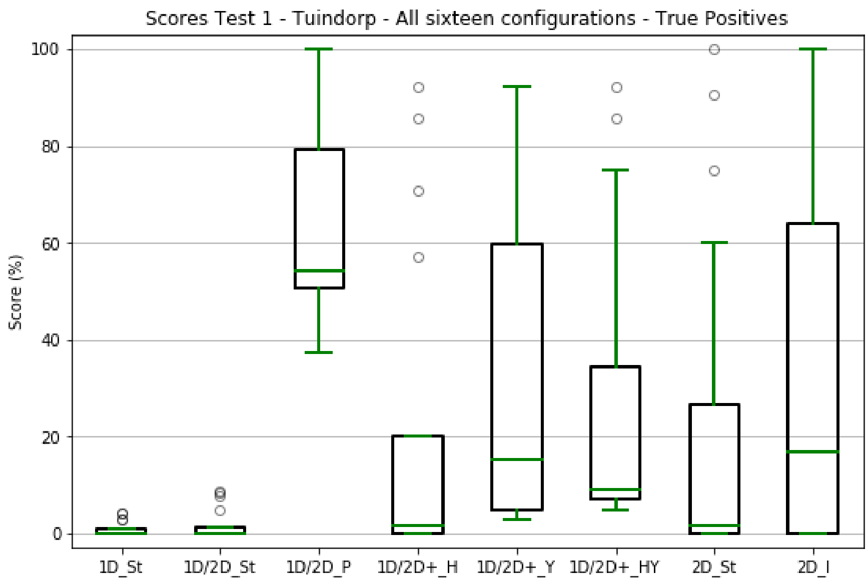


Figure 26- Results of test 1 for test case Tuindorp

Loenen – no specific storm event (Bui 09) – True Positives – Model 1D_St				
Grid (m)	**h_f: 5 cm closest manhole**	**h_f: 10 cm closest manhole**	**h_f: 5 cm surrounding cells**	**h_f: 10 cm surrounding cells**
20	*0%*	*0%*	*0%*	*0%*
30	*0%*	*0%*	*0%*	*0%*
40	*0%*	*0%*	*0%*	*0%*
50	*0%*	*0%*	*0%*	*0%*
Score (average)		**0%**		
Score (median)		**0%**		

Table 15 – Results for test case Loenen of models 1D_St. True positive scores. h_f stands for water depth threshold in a grid cell. For explanation on 'closest manhole' and 'surrounding cells', see § 6.1.3.

Configuration				Loenen – no specific storm event (Bui 09) – True Positives													
h_f (cm)	**$A_{f,1}$** (m²)	**$A_{f,2}$** (m²)	**Grid** (m)	**1D/2D_St % + Rank**		**1D/2D_P % + Rank**		**1D/2D+_H % + Rank**		**1D/2D+_Y % + Rank**		**1D/2D+HY % + Rank**		**2D_St % + Rank**		**2D_I % + Rank**	
5	**5**	**5**	**20**	*0%*	*7*	*100%*	*6*	*100%*	*6*	*100%*	*6*	*100%*	*6*	*100%*	*6*	*100%*	*6*
10	**5**	**5**	**20**	*0%*	*7*	*0%*	*7*	*0%*	*7*	*33%*	*3*	*0%*	*7*	*100%*	*2*	*100%*	*2*
10	**5**	**1**	**20**	*0%*	*7*	*0%*	*7*	*0%*	*7*	*33%*	*3*	*0%*	*7*	*100%*	*2*	*100%*	*2*
10	**1**	**1**	**20**	*0%*	*7*	*33%*	*3*	*0%*	*7*	*33%*	*3*	*33%*	*3*	*100%*	*1*	*100%*	*1*
5	**5**	**5**	**30**	*0%*	*7*	*100%*	*6*	*100%*	*6*	*100%*	*6*	*100%*	*6*	*100%*	*6*	*100%*	*6*
10	**5**	**5**	**30**	*0%*	*7*	*0%*	*7*	*0%*	*7*	*100%*	*4*	*100%*	*4*	*100%*	*4*	*100%*	*4*
10	**5**	**1**	**30**	*0%*	*7*	*0%*	*7*	*0%*	*7*	*100%*	*4*	*100%*	*4*	*100%*	*4*	*100%*	*4*
10	**1**	**1**	**30**	*0%*	*7*	*100%*	*5*	*0%*	*7*	*100%*	*5*	*100%*	*5*	*100%*	*5*	*100%*	*5*
5	**5**	**5**	**40**	*0%*	*7*	*100%*	*6*	*100%*	*6*	*100%*	*6*	*100%*	*6*	*100%*	*6*	*100%*	*6*
10	**5**	**5**	**40**	*0%*	*7*	*50%*	*4*	*0%*	*7*	*50%*	*4*	*0%*	*7*	*100%*	*2*	*100%*	*2*
10	**5**	**1**	**40**	*0%*	*7*	*50%*	*4*	*0%*	*7*	*50%*	*4*	*0%*	*7*	*100%*	*2*	*100%*	*2*
10	**1**	**1**	**40**	*0%*	*7*	*50%*	*5*	*0%*	*7*	*50%*	*5*	*50%*	*5*	*100%*	*2*	*100%*	*2*
5	**5**	**5**	**50**	*0%*	*7*	*100%*	*6*	*100%*	*6*	*100%*	*6*	*100%*	*6*	*100%*	*6*	*100%*	*6*
10	**5**	**5**	**50**	*0%*	*7*	*100%*	*5*	*0%*	*7*	*100%*	*5*	*100%*	*5*	*100%*	*5*	*100%*	*5*
10	**5**	**1**	**50**	*0%*	*7*	*100%*	*5*	*0%*	*7*	*100%*	*5*	*100%*	*5*	*100%*	*5*	*100%*	*5*
10	**1**	**1**	**50**	*0%*	*7*	*100%*	*5*	*0%*	*7*	*100%*	*5*	*100%*	*5*	*100%*	*5*	*100%*	*5*
Score (average)				**0% (7)**		**61% (5)**		**25% (6)**		**84% (3)**		**68% (4)**		**100% (2)**		**100% (2)**	
Score (median)				**0% (7)**		**100% (5)**		**0% (7)**		**100% (5)**		**100% (5)**		**100% (5)**		**100% (5)**	

Table 16– Results for test case Loenen of 7 models (all except 1D_St). True positive scores. H_f stands for water depth threshold in a grid cell. A_f stands for the areal threshold for flooding in a grid cell. The subscripts '1' and '2' divide between the value for flooded and non-flooded locations respectively. The rank function picks the lower rank when two models score the same.

Figure 27 – Results of test 1 for test case Loenen

7.3 Test 2 – Test Results

Test 2 is a method to assess the in-sewer performance by comparing the modelling results in a certain manhole with in-sewer monitoring data. Three indicators were introduced: the RMSE, NSE and KGE. In the case of Tuindorp and Ulvenhout, two storms events were used. In the case of Tuindorp, monitoring results from fourteen manholes was used for both storm events. In the case of Ulvenhout, monitoring data from one manhole was used. For Loenen, no monitoring data was available. Only the models based on a 1D, 1D/2D and 1D/2D+ concept are tested, as the 2D_St and 2D_I model not contain a hydrodynamic sewer model. The four phases in the calculation of RMSE represent the dry weather flow (P1), filling up (P2), peak water level (P3) and emptying (P4) of the system.

7.3.1 Test case Ulvenhout

In the case of Ulvenhout, the monitored manhole is a pump sump of which the on and off level were not implemented correctly in the model. For this reason, the run time before the storm event was shortened and P1 is not considered. As the amount of rainfall (53 mm for both storms) is significantly higher than the amount of waste water usually present in a sewer system (a small portion of the total storage, which is around 7 mm in a combined sewer system), the influence of this on the peak water level is expected to be small.

The results for test case Ulvenhout are presented in Table 17 and Table 18. The graphs with water levels in the corresponding manhole are shown in Figure 28 and Figure 29. Information on the location of the manhole is provided in Appendix L. For both storm events, the RMSE is highest during the filling up of the system (P2) and lowest during peak water level (P3). In the filling up of the system, a small time shift between model and observation can result in a large difference in water depth for a certain moment in time. It is therefore likely that the largest RMSE is found in this phase.

Considering both storm events, there is no model (or modelling concept) that clearly performs best according to the indicators. For the 2016 storm event, the 1D/2D_St and 1D/2D_P model score highest. The 1D/2D+_H and 1D/2D+_HY model score highest for the 2017 storm event. The 1D/2D_P model scores best for the 2016 storm event, but worst for the 2017 storm event. The 1D/2D+_Y model scores low for both storm events. The RMSE for the 2016 storm event, 36.25 cm, is 10 cm higher than the second worst scoring model (1D/2D+_H with 26.34 cm). For that same storm event, the NSE is negative (-0.048). This means that the mean of the observed values (a single, constant number) is a better estimator of the water level than the model. Both storm events did not result in overflowing of the monitored manhole (ground level +4.33 m). As there is a difference in the results of the 1D_St and 1D/2D_St model for both test cases, it indicates that street flooding did occur in other parts of Ulvenhout during both storm events.

The beta-term (β) of the KGE is composed of the ratio between the mean of the observed and the mean of the monitored data. It can be used as a rough indicator for overestimation or underestimation of sewer inflow. For the 2016 storm event, all models score < 1. This means that the average of all models is below the average of the monitored data, indicating underestimation of sewer inflow. Figure 28 confirms the numbers. The 1D/2D+_H and 1D/2D+_Y model have the lowest β (0.954 and 0.943). For the 2017 storm event, overall β-values are closer to zero. Again, the 1D/2D+_H and 1D/2D+_Y model score < 1 (0.995 and 0.950), indicating underestimation of sewer inflow.

The 2017 storm event runs over a period of eleven hours. At the beginning of the storm event, a minor storm event passes by (6 mm in two hours). All models show an increase in the water level for this smaller storm event (see Figure 29). However, the monitoring shows no increase at all in the water level in the sewer system. The difference can be caused by a difference in representation of the surface runoff for a smaller storm event. However, there is also a suspected difference between the models and measurements in the form of an extra conduit (200 mm), which is depicted in Appendix L. Besides the reaction to the initial storm event, the resulting water level of the 1D/2D+_Y model, which is least influenced by the minor storm event, shows that the pump on and off level is not represented correctly in the provided 1D (1D_St) model.

7.3.2 Test case Tuindorp

For Tuindorp, the originally provided 1D_St model shows a similar deviation from the measurements in a large amount of manholes. The initial water level is lower and more constant in the model, while the peak (especially the first) is overestimated. According to communication with the person who calibrated the model, calibration of the hydrological model faced difficulties as the measuring equipment could not be hung deep enough in the manhole (M. van Bijnen 2018, personal communication, 9 April). The calibration of the model therefore focussed on the 'usable' phases of the rain events, after the initial filling up of the system (M. van Bijnen 2018, personal communication, 9 May). Furthermore, the parameter set for runoff parameters is retrieved from the *Leidraad Riolering*. The difficulties in calibration and the default parameter set influenced the performance of the model at the start of a storm event. Appendix M shows some examples of the described deviation.

The results for test case Tuindorp are presented in Table 19 and Table 20. Figure 30 and Figure 31 show the water levels in one of the monitored manholes. Information on the location of the manhole is provided in Appendix L. For both storm events, the highest RMSE's are found in P2, just as in test case Ulvenhout. The results for test case Tuindorp are more uniform over the two storms events than in the Ulvenhout test case. This was to be expected, as more manholes were taken into account.

For both storm events, the 1D/2D+_H model scores best on every indicator. The RMSE is at least a centimer lower than the second best scoring model, of which the ID differs per test case. The KGE and NSE are closest to unity. The 1D/2D+_HY model scores somewhat worse (RMSE +1.2 cm and +2.4 cm), despite the added subdivision of yards. The models based on a 1D/2D+ concept in general score lower (= better) in the RMSE of P2, which indicates that the modelling of fully distributed runoff improves the modelling of sewer inflow. The 1D/2D_P model scores worst on every indicator and for both storm events. Especially the RMSE in period 3 (peak water levels) is large: two to three times higher than the other models. The β-term of the 1D/2D_P model is positive (1.107 for 2014, 1.235 for 2015 storm event), conforming overestimation of sewer inflow. Just as in test case Ulvenhout, the 1D/2D+_Y model underestimates sewer inflow. The β-term of the KGE is negative for both storm events (0.762 and 0.841). Figure 30 and Figure 31 confirm this.

In comparison to the Ulvenhout case, the average RMSE over the fourteen manholes is significantly lower. A possible explanation for this is that the provided 1D model of Tuindorp is calibrated more recently and that the sewer system of Tuindorp was recently cleaned at the time of the measurements. However, the other indicators (NSE and KGE) are higher for the Ulvenhout case. The KGE values for the Tuindorp case are strongly dominated by the α-term (variability ratio). As the variability ratio is > 1, it implicates that the standard deviation of modelled water levels is higher than the measured water levels. This was to be expected, as it was stated that the model tends to underestimate initial water level and overestimate peak water levels. The NSE is influenced by the mean of the observations. If the mean value is generally close to the observations (for instance in a storm event with a sharp peak and prolonged peak water level), the NSE will be lower as the mean of the observations is a relatively stronger estimator of water level.

Ulvenhout – 2016 Storm Event – one manhole										
Model ID	**RMSE P1** (cm)	**RMSE P2** (cm)	**RMSE P3** (cm)	**RMSE P4** (cm)	**RMSE total** (cm)	**KGE r** (-)	**KGE beta** (-)	**KGE alpha** (-)	**KGE total** (-)	**NSE total** (-)
1D_St	*N/A*	*35.25*	*9.88*	*17.06*	**24.24**	*0.933*	*0.960*	*1.211*	**0.775**	**0.531**
1D/2D_St	*N/A*	*30.58*	*12.92*	*19.11*	**22.97**	*0.939*	*0.959*	*1.113*	**0.865**	**0.579**
1D/2D_P	*N/A*	*29.96*	*5.82*	*17.59*	**21.55**	*0.931*	*0.968*	*1.193*	**0.792**	**0.630**
1D/2D+_H	*N/A*	*37.27*	*13.36*	*19.44*	**26.34**	*0.930*	*0.954*	*1.206*	**0.778**	**0.447**
1D/2D+_Y	*N/A*	*57.46*	*9.83*	*17.61*	**36.25**	*0.941*	*0.943*	*1.585*	**0.410**	**-0.048**
1D/2D+_HY	*N/A*	*35.25*	*9.88*	*17.06*	**24.24**	*0.933*	*0.960*	*1.211*	**0.775**	**0.531**

Table 17 – Results for test case Ulvenhout, 2016 storm event

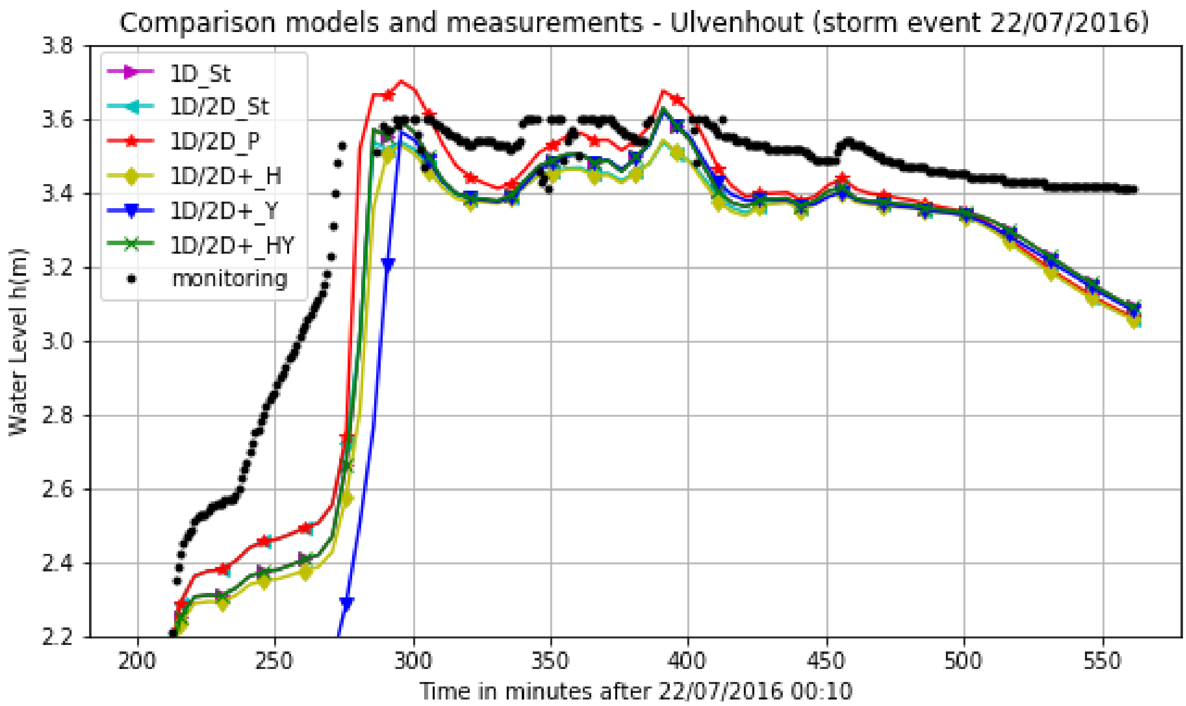


Figure 28 – Monitoring and modelling results for the 2016 storm event in Ulvenhout

Ulvenhout – 2017 Storm Event – one manhole										
Model ID	**RMSE P1** (cm)	**RMSE P2** (cm)	**RMSE P3** (cm)	**RMSE P4** (cm)	**RMSE total** (cm)	**KGE r** (-)	**KGE beta** (-)	**KGE alpha** (-)	**KGE total** (-)	**NSE total** (-)
1D_St	*N/A*	*49.47*	*12.86*	*17.94*	**31.84**	*0.878*	*1.012*	*0.932*	**0.860**	**0.956**
1D/2D_St	*N/A*	*50.40*	*13.24*	*17.97*	**32.41**	*0.874*	*1.012*	*0.939*	**0.859**	**0.954**
1D/2D_P	*N/A*	*50.40*	*16.24*	*17.68*	**32.83**	*0.874*	*1.016*	*0.961*	**0.867**	**0.953**
1D/2D+_H	*N/A*	*44.48*	*8.08*	*17.98*	**28.56**	*0.906*	*0.995*	*1.012*	**0.905**	**0.965**
1D/2D+_Y	*N/A*	*48.39*	*18.25*	*16.21*	**31.89**	*0.958*	*0.950*	*1.235*	**0.756**	**0.956**
1D/2D+_HY	*N/A*	*43.68*	*11.75*	*17.29*	**28.43**	*0.904*	*1.006*	*0.974*	**0.900**	**0.965**

Table 18 – Results for test case Ulvenhout, 2017 storm event

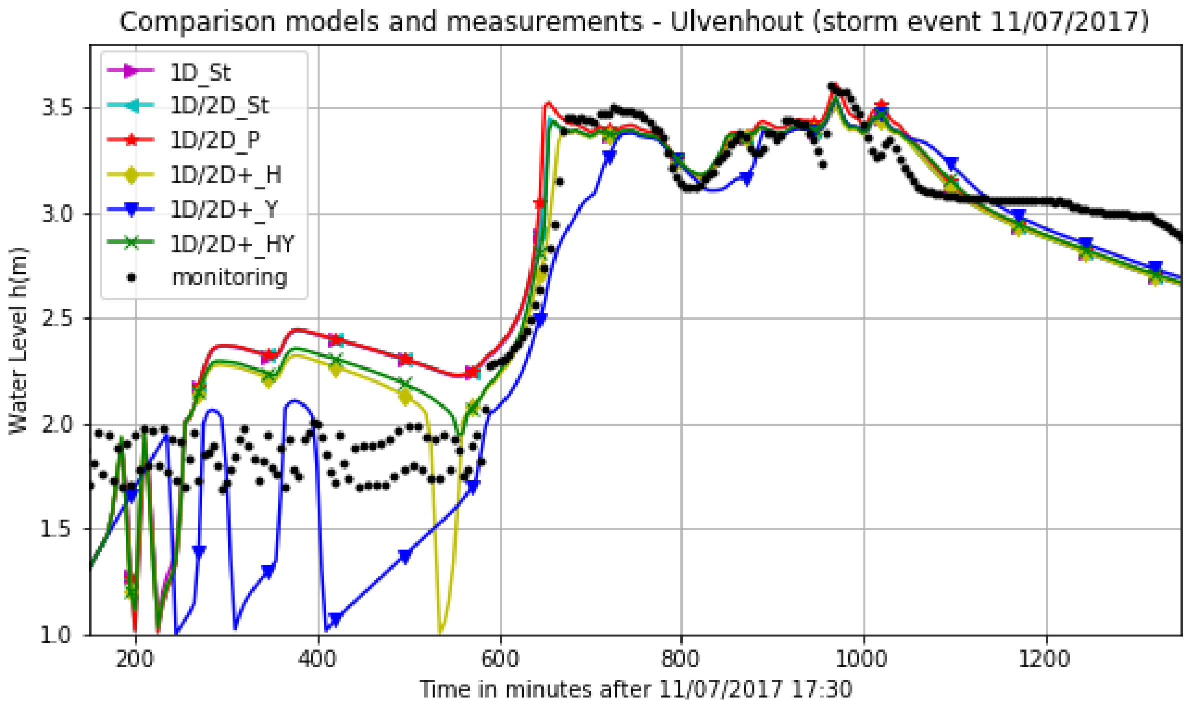


Figure 29 – Monitoring and modelling results for the 2017 storm event in Ulvenhout

Tuindorp – 2014 Storm Event – fourteen manholes										
Model ID	**RMSE P1** (cm)	**RMSE P2** (cm)	**RMSE P3** (cm)	**RMSE P4** (cm)	**RMSE total** (cm)	**KGE r** (-)	**KGE beta** (-)	**KGE alpha** (-)	**KGE total** (-)	**NSE total** (-)
1D_St	*12.83*	*32.97*	*16.47*	*9.85*	**16.12**	*0.942*	*0.942*	*1.391*	**0.541**	**0.409**
1D/2D_St	*12.84*	*33.38*	*16.11*	*9.84*	**16.11**	*0.942*	*0.943*	*1.391*	**0.541**	**0.406**
1D/2D_P	*12.84*	*58.50*	*32.19*	*9.76*	**25.05**	*0.920*	*1.107*	*1.767*	**0.170**	**-0.441**
1D/2D+_H	*12.91*	*25.52*	*14.59*	*11.76*	**15.11**	*0.956*	*0.885*	*1.313*	**0.595**	**0.506**
1D/2D+_Y	*13.11*	*36.28*	*14.32*	*18.91*	**18.70**	*0.933*	*0.762*	*1.172*	**0.601**	**0.437**
1D/2D+_HY	*12.91*	*27.12*	*16.61*	*12.82*	**16.30**	*0.949*	*0.904*	*1.362*	**0.554**	**0.448**

Table 19 – Results for test case Tuindorp, 2014 storm event

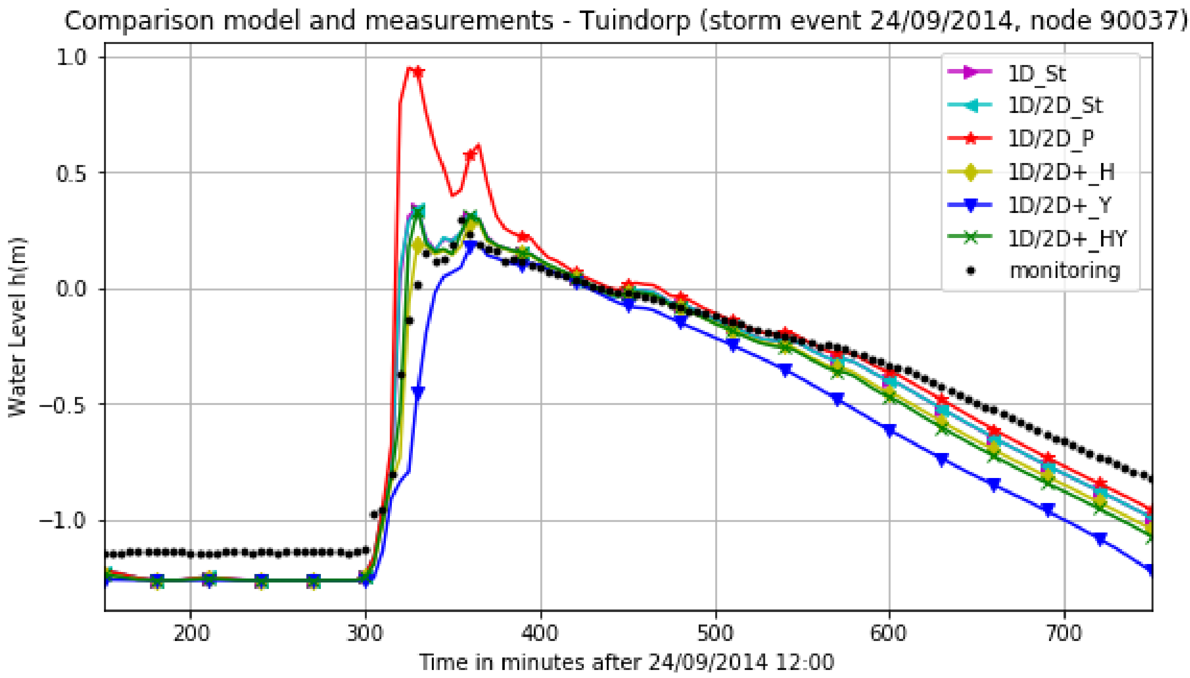


Figure 30 – Monitoring and modelling results for the 2014 storm event in Tuindorp. Manhole/node 90037 is selected.

Tuindorp – 2015 Storm Event – fourteen manholes										
Model ID	**RMSE P1** (cm)	**RMSE P2** (cm)	**RMSE P3** (cm)	**RMSE P4** (cm)	**RMSE total** (cm)	**KGE r** (-)	**KGE beta** (-)	**KGE alpha** (-)	**KGE total** (-)	**NSE total** (-)
1D_St	*16.06*	*46.11*	*10.79*	*11.40*	**18.59**	*0.901*	*1.034*	*1.278*	**0.664**	**0.461**
1D/2D_St	*16.14*	*46.36*	*10.95*	*11.41*	**18.70**	*0.899*	*1.034*	*1.278*	**0.663**	**0.456**
1D/2D_P	*16.15*	*88.93*	*29.30*	*9.58*	**32.93**	*0.828*	*1.235*	*1.696*	**0.208**	**-0.758**
1D/2D+_H	*17.17*	*19.37*	*8.97*	*12.05*	**13.13**	*0.952*	*0.975*	*1.256*	**0.699**	**0.618**
1D/2D+_Y	*19.98*	*16.41*	*11.46*	*16.12*	**14.63**	*0.963*	*0.841*	*1.125*	**0.694**	**0.590**
1D/2D+_HY	*17.27*	*29.54*	*11.06*	*12.83*	**15.58**	*0.934*	*1.007*	*1.311*	**0.645**	**0.529**

Table 20 – Results for test case Tuindorp, 2015 storm event

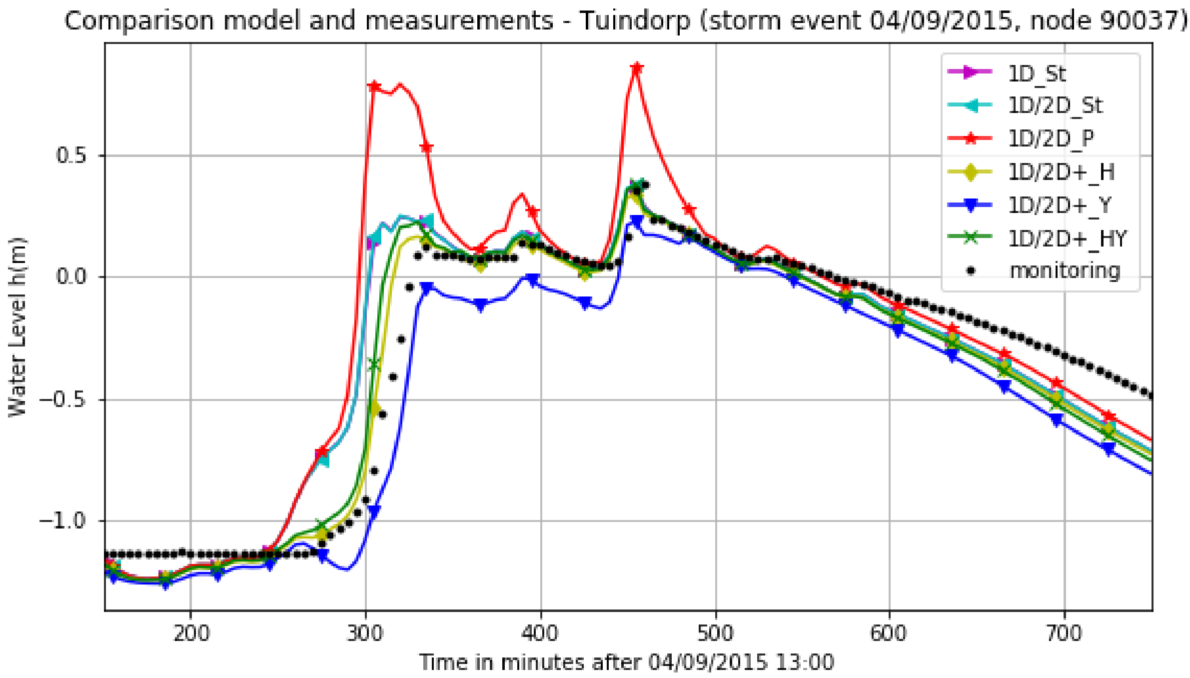


Figure 31 – Monitoring and modelling results for the 2015 storm event in Tuindorp. Manhole/node 90037 is selected.

8. Discussion

In Chapter 7, the results were presented. Based on both tests and all three test cases, no modelling concept should be clearly favoured above the others without specific knowledge of the case. However, the test results did show considerable differences between the modelling concepts in general and the four additional aspects. In this chapter, the results are discussed. The chapter consists of three parts. The first part of the discussion (§ 8.1) spans the four aspects that were determined in § 4.3. The second part (§ 8.2) concerns the general differences between main modelling concepts and aspects of attention when applying a certain modelling concepts. The last part of the discussion concerns the research method and the test data (§ 8.3). To confirm, check or highlight differences between models, references will be made to the flood maps in Appendix I, J and K.

8.1 Four additional aspects

The four additional aspects, for which an extra model was created, are discussed in this paragraph.

8.1.1 The inclusion of pervious surface area with semi-distributed runoff

The inclusion of pervious surface area with semi-distributed runoff was tested by creating the 1D/2D_P model to compare with the 1D/2D_St model. The 1D/2D_P model is the same as the 1D/2D_St model, except for the addition of all pervious surface area to the sub catchments. Flood comparison maps for the three test cases are found in Appendix K: Figure 69, Figure 73 and Figure 77 for test case Ulvenhout, Tuindorp and Loenen respectively.

Based on the both tests, the differences between the 1D/2D_St and 1D/2D_P model are significant. The 1D/2D_P model overestimates the amount of street flooding and sewer inflow. In test 1, the 1D/2D_P model scores much higher in the recognition of flooded grid cells (100% for test case Ulvenhout and default test settings) than the recognition of non-flooded grid cells (57%), indicating overestimation of street flooding. In the Tuindorp case in test 2, the 1D/2D_P has a RMSE of more than fifteen centimeters larger compared to 1D/2D_St model for both storm events during peak level. The reason for this is seen in Figure 30 and Figure 31. In models based on a 1D/2D concept, runoff is modelled semi-distributed. Surface runoff of pervious surface occurs when the surface storage of 6 mm is filled and the rainfall intensity surpasses the infiltration capacity (Horton initial f_0 is set at 5 mm/h). At the beginning of a storm event, the 1D/2D_St and 1D/2D_P model perform equal. When the first 6 mm of rain (plus infiltration losses) has fallen, the pervious surface area in the 1D/2D_P model starts to produce additional runoff, resulting in increased sewer inflow. The largest differences between the models are therefore found during the last part of the filling up of the system and during peak water levels. The flood comparison maps confirm the large difference between the two models: differences of more than 10 cm in flood depth are modelled.

A relatively small difference is found in the results of test 2 for Ulvenhout. The RMSE differs 1.4 cm for the 2016 storm event and 0.4 cm for the 2017 storm event. KGE and NSE differed not more than 0.07. A reason for this may be the location of the monitored manhole: a pump sump situated close to a weir (combined sewer overflow). During both storms, the nearby weir was overflowing (the crest level is at +3.34 m). It is therefore likely that part of the increased amount of sewer inflow is discharged via the weir. This theory is confirmed by the modelling results:

during the 2017 storm event, the weir discharged 764 m^3 (1D/2D_P model) and 340 m^3 (1D/2D_St model). Another reason may be the amount of pervious surface area. The monitored manhole is a pump sump that discharges water from a partly separated part of the sewer system (see Appendix L). This partly separated part has a relatively high degree of imperviousness. The additional amount of connected (pervious) surface area is therefore relatively small.

In general, the addition of pervious surface area does influence the amount of surface runoff. However, the 1D/2D_P model significantly overestimates the amount of surface runoff. There are two possible reasons for this. First, the infiltration parameters were based on the Dutch Leidraad Riolering C2100. Surface storage was set at 6 mm and Horton initial infiltration at 5 mm/h. The infiltration capacity is significantly lower compared to other literature (for instance, (US Environmental Protection Agency, 1999)). The parameters in the Leidraad Riolering C2100 are chosen to be 'on the safe side', according to the document (Stichting RIONED, 2004). Apparently, the infiltration parameters are too low. Secondly, all pervious surface area is added to the sub catchments. In a real situation, not all runoff from pervious area will reach the sewer system. A strip of grass close to a water body, a soccer field or a park is not likely to contribute to sewer inflow, even under extreme rainfall. If it is therefore decided to add pervious surface area with semi-distributed runoff, calibrated infiltration parameters and proper flow path analysis are vital elements. However, in practice, a proper calibration of pervious surface area will be hard. Runoff delay and surface storage will be hard to calibrate as a single value, as urban soils are complex and non-uniform.

8.1.2 The representation of roof surfaces in 1D/2D+ models

Two models were created with a different representation of roof surfaces: the 1D/2D+_Y and 1D/2D+_HY model. In the case of the 1D/2D+_Y model, all runoff is modelled fully distributed and houses are modelled as elevated plains. In the 1D/2D+_HY model, houses are modelled with semi-distributed runoff (referred to as hybrid runoff). Flood comparison maps for the three test cases are found in Appendix K: Figure 70, Figure 74 and Figure 78 for test case Ulvenhout, Tuindorp and Loenen respectively.

The results of the 1D/2D+_Y model show an underestimation of the amount of sewer inflow and a time shift in the filling up of the system. For all four storm events in test 2, the 1D/2D+_Y model has a lower β-term than the 1D/2D+_HY model, indicating that less water reaches the sewer system. The difference varies from 0.017 (Ulvenhout, 2016 storm event) to 0.166 (Tuindorp, 2015 storm event). The flood comparison maps confirm the difference: the 1D/2D+_Y model shows more flooding around buildings, especially in back yards, than the 1D/2D+_HY model. Less flooding on the streets is present. Apparently, a large portion of the rainfall that falls onto a roof surface does not reach the sewer system and infiltrates or accumulates around the building. The accumulation of water around buildings is confirmed by the flood contour maps (Appendix J) and flood difference maps (Appendix K). Furthermore, the sewer inflow is delayed, as the surface runoff reaches the gullies in a slower pace than in an actual situation (with rain gutters) or by assignment of surface area with fast runoff (semi-distributed runoff).

In contrast to test 2, the 1D/2D+_Y model scores almost equal to the 1D/2D+_HY model in test 1 for test case Ulvenhout (87% and 89% with default test parameters). As both models score high in true positives for both depth threshold configurations (5 cm and 10 cm), it indicates that despite the difference, both models generally result in more than 10 cm of flooding at the vulnerable locations.

In summary, the representation of roof surfaces is a significant factor in building a model based on a 1D/2D+ concept. The hybrid modelling of runoff cancels out part of the negative effect of the lack of private connections (which was also described in (Pina, et al., 2016)), whilst benefiting from the better representation of surface runoff (thus sewer inflow) due to fully distributed surface modelling.

8.1.3 The influence of paved yards on the amount of street flooding

The influence of paved yards on the amount of street flooding can be assessed by comparing the modelling results of the 1D/2D+_H and 1D/2D+_HY model. In the 1D/2D+_H model, it is assumed that all yards are pervious surface area with a high infiltration capacity. In the 1D/2D+_HY model, yards are further subdivided according to the NDVI-index based on an 80 cm satellite image. Flood comparison maps for the three test cases are found in Appendix K: Figure 71, Figure 75 and Figure 79 for test case Ulvenhout, Tuindorp and Loenen respectively.

In both test 1 and test 2, considerable differences are found between the 1D/2D+_H and 1D/2D+_HY model. In test 1 and for test case Ulvenhout, the 1D/2D+_HY reached a combined score of 89% for the default parameter settings. The 1D/2D+_H model scores 66%. The difference is confirmed in the sensitivity analysis: the 1D/2D+_HY model scores 11 percent points higher in the average of all sixteen test parameter configurations. Furthermore, the 1D/2D+_HY model most likely overestimates the amount of flooding. For all sixteen test parameter configurations, the model scores 94% for true positives and 61% for true negatives. The 1D/2D+_H model is the only model in test 1 where scores for true negatives are higher: 55% true positives, 77% true negatives. In test 2, none of the two models clearly perform better. However, differences in RMSE of up to 2 cm are recorded. As infiltration parameters are not validated for use in this research, it is possible that the 1D/2D+_H model performs better than the 1D/2D+_HY model, despite the addition of yard sub classification. In general, the models that use the infiltration parameters based on (US Environmental Protection Agency, 1999) seem to overestimate street flooding according to the Ulvenhout test case and the chosen test parameters.

The comparison of the 1D/2D+_H and 1D/2D+_HY model shows that paved yards have a significant influence on the amount of street flooding. The difference in street flooding is confirmed shown by comparing the flood maps and flood comparison maps. However, proper validation of infiltration parameters is very important and can impose a larger error that the exclusion of runoff from yards. Appendix N explores the variation of two input parameters for infiltration and confirms this statement. Furthermore, the addition of more surface types increases the amount of modelling parameters. This makes it harder to calibrate the model and increases the chance of over parametrisation: a too large amount of parameters causes a large amount of parameters combinations to give the same result. Finally, only one method of representing yards is used, based on the NDVI-index of a satellite image with 80 cm resolution. The resolution of the satellite image, resolution of the DEM and chosen grid size are most likely not good enough to accurately represent runoff from yards.

8.1.4 The influence of infiltration zones on the results of a 2D model

The 2D_St and 2D_I model were created to test the influence of infiltration zones on the results of a 2D model. In the 2D_St model, a standard amount of 20 mm/h was subtracted from the rainfall to account for infiltration and sewer discharge. In the 2D_I model, paved surfaces are assigned an 'infiltration' capacity of 60 l/s/ha and impervious surface 5 mm/h. Flood comparison maps for the three test cases are found in Appendix K: Figure 72, Figure 76 and Figure 80 for test case Ulvenhout, Tuindorp and Loenen respectively.

The differences between the two models were only assessed in test 1, as models based on a 2D modelling concept do not contain a hydrodynamic sewer model and are therefore not tested in test 2. In test 1 and for test case Ulvenhout, the 2D_I model scores slightly higher with a 69% combined score for all test parameter configurations, while the 2D_St model scores 66%. The difference between the scores for true positives and true negatives are smaller for the 2D_I model. Both models score higher in the recognition of true positives. The 2D_I model scores 96% and 41% for true positives and true negatives respectively, the 2D_St model scores 76% and 56%. For test case Tuindorp, both models score low in the recognition of flooded grid cells. Still, the 2D_I model scores higher than the 2D_St model (35% true positives compared to 22%).

The results indicate that both models tend to overestimate street flooding and that the 2D_I model overestimates the amount of street flooding more than the 2D_St model. The flood comparison maps confirm the difference: the 2D_I model results in more street flooding than the 2D_St. Most likely reason is the difference in 'average' infiltration/discharge capacity. The 2D_St model has a standard capacity of 20 mm/h, the 2D_I model has 21.6 mm/h for impervious surface and 5 mm/h for impervious surface.

Based on the test (test 1), no valid conclusions can be drawn on the benefits of adding infiltration zones. The theoretical benefit of adding infiltration zones would be that more flooding occurs in areas with a higher imperviousness. This cannot be tested by using the method proposed in this research. In the recommendations, some ideas about a better comparison between the two 2D models are shared.

8.2 General modelling concepts

This part of the discussion concerns the differences between the main classes of modelling concepts: 1D, 1D/2D, 1D/2D+ and 2D. This paragraph combines the overall modelling results and the four additional aspects to assess the difference between modelling concepts and points of attention for applying a certain modelling concept.

8.2.1 Comparison of 1D and 1D/2D modelling concepts

A 1D and a 1D/2D model (1D_St and 1D/2D_St) are assessed in both tests. As 1D models do not contain a surface model, test 1 is performed in a somewhat different manner (see § 6.1.3). For test case Tuindorp and Loenen, nearly the same scores are reached as both models mostly fail in recognizing the flooded locations.

For test case Ulvenhout, there are considerable differences. The 1D_St model scores almost equal to the 1D/2D_St model with default test parameter settings (86% and 85% combined score). However, the combined average of the sixteen test parameter configurations differs in favour of the 1D/2D_St model (81% compared to 73%). Furthermore, the 1D_St model shows a larger range in the sensitivity of the test parameters (64% – 86% compared to 76% – 87%).

Lastly, the results of the 1D_St model show a large imbalance between true positives and true negatives: 94% vs. 50% for the average of all sixteen test parameter configurations. The imbalance may be due to the flood modelling in models based on a 1D concept. A water depth of 10 cm in the virtual flood cone will most likely represent less flood volume than a water depth of 10 cm on an actual surface model. The threshold of water depth is therefore reached more often, resulting in more grid cells being marked 'flooded' according to the modelling results. This increases the score for true positives and decreases the score for true negatives. As this process has a minor influence on the combined score, it may be concluded that the 1D/2D_St model is better in estimating the locations of street flooding than the 1D model.

As a 1D and a 1D/2D model use the same hydrological model, they should perform exactly the same when no street flooding occurs or parameters are changed. In test 2, the two storm events in Tuindorp did not produce a large amount of street flooding. Therefore, the 1D and 1D/2D model performed very similar to each other. In the two storms tested in Ulvenhout, street flooding did occur. In the case of the 2016 storm event, the addition of an elevation model (1D/2D) results in an improvement of the RMSE, KGE and NSE. However, the performance indicators decrease in the 2017 storm event. The addition of an elevation model to allow 2D overland flow does not guarantee an improvement of the in-sewer performance indicators. The addition of 2D overland flow allows surface interaction between flooded manholes. If a 1D model overestimates the amount of sewer inflow, the resulting RMSE will be relatively high. If in the 1D/2D model, overland flow from another manhole causes even more sewer inflow, the resulting RMSE will be even higher. The results therefore show that for large storm events, differences are present between a 1D and a 1D/2D model. The addition of runoff from pervious surfaces (1D/2D_P model), has no influence on the difference between a 1D and a 1D/2D model, as the hydrological models are the same. Based on findings in (Freni, et al., 2010) and (Vojinovic & Tutulic, 2009), the introduction of overland flow results in a better modelling of urban flooding. Although this theory is not completely confirmed by the modelling results in this research, the conclusions are most likely to be true if the models would have been re-calibrated.

8.2.2 Comparison of 1D/2D and 1D/2D+ modelling concepts

In both tests, models based on a 1D/2D and 1D/2D+ modelling concept were compared. In test 1, the differences between the individual models (to investigate the four additional aspects) were larger than the general differences between the main modelling concepts. Especially the 1D/2D+_Y model deviates a lot from the other two 1D/2D+ models. The hybrid runoff model was therefore decided to provide significant benefits (§ 8.1.2).

The range in scores of the 1D/2D models over the sixteen test parameter configurations was smaller than the range in scores of the 1D/2D+ models, considering the combined score for the Ulvenhout test case. This difference in range is explained by the difference in runoff. In the case of a 1D/2D model, runoff is modelled semi-distributed: sewer inflow is modelled by the assignment of surface area to the manholes in the sewer system. Only when a manhole overflows, water enters the 2D surface model. Surface flooding therefore 'tends' to stay close to the manholes. In a 1D/2D+ model, rain water is distributed equally over each grid cell. This means that if water cannot reach a gully, forms a puddle (in a topographic depression) or if a gully cannot cope with the amount of inflow, the water accumulates on the surface without entering the sewer system. A 1D/2D+ model therefore tends to show flooding at more locations then only at overflowing manholes. As these puddles or overstressed gullies may or may not be recognized by the chosen parameters in test 1, test results deviate more than in the case of a

1D/2D model. The flood maps (Appendix I) and especially the flood contour maps (Appendix J) support the explanation.

In test 2 and for the 2016 storm event in Ulvenhout, the performance of the 1D/2D models is roughly equal to the 1D/2D+ models, while the 1D/2D+ models score higher for all overall indicators for the 2017 storm event. A possible reason for this seems to be the reaction to the smaller storm event (about 6 mm). The 1D/2D+ models show a smaller increase in water levels compared to the 1D/2D model, therefore resulting in a lower RMSE. The difference can be caused by either a difference in surface runoff model or the suspected extra conduit, which is depicted in Appendix L. It may therefore not be concluded safely that the 1D/2D+ performed better in terms of runoff for the 2017 test case.

In test case Tuindorp, the 1D/2D+ models score better in terms of RMSE during the filling up of the system (P2) compared to 1D/2D models. Especially for the 2015 storm event the difference is large: the 1D/2D+_H model has an RMSE of 19.37 cm for P2, the 1D/2D_St model 46.36 cm. The most likely reason for this is (again) the difference in runoff modelling. The 1D/2D model models runoff semi-distributed, with predefined parameters for runoff delay and surface storage. In the initial observation of the modelling results (§7.2.2), it was pointed out that the original 1D model struggled to model the initial sewer inflow and the first peak (Appendix M). As the provided 1D model (1D_St) uses the same hydrological model as the 'basic' 1D/2D model (1D/2D_St), it was to be expected that the bias of the 1D/2D in the filling up of the system would also be large. The 1D/2D+ models simulate runoff in a different manner. No surface dependant constants are used for runoff delay and surface storage is modelled by the filling up of actual topographic depressions in the elevation map.

According to the results, fully distributed runoff modelling provides benefits in the modelling of the filling up of the system. However, sewer inflow is likely to be underestimated as the β-term in test 2 is less than zero and flood maps in Appendix I confirm more water between houses and in topographic depressions. Modelling runoff as a hybrid provides benefits: the β-term is higher and more water reaches the sewer system instead of accumulating around buildings. By using a hybrid runoff model and high resolution grid, the negative side effect seems to of less influence. In Dutch test cases and by using a high resolution grid and hybrid runoff, 1D/2D+ models therefore have a higher potential in flood estimation than 1D/2D models. However, a large amount of input data is needed and run time is large.

8.2.3 2D modelling concepts

The performance of a 2D model was only tested in test 1. Both 2D models (2D_St and 2D_I) scored higher in the recognition of flooded grid cells (true positives) than the recognition of non-flooded grid cells (true negatives), indicating overestimation of street flooding. The overestimation is confirmed when the flood maps (Appendix I) of all three test cases are assessed. Based on this research, 2D models are good estimators of flooded locations (true positives), but the extent is generally overestimated. The most likely reason for this is the choice for infiltration/discharge parameters. The amount of test data and the characteristics of the method lacks further argumentation on the extent of overestimation and types of flooding that are (or are not) recognized properly, as was already stated in § 8.1.4.

8.3 Research methods and test data availability

This paragraph discusses the method of comparing models and the information provided for that.

8.3.1 Research methods

In test 1, the core of the test is the translation from two different entities (information in multiple forms vs. modelled street flooding from InfoWorks ICM) into a form in which it is comparable. Precise flood depths, flood volumes and a flood extent are almost never known for a specific flood event. By comparing the two entities, it is assumed that the data for observed flooding has perfect accuracy. Perfect accuracy in this context means the combination of a lack of temporal differences in observations and a spatial resolution which is higher than the resolution of the grid cells in the tests. In reality, the various sources of information have a different non-perfect accuracy.

The proposed test 1 inhabited various parameters: grid size, areal threshold, depth threshold and three thresholds for assignment of flooding to a grid cell based on the observations (one for point data, one for string data and one for polygon data). For each test case, a set of default test parameters was selected. In order to test the sensitivity of part of some of these parameters (grid size, areal threshold, and depth threshold), sixteen test parameter configurations were chosen and tested for the three cases. Four grid sizes were tested. Four combinations of depth and areal thresholds were tested for models with a 2D surface model. For the 1D_St model, the sensitivity of depth threshold and the method of conversion from manhole flooding to flood extent were tested.

In general, the test parameters have a large influence on the scores of test 1. For the combined score (Ulvenhout), the smallest range in scores is 13 percent points (76-89%, 1D/2D_St model). The largest range is 39 percent points (52-91%, 1D/2D+_HY model). As already explained, the depth and area threshold influence the balance between true positive and true negative results. For only true positives or true negatives, the ranges are therefore even large (up to 100 percent points). Considering the individual parameters that were tested, the varying of the depth threshold proved to be the largest influence on the ranking of the models.

The highest scores for true positives were reached for a low threshold for depth and area. This makes sense, as the combination of an aerial threshold of 5 m^2 and a depth threshold of 5 cm results in a low boundary for a grid cell to be marked flooded. The size of the grid influences the modelling results: differences of more than 30 percent points and multiple places in (relative) ranking are found. Based on the results, no grid size should be clearly favoured above the others. In the choice of a proper grid size, a balance should be found. A too large grid size causes inaccuracy and possible loss of information (two parallel streets may be caught in one grid cell). A small grid size is only possible if the accuracy of the data is good enough. The ideal grid size is therefore different for each test case and cannot be recommended as a single value.

Test 2 is a method of quantifying the difference between the models and measurements. The RMSE, in combination with the partial values for the four phases, proved to be a good indicator of differences between modelling concepts. Most of the influences on the RMSE are explainable and confirmed by assessing flood maps. The NSE, KGE and RMSE generally agree with each other. For test case Tuindorp, where the largest amount of information was available, the best and worst scoring models are the same for the three indicators. The β-term of the KGE, which is the ratio of the mean of the observations and corresponding model, proved to be a good measure for overestimation or underestimation of a model. The indicator was also used in (Pina, et al., 2016). As the KGE of a specific flood event may be strongly dominated by a certain term (for instance, α-term in Tuindorp), assessing the partial values are important in understanding the test case and the value of possible conclusions. Also in test case Ulvenhout, local conditions (the suspected extra conduit) are suspected to influence the results. Good understanding of the local situation is therefore needed for a strong assessment.

8.3.2 Data availability

Test 1 was only executed fully for test case Ulvenhout. The results provided more information on differences between models than the other test cases. In the case of Tuindorp, no information was available on non-flooded locations. In the case of Loenen, only one location of flooding was available and the locations were not linked to a specific storm event. By comparing the imbalance between true positive and true negative scores for different models, overestimation and underestimation can be detected. This overestimation and underestimation is strongly influenced by the chosen test parameters. The 'full' test therefore provided significantly more information.

Test 2 was executed for test case Tuindorp and Ulvenhout. Results were more constant for Tuindorp. Possible explanation for this is the number of manholes and that the Tuindorp sewer system was cleaned in the time before the monitoring campaign. In general, the quality of the tests depends on the amount of information that is put in. When more manholes are assessed, local deviations between the models and measurements are likely to be averaged out.

Considering the data that was provided for this research, more data would be needed for a better quantification of the differences between modelling concepts. The data did provide arguments to explain differences between modelling concepts, but differences varied strongly between test cases.

In test 1, the best available source of information, the video footage from Ulvenhout, still only covered a fraction of the total street surface. Aerial image from satellites or drones could provide an enormous database of consistent (in time) and reliable information to be used in the test. However, the chance that a satellite flies over during the time of flooding is small. Furthermore, privacy rules and foliage provide barriers for the use of drones.

9. Conclusions and Recommendations

This chapter sums up the most important conclusions and provides recommendations.

9.1 Conclusions

This research introduced a method to assess different concepts often used in modelling the urban drainage system. The method consists of two tests and the method was tested using three different test cases: Loenen, Ulvenhout and Tuindorp. In the comparison of main modelling concepts (1D, 1D/2D, 1D/2D+ and 2D) and assessment of additional aspects, some significant conclusions can be drawn based on this research.

Extending a 1D model to a 1D/2D model by adding a 2D surface model improves its abilities to represent street flooding for heavy storm events. The 'standard' 1D/2D model (1D/2D_St) scored 81% as an average of sixteen test parameter configurations and for test case Ulvenhout. The 1D model (coded 1D_St), which used the exact same hydrological model, scored 73%. When no street flooding occurs, the models should perform the same. In the assessment of in-sewer monitoring data, this was indeed the case: the RMSE varied only 0.01 cm for a storm event in Tuindorp (2014) where no significant street flooding occurred. The difference in RMSE for a heavier storm event increases to 1.3 cm (Ulvenhout, 2016 storm event). Despite the difference, the 1D/2D model did not always perform better than a 1D model. The reason for this is that the models were not re-calibrated. The expected benefits of allowing overland flow agrees with findings in literature (Vojinovic & Tutulic, 2009) (Freni, et al., 2010).

When creating a model with semi-distributed runoff (1D or 1D/2D), the addition of pervious surface area influences the amount of street flooding. The 1D/2D_P model that was created scored higher in the recognition of flooded grid cells than the 1D/2D_St model. The score for true positives increased from 8% to 96% for test case Tuindorp and using default test parameters. However, the model scores worse in the recognition of non-flooded grid cells, indicating overestimation of street flooding. The results imply that the set of parameters often used in the Netherlands (based on Leidraad Riolering C2100) provides unrealistically low parameters for infiltration and surface storage of pervious surface types. Furthermore, not all pervious surface area is likely to produce runoff, even under extreme rainfall conditions. Proper validation of infiltration parameters and flow path analysis are advised.

A model with fully distributed runoff provides benefits over semi-distributed runoff. The 1D/2D+ models were better estimators of sewer inflow and first peaks than the 1D or 1D/2D models. The best performing 1D/2D+ model for the 2015 storm event in Tuindorp (1D/2D+_H) granted a reduction of 27 cm in RMSE during the filling up of the system and 2 cm for peak water levels. However, models with fully distributed runoff tend to overestimate flooding around houses and underestimate sewer inflow due to incorrect representation of private connections, such as roof gutters. The findings agree with (Pina, et al., 2016). The use of a hybrid runoff model, where runoff from roofs is modelled semi-distributed, cancels part of the negative tendency and is advised to be used.

Runoff from paved yards also influences on the amount of street flooding. The 1D/2D+_H model, which assumes that all yards are pervious, scores 39%, 71% and 0% in the recognition of flooded grid cells for the test cases Ulvenhout, Tuindorp and Loenen respectively (default test parameters). The 1D/2D+_HY model, which subdivides yards into pervious, semi-pervious and impervious, scores 89%, 75% and 100% for the three test cases. However, as infiltration parameters were not validated in this research, the addition (semi-)impervious zones in yards did not always results in a better estimation of sewer inflow or flooding without overestimation of underestimation. The values for the infiltration parameters are key elements in the calibration of a model with fully distributed runoff and can pose an even larger error than the method for representing yards.

On the inclusion of infiltration zones in 2D models, no conclusion was deemed valid due to lack of test data and the inability of the test methods to cover correlation between imperviousness and flooding. The differences that were found are most likely due to the difference in parameters for discharge and infiltration.

9.2 Recommendations

Based on this research, the following recommendations are made for further research:

- Gathering of data during an intense storm event

The largest threat for the accuracy of test 1 is the quality of input data. The test assumes equally (highly) accurate data. This research recommends more gathering of flood data during intense storm events. The use of aerial photographs by drones (or satellites) could provide a useful source on information. The images are affordable, high quality images can be taken that cover a large amount of area. The data is furthermore equal in time, which is not the case for a driving vehicle. However, privacy regulations and foliage make widespread implementation difficult.

- Addition of more test cases

In this research, three test cases were used: Tuindorp, Ulvenhout and Loenen. By adding more test cases with more, the differences between models can be quantified in more detail.

- Method for assessing differences between 2D models

In § 8.1.4, it was concluded that a valid comparison between characteristics of a 2D model is not possible with the method proposed and the information provided. In order to test differences between the 2D models, the following method is advised: first a map with imperviousness is needed. Then, the correlation between the model and the imperviousness is determined and compared with the correlation of the 2D_St model and other well-performing models. In this way, it can be tested whether the addition of infiltration zones improves flood estimation.

- Automating the tests

In this research, test 1 was mostly executed by making use of GIS-software. Most operations were performed only one step at the time. By making use of scripts, larger parts of the processes can be automated. This can save a lot of time if multiple test cases are assessed.

- Calibration of models

Re-calibration of modelling parameters with monitoring data was not part of this research. As a result of this, the addition of elements to the model that should most certainly improve results, sometimes had a negative effect.

Bibliography

Actueel Hoogtebestand Nederland, 2018. *AHN3,* Amersfoort: AHN Nederland.

Adeogun, A., Daramola, M. & Pathirana, A., 2015. Coupled 1D-2D hydrodynamic inundation model for sewe overflow: influence of modelling parameters. *Water Science,* Volume 29, pp. 146-155.

Bazalgette, J. & Forrest, J., 1865. *On the main drainage of London : and the interception of the sewage from the River Thames.* London: W. Clowes and Sons.

Bosselaar, J., 1940. *Stedelijke Rioleeringen.* Deventer: A.E. Kluwer.

Clemens, F., 2001. *Hydrodynamic models in urban drainage: application and calibration (PhD Thesis).* Delft: DUP Science.

Deltaprogramma, 2017. *Deltaprogramma 2018,* s.l.: Ministerie van Infrastructuur en Milieu & Ministerie van Economische Zaken.

DINOLoket, 2017. *Ondergrondgegevens.* [Online]
Available at: https://www.dinoloket.nl/ondergrondgegevens
[Accessed 30 November 2017].

Djordjevic, S., Prodanovic, D. & Maksimovic, C., 1999. An approach to simulation of dual drainage. *Water Science & Technology,* 39(9), pp. 95-103.

Environment Agency, 2009. *Desktop review of 2D hydraulic modelling packages,* Bristol, UK: Environment Agency.

Environment Agency, 2013. *Benchmarking the latest generation of 2D hydraulic modelling packages,* Bristol, UK: Environment Agency.

Freni, G., La Loggia, G. & Notaro, G., 2010. Uncertainty in urban flood damage assessment due to urban drainage modelling and depth-damage curve estimation. *Water Science & Technology,* Volume 61, pp. 2979-2993.

Gemeente Enschede, 2018. *Riolering.* [Online]
Available at: https://www.enschede.nl/openbare-ruimte/grondwaterenriool/riolering
[Accessed 2 May 2018].

Gupta, H., Kling, H., Yilmaz, K. & Martinez, G., 2009. Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *Journal of Hydrology,* Volume 377, pp. 80-91.

Henckens, G., Langeveld, J. & van Berkum, P., 2003. Kalibratie van het hydrodynamische rioleringsmodel van Loenen. *Rioleringswetenschap en techniek,* 3(11), pp. 45-60.

Henderson Water Utility, 2018. *Combined Sewer System.* [Online]
Available at: https://www.hkywater.org/departments/wastewater/combined-sewer-system
[Accessed 3 April 2018].

Henonin, J., Russo, B., Mark, O. & Gourbesville, P., 2013. Real-time urban flooding forecasting and modelling - a state of the art. *Journal of Hydroinformatics,* Volume 15, pp. 717-736.

Horton, J., 1940. An approach towards a physical interpretation of infiltration capacity. *Proceedings of the Soil Science of America,* Volume 5.

Kadaster, 2018. *BGT.* [Online]
Available at: https://www.kadaster.nl/bgt
[Accessed 25 February 2018].

Kling, H., Fuchs, M. & Paulin, M., 2012. Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios. *Journal of Hydrology,* Volume 424-425, pp. 264-277.

KNMI, 2014. *KNMI Klimaatscenarios voor Nederland,* De Bilt: Koninklijk Nederlands Meteorologisch Instituut.

Langeveld, J., 2004. *Interactions within wastewater systems (PhD Thesis),* Delft: Delft University of Technology.

Leandro, J., Chen, A., Djordjevic, S. & Savic, D., 2009. Comparison of 1D/1D and 1D/2D Coupled (Sewer/Surface) Hydraulic Models for Urban Flood Simulation. *Journal of Hydraulic Engineering,* 135(6), pp. 495-504.

Mann, E., 2016. *Story of cities #14: London's Great Stink heralds a wonder of the industrial world.* [Online]
Available at: https://www.theguardian.com/cities/2016/apr/04/story-cities-14-london-great-stink-river-thames-joseph-bazalgette-sewage-system
[Accessed 3 April 2018].

Martens, H. & Magni, M., 2001. *Multivariate Analysis of Quality.* 1st ed. Chichester, England: John Wiley & Sons Ltd..

Martin, A., 2011. *Modelling Road Gullies.* Staplefield (UK), International Flood and Modelling Conference.

Merz, B., Kreibuch, H., Thieken, A. & Schmidtke, R., 2004. Estimation uncertainty of direct monetary flood damage to buildings. *Natural Hazards and Earth System Sciences,* Volume 4, p. 153–163.

Motohka, T., Nasahara, K., Oguma, H. & Tsuchida, S., 2010. Applicability of Green-Red Vegetation Index for Remote Sensing of Vegetation Phenology. *Remote Sensing,* Volume 2, pp. 2369-2387.

Nash, J. & Sutcliffe, J., 1970. River flow forecasting through. Part I. A conceptual models discussion of principles. *Journal of Hydrology,* Volume 10, pp. 282-290.

Netherlands Space Office, 2018. *Triplesat.* [Online]
Available at: https://www.spaceoffice.nl/nl/satellietdataportaal/beschikbare-data/triplesat/
[Accessed 7 February 2018].

Newman, P., 2001. *Daily Life in the Middle Ages.* 1st ed. Jefferson, U.S.: McFarland.

Nieuwenhuis, E. et al., 2018. Statistical modelling of Fat, Oil and Grease (FOG) deposits in wastewater pump sumps. *Water Research,* Volume 135, pp. 155-167.

Pina, R. et al., 2016. Semi vs. Fully distributed Urban stormwater models: model set up and comparison with two real case studies. *Water,* 8(58), pp. 1-20.

Post, J., 2016. *A statistical approach to guide the management of the anterior part of the sewer (PhD Thesis),* Delft: Delft University of Technology.

Saint-Venant, A., 1871. Théorie du mouvement non permanent des eaux, avec application aux crues des rivières et a l'introduction de marées dans leurs lits. *Comptes rendus de l'Académie des Sciences,* Volume 73, pp. 147-154, 237-240.

Sanchez, A., 2018. *Question Dutch benchmark (Email),* s.l.: s.n.

Snow, J., 1855. *On the Mode of Communication of Cholera.* 2nd ed. London: John Churchill.

Spekkers, M., 2015. *On rainstorm damage to building structure and content (PhD Thesis),* Delft: Delft University of Technology.

Standards for Highways, 2017. *Design Manal for Roads and Bridges, Volume 4.* [Online]
Available at: http://www.standardsforhighways.co.uk/ha/standards/dmrb/vol4/section2.htm
[Accessed 11 April 2018].

Stichting RIONED, 2004. *Leidraad Riolering: Module C2100,* Ede: Stichting RIONED.

Stichting RIONED, 2014. *Ervaringen met de aanpak van regenwateroverlast in bebouwd gebied,* Ede: Stichting RIONED.

Stichting RIONED, 2017a. *Over de Kennisbank.* [Online]
Available at: https://www.riool.net/over-de-kennisbank
[Accessed 27 March 2018].

Stichting RIONED, 2017b. *Stichting RIONED in het kort.* [Online]
Available at: https://www.riool.net/over-rioned/stichting-rioned-in-het-kort
[Accessed 27 March 2017].

STOWA, 2017. *Benchmark Inundatiemodellen,* Amersfoort: STOWA.

STOWA, 2018. *Neerslagstatistieken voor korte duren. Actualisatie 2018,* Amersfoort: STOWA.

ten Veldhuis, J. & Clemens, F., 2010. Flood risk modelling based on tangible and intangible urban flood damage quantification. *Water Science & Technology,* Volume 62, pp. 189-195.

U.S. Geological Survey, 2018. *Unsaturated Zone.* [Online]
Available at: https://water.usgs.gov/ogw/unsaturated.html
[Accessed 3 April 2018].

US Environmental Protection Agency, 1999. *Infiltration through disturbed urban soils and compost-amended soil effects on runoff quality and quantity,* Washington, US: Office of Research and Development.

Vaes, G. & Berlamont, J., 1996. Composietbuiten als neerslaginvoer voor rioleringsberekeningen. *Water,* Volume 88, pp. 143-148.

van Bijnen, M., Korving, H., Langeveld, J. & Clemens, F., 2017. Calibration of hydrodynamic model-driven sewer maintenance. *Strcture and Infrastructure Engineering,* 13(9), pp. 1167-1185.

Van de Ven, F., 2016. *Water Management in Urban Areas (Lecture Notes TU Delft course CIE5510).* Delft: Delft University of Technology.

Vereniging VPB, 2008. *Handboek Rioleringstechniek,* Woerden: Vereniging VPB.

Vojinovic, Z. & Tutulic, D., 2009. On the use of 1D and coupled 1D-2D modelling approaches for assessment of flood damage in urban areas. *Urban Water Journal,* 6(3), pp. 183-199.

Weier, J. & Herring, D., 2000. *Measuring Vegetation (NDVI & EVI).* [Online]
Available at: https://earthobservatory.nasa.gov/Features/MeasuringVegetation/
[Accessed 28 March 2018].

Wiesmann, U., Choi, I. & Dombrowski, E., 2006. *Fundamentals of Biological Wastewater Treatment, 1.* 1st ed. Hoboken, U.S.: John Wiley & Sons.

Appendices

A. Three Sources of Model Data in the Netherlands

In the Netherlands, data availability for modelling urban drainage is generally good. The following three sources contribute to this.

Algemeen Hoogtebestand Nederland (AHN)

The AHN is a freely available DEM (Actueel Hoogtebestand Nederland, 2018)(see Figure 32). The AHN2 has a resolution of 0.5 m and a systematic and stochastic error of not more than 5 cm. A corrected version is available where trees, houses and cars have been filtered out. The most recent version is the AHN3, but this version is not yet available for the whole Netherlands. The AHN can be downloaded or viewed in an ArcGIS online application.

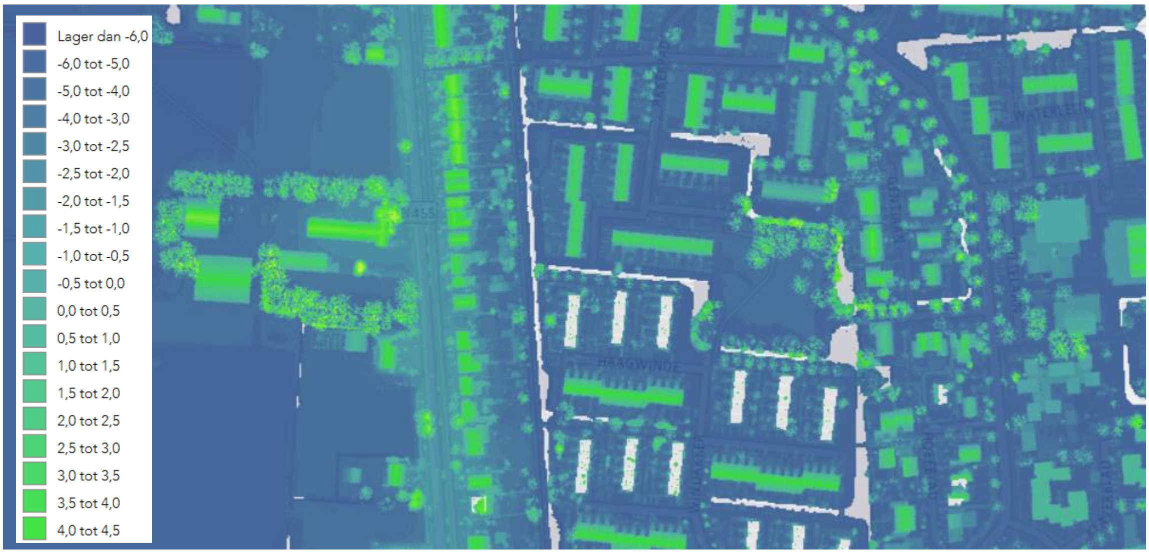


Figure 32 – The AHN, which can be downloaded or viewed in an ArcGIS online application (Actueel Hoogtebestand Nederland, 2018). The unit for the numbers provided in the legend is meters.

Kennisbank Stedelijk Water

The Kennisbank Stedelijk Water (translated: Knowledge Database Urban Water), formerly known as the Leidraad Riolering (translated: Sewer Guideline), is a database of documents concerning sewers, including modules on design principles, maintenance, laws, research, finances and sewer models (Stichting RIONED, 2017a). The database is maintained by the RIONED Foundation, the umbrella organisation for urban water management and sewerage in the Netherlands (Stichting RIONED, 2017b). For the modelling of urban drainage, the database provides information on sewer parameters and design storms.

Basisregistratie Grootschalige Topografie (BGT)

The BGT is a freely available database with information on land use (Kadaster, 2018). Houses, streets, pavement, yards, waterways and parks are all defined. This allows for an easy definition of paved and unpaved surfaces. The information can be downloaded and used directly in a GIS application.

B. Land use classification according to the Dutch Leidraad Riolering

In the Dutch Leidraad Riolering module C2100, a set of input parameters is provided for twelve possible surface classes. These classes can be used as a guideline for the creation of sub catchments. Table 21 shows the twelve categories and the corresponding parameters. The following remarks should be made:

- Four categories of surface are defined: closed impervious, open impervious, roofs and pervious. Close impervious surface is surface with no infiltration capacity, such as asphalt and concrete. Open impervious surface is surface with a small infiltration capacity, such as tiled pavement.
- Three types of slope are defined: inclined, flat and flat wide. A ‘flat, wide’ surface has a higher surface storage and slower runoff delay compared to a ‘flat’ surface.
- Storage is a combination of surface storage and initial losses.
- Two values are shown for the Horton k (see equation 1 for the parameters). A distinction is made between the value for decay and recovery. After the storage capacity of the unsaturated zone is depleted, the storage slowly empties because of groundwater recharge and evaporation. These processes are captured in the Horton k for recovery.

One may notice that the values for the infiltration parameters are somewhat low. According to the document, parameters are chosen ‘on the safe side’ to prevent an underestimation of sewer inflow (Stichting RIONED, 2004). In practice (and in the test cases used in this research) not all of the twelve categories are distinguished. Usually, only four categories are used: closed impervious (flat), open impervious (flat), roofs (inclined) and roofs (flat). Reasons for this are:

- Inclined surfaces are quite uncommon in most parts of the Netherlands.
- The difference between ‘flat’ and ‘flat, wide’ is quite arbitrary and wide open surfaces are not common in villages and cities.
- Pervious terrain is often not taken into account as runoff from pervious surface is assumed negligible; only impervious surface area is added to the sub catchments.

Type of surface	Type of Slope	Runoff Delay	Storage	Horton f_0	Horton f_c	Horton k_{decay}	Horton $k_{recovery}$
		sec	*mm*	*mm/h*	*mm/h*	*1/h*	*1/h*
Closed impervious	**Inclined**	*120*	*0.0*	-	-	-	-
	Flat	*300*	*0.5*	-	-	-	-
	Flat, wide	*500*	*1.0*	-	-	-	-
Open impervious	**Inclined**	*120*	*0.0*	*2.0*	*0.5*	*3.0*	*0.1*
	Flat	*300*	*0.5*	*2.0*	*0.5*	*3.0*	*0.1*
	Flat, wide	*500*	*1.0*	*2.0*	*0.5*	*3.0*	*0.1*
Roofs	**Inclined**	*120*	*0.0*	-	-	-	-
	Flat	*300*	*2.0*	-	-	-	-
	Flat, wide	*500*	*4.0*	-	-	-	-
Pervious	**Inclined**	*120*	*2.0*	*5.0*	*1.0*	*3.0*	*0.1*
	Flat	*300*	*4.0*	*5.0*	*1.0*	*3.0*	*0.1*
	Flat, wide	*500*	*6.0*	*5.0*	*1.0*	*3.0*	*0.1*

Table 21 – The twelve categories of land use as defined in the Leidraad Riolering (Stichting RIONED, 2004)

C. Infiltration Parameters for Infiltration Zones

In 1999, the US Environmental Agency researched the infiltration of disturbed urban soils (US Environmental Protection Agency, 1999). Eight different categories of soils were defined. The categories and the number of field tests per category are shown in Figure 33. The recorded infiltration rates were fitted to the Horton infiltration equation (Equation 1). A summary of the composed box plot probabilities is shown in Figure 34.

All three test cases have a dominant sandy soil type (see Chapter 5) (DINOLoket, 2017). For the test cases Tuindorp and Ulvenhout, the 'median' values for compact sandy soil are selected. The corresponding values for f_0, f_c and k are 127 mm/h, 12.7 mm/h and 6 respectively. In the case of Loenen, the groundwater table is very low, > 10 meters below ground level (DINOLoket, 2017). It is therefore expected that the infiltration capacity will be higher. The '75%' values are selected. The corresponding values for f_0, f_c and k are 304.8 mm/h, 31.75 mm/h and 12 respectively.

Category	Moisture	Texture	Compaction	Number of Tests
1	Saturated	Clay	Compacted	18
2	Saturated	Clay	Non-compacted	27
3	Saturated	Sand	Compacted	18
4	Saturated	Sand	Non-compacted	12
5	Dry	Clay	Compacted	15
6	Dry	Clay	Non-compacted	17
7	Dry	Sand	Compacted	21
8	Dry	Sand	Non-compacted	24

Figure 33 – Amount of tests per soil type (US Environmental Protection Agency, 1999)

Infilatration Parameter	Soil Group	90%	75%	Median	25%	10%
f_o (in/hr)	Clay - Dry Noncompact	42	24	11	7	5
	Clay - Other	7	3.75	2	1	0
	Sand - Compact	42	12	5	1.5	0
	Sand - Noncompact	52	46	34	24	0.25
f_c (in/hr)	Clay - Dry Noncompact	20	12	3	0.75	0.25
	Clay - Other	0.75	0.5	0.25	0	0
	Sand - Compact	5	1.25	0.5	0.25	0
	Sand - Noncompact	24	19	15	9	0
k	Clay - Dry Noncompact	18	13	9.5	4.5	3
	Clay - Other	11	6.5	3.75	1.75	0
	Sand - Compact	17	12	6	3	1
	Sand - Noncompact	19	12	5	2	0
15 minutes averaged (in/hr)	Clay - Dry Noncompact	28	14	6	3	2
	Clay - Other	4	2	1	0.25	0
	Sand - Compact	12	8	4	2	0.5
	Sand - Noncompact	37	29	25	17.5	6.5
30 minutes averaged (in/hr)	Clay - Dry Noncompact	23	19	6	2	1.75
	Clay - Other	2.5	1.75	1	0.25	0
	Sand - Compact	8	6	2.75	1.75	0.25
	Sand - Noncompact	29	26	20	16	5
60 minutes averaged (in/hr)	Clay - Dry Noncompact	23	17	6	2	1.5
	Clay - Other	2	1	0.5	0.25	0
	Sand - Compact	0.75	5	2	1	0.25
	Sand - Noncompact	26	22	17.5	12	4
120 minutes averaged (in/hr)	Clay - Dry Noncompact	22.5	16	5	1	0.75
	Clay - Other	1.25	0.75	0.5	0.25	0
	Sand - Compact	6	4	1	0.5	0
	Sand - Noncompact	24	20	16	11	3

Figure 34 – Horton infiltration parameters retrieved from the tested soils. The units are in inch/hr (1 inch/hr = 25.4 mm/hr) (US Environmental Protection Agency, 1999)

D. Classification of yards based on the NDVI-index

As mentioned in § 4.3.3, the Dutch database for land cover, BGT, does not differentiate between paved front and back yards (and some private terrain belonging to companies). The influence of paved yards on the amount of surface flooding is an aspect that is analysed in this research. This appendix will explain the steps that are taken to subdivide yards according to the NDVI-index (Normalized Difference Vegetation Index).

In a colourized photograph, three colour bands are distinguished: red, green and blue (hence the common abbreviation, 'RGB-image'). Each object on which a ray of light falls reflects light in a different manner and that is the reason why we humans experience different colours. Nowadays, satellites are able to distinguish more colour bands than humans can visually experience. (Near) infrared (NIR) imagery is an important example of this.

Healthy vegetation has the important property that most of the visual light is not reflected, but used for photosynthesis. In contrast, a large portion of the near infrared light is reflected (Weier & Herring, 2000). The NDVI index is a normalized ratio of red and (near) infrared reflectance, and is defined as follows (Motohka, et al., 2010):

$$NDVI = \frac{\rho_{nir} - \rho_{red}}{\rho_{nir} + \rho_{red}} \qquad \text{(Equation 3)}$$

ρ_{nir} = reflectance of near infrared light (-)
ρ_{red} = reflectance of near red light (-)

The value of the NDVI is always between minus one (-1) and plus one (+1). When no vegetation is present, the value is close to zero. A value close to one indicates a high density of green leaves.

The *Netherlands Space Office* provides a freely available database with images of various satellites (Netherlands Space Office, 2018). The four-band (RGB + NIR) images of the Triple sat, with a resolution of 80 cm, are used for the classification in this research. For the best possible classification, the following conditions should apply:

- A resolution of the image should be as high as possible.
- Optimum time of the day: around 13:00, when the sun causes the least amount of shadow. Shadow causes difficulties for the classification.
- Optimum time of the year: as the reflectance of green leaves is assessed, images taken during the late autumn or winter worsens the discrimination.
- Minimal cloud cover. Classification is not possible under clouded parts.
- An optimal angle of the satellite, represented by the Off-Nadir angle.

The satellite images can be imported as a raster into GIS. The steps of the process from a downloaded image to the classified yards are as follows:

- A satellite image is retrieved and important into GIS.
- Based on the BGT, terrain is divided into impervious surface area, pervious surface area and yards.
- The NDVI index is calculated for the area, according to the formula as presented in Equation 3.
- The shapes (polygons) with front and back yards are clipped against the indexed satellite image. This means that the front and back yards are cut out. As the BGT already differentiates houses, roads, parks and pavements with a decent accuracy, a classification according to the NDVI index would most likely worsen the accuracy of the polygons.
- Classes are made. In this research, all terrain with $NDVI < 0.1$ is classified as impervious. Terrain with $0.1 \leq NDVI < 0.4$ is classified semi-pervious. Terrain with $NDVI \geq 0.4$ is classified pervious. These values are chosen quite arbitrarily; they are determined by examining the results of different values.

Figure 35, Figure 36, Figure 37 and Figure 38 show the classification process for test case Ulvenhout. Figure 35 shows an aerial image of Ulvenhout. In Figure 36, the information from the BGT is used to classify the area into impervious terrain (black), pervious terrain (green) and yards (blue). It is clear from the picture that the fraction of front yards, back yards and other undefined private terrain can be very high. Figure 37, the results of the NDVI-classification are shown. Figure 38 shows the final result of the classification.

Figure 35 – Aerial image of Ulvenhout

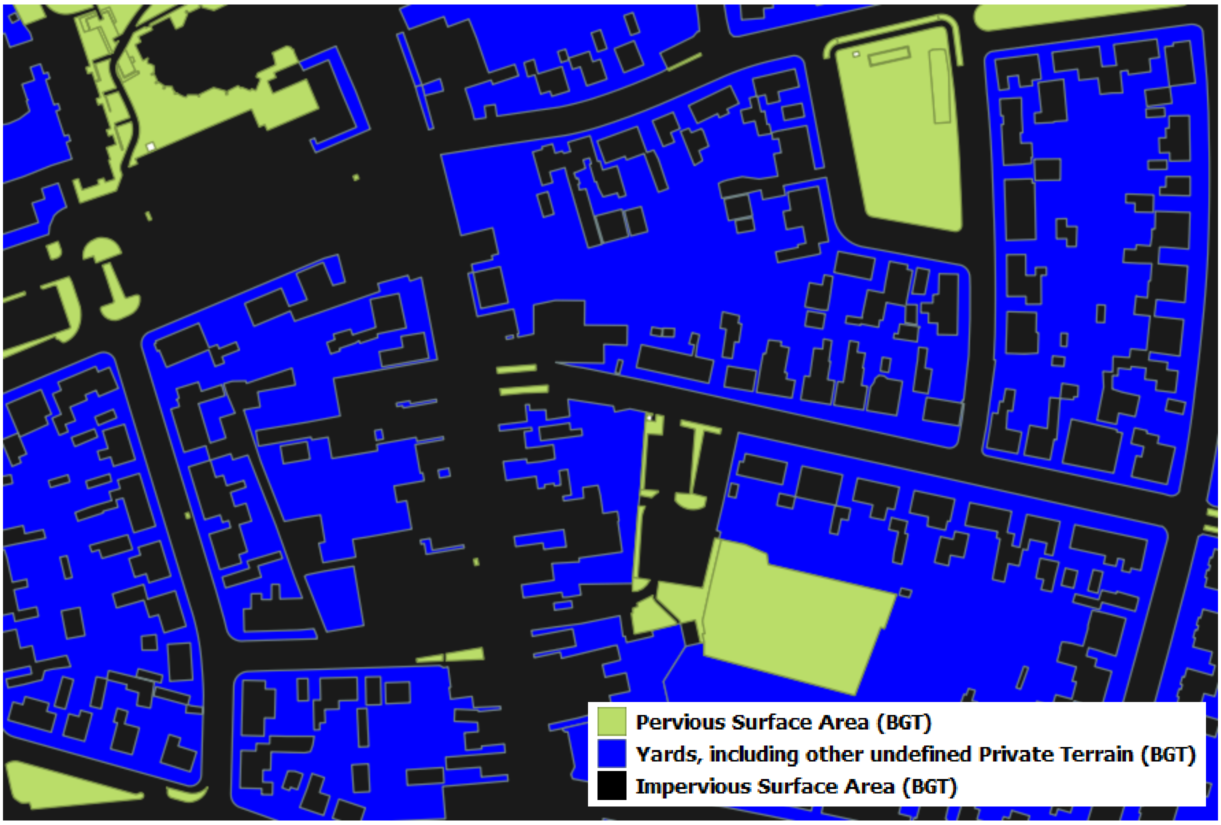


Figure 36 – Classification according to the BGT. All impervious surface areas (such as roads, houses or pavement) are combined. All pervious surface area (parks, grass) is combined. The yards and undefined private terrains are depicted in blue.

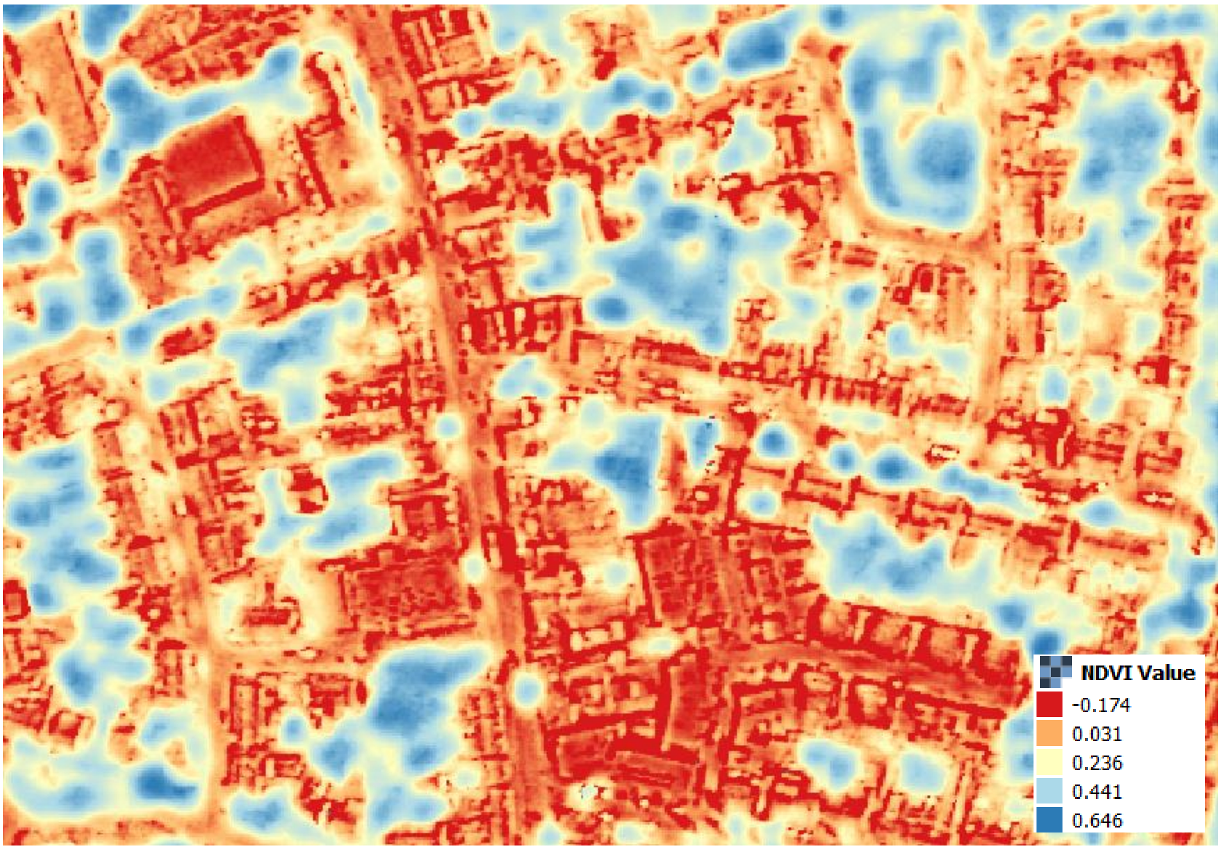


Figure 37 – NDVI classification of the whole neighbourhood

Figure 38 – Classification of the yards and undefined private terrain

E. Input Parameters InfoWorks ICM

This Appendix gives an overview of the input parameters used in InfoWorks ICM.

Category	Parameter	Value or Selected Option	Unit
Nodes	*Benching Method*	*Full Benching*	-
Conduits	*Sediment Depth*	*0*	*mm*
	Headloss Type	*NONE*	-
	Settlement Efficiency	*0*	*%*
	Roughness Type	*Colebrook-White*	-
	Bottom Roughness Concrete	*3.000*	*mm*
	Top Roughness Concrete	*3.000*	*mm*
	Bottom Roughness PVC	*0.400*	*mm*
	Top Roughness PVC	*0.400*	*mm*
	Base Height	*0*	*mm*
	Base Infiltration Loss	*0*	*mm/hr*
	Side Infiltration Loss	*0*	*mm/hr*
Sub Catchments	*Base Flow*	*0*	m^3/s
	Additional Foul Flow	*0*	m^3/s
	Unit Hydrograph Definition	*User-Tp-Tb*	-
	Baseflow Calculation	*PDM*	-
	Soil Moisture Deficit	*PDM*	-
2D Zone	*Maximum Triangle Area*	*2.000*	m^2
	Minimum Element Area	*1.000*	m^2
	Minimum Angle	*25.00*	*degrees*
	Maximum Height Variation	*1.000*	*m*
	Manning's Roughness	*0.0125*	-
	Boundary Condition	*Normal Boundary*	-
Simulation General	*Rainfall Smoothing*	*Not Applied*	-
	Dry Weather Flow Mode Multiplier	*32*	*x*
1D Simulation	*Timestep*	*60*	*s*
	Results Timestep	*5*	*x*
	Gauge Timestep	*1*	*x*
2D Simulation	*Timestep*	*1*	*s*
	Results Timestep	*300*	*x*
	Gauge Timestep	*1*	*x*
	Timestep Stability Control	*0.95*	-
	Maximum Velocity	*10*	*m/s*
	Theta	*0.9*	-
	Inundation Mapping Depth Threshold	*0.01*	*m*
	Damage Timestep Multiplier	*1*	*x*
	State Search Radius	*100*	*m*
	State Power Parameter	*2*	-
	Depth Tolerance	*0.001*	*m*
	Momentum Tolerance	*0.001*	*m*
	Velocity Tolerance	*0*	*m/s*
	Link 1D and 2D Calculations	*Link At Minor Time Step*	-

Table 22 – Overview of the used input parameters for InfoWorks ICM. Most parameters are left at their default value.

F. Dutch Design Storms

In the process of choosing a storm event, a distinction should be made whether statistics are applied on the storm event or on the flooding of the urban water system ‘afterwards’. The distinction should be made because sewer systems show a strong non-linear behaviour (Vaes & Berlamont, 1996). This means that a storm event with a certain return period does not necessarily result in a flood event with the same return period. Sewer systems may be more vulnerable to storms with a specific shape or duration.

When it is chosen to perform statistics on the flood event, usually a long time series is modelled, preferably multiple decades for evaluating urban flooding. This, however, requires a constant time series of good quality and high sampling frequency. Furthermore, this would imply a calculation time of days for calculations with a 2D surface model. Also, policy makers usually prefer to set a norm, and setting a (universal) norm is easier by choosing a certain storm event. It is therefore common to perform statistics on the storm event. Usually, a precipitation series is selected from local weather stations. The data is then processed into design storms. In the Netherlands, ten ‘design storms’ with a predefined shape and return period were created based on a 15-minute dataset from the period 1955-1979 (Stichting RIONED, 2004). Internationally, composite storms, based on intensity-duration-frequency relations, are often used. The concept is applied in (Vaes & Berlamont, 1996).

In this research, the original models are tested with Dutch Storm 8 (Bui 08). This storm represents a return period of two years and is a design standard for most newly designed sewer systems in the Netherlands. The storm is therefore a good measure of locations that do not entirely perform according to the design standard. Recently, renewed rainfall characteristics for storm events with a short duration were presented (STOWA, 2018). The characteristics were based on 10-minute data from 31 automatic rain gauges in the period 2003-2016. The hourly rainfall for a storm event with a return period of two years was calculated at 20.0 mm. This number corresponds well with the 19.8 mm of rainfall from Bui 08.

Case-specific storm events are preferably used when a flood situation is assessed. This means that for a certain storm event, rain data is available from local rain gauges and the flooded locations corresponding to the storm are known. The storm data can then be used to test the performance of the modelled system under the same conditions.

If no specific flood event is linked to the available information on areas prone to flooding, Dutch Design Storm 9 (Bui 09) is used. This storm roughly represents a return period of five years (Figure 40). As this storm is more intense than the design standard for most Dutch sewer systems, the storm event should be suitable for recognizing the vulnerable areas. In the new rainfall characteristics by STOWA, the hourly rainfall for a storm event with a return period of five years was calculated at 25.8 mm (STOWA, 2018). This number is somewhat lower than the rainfall with Bui 09 (29.4 mm).

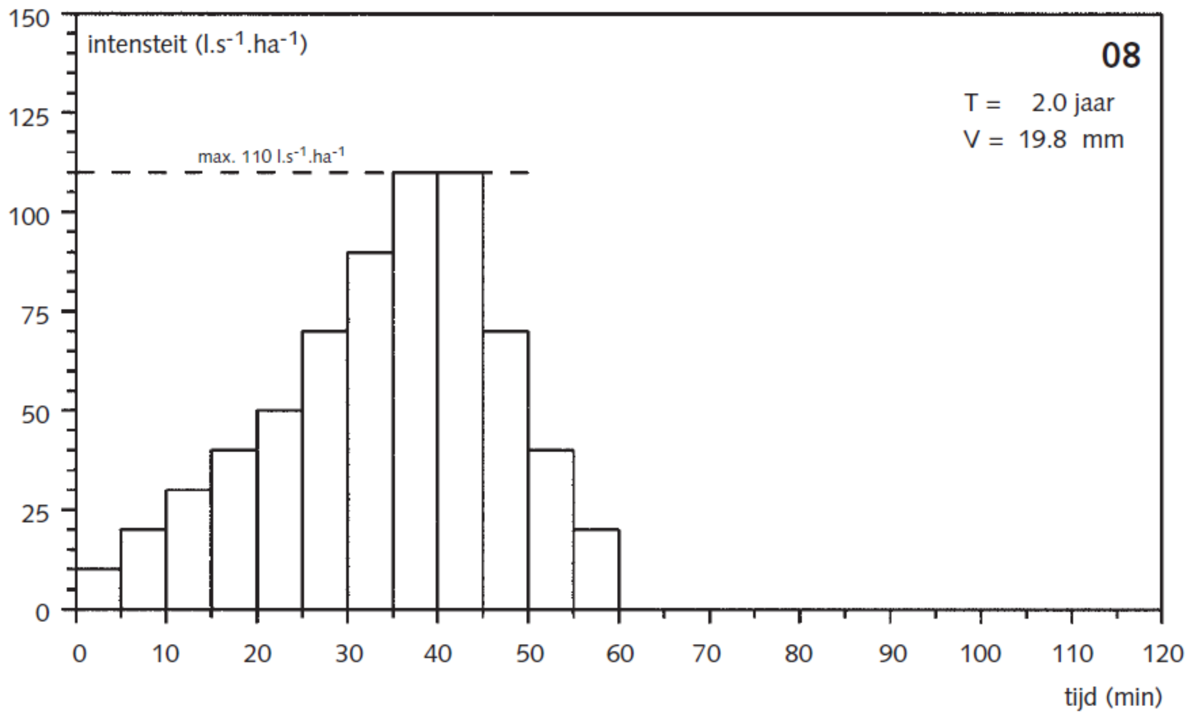


Figure 39 – Dutch Design Storm 8 or Bui 08. This storm represents a return period (T) of two years (Stichting RIONED, 2004)

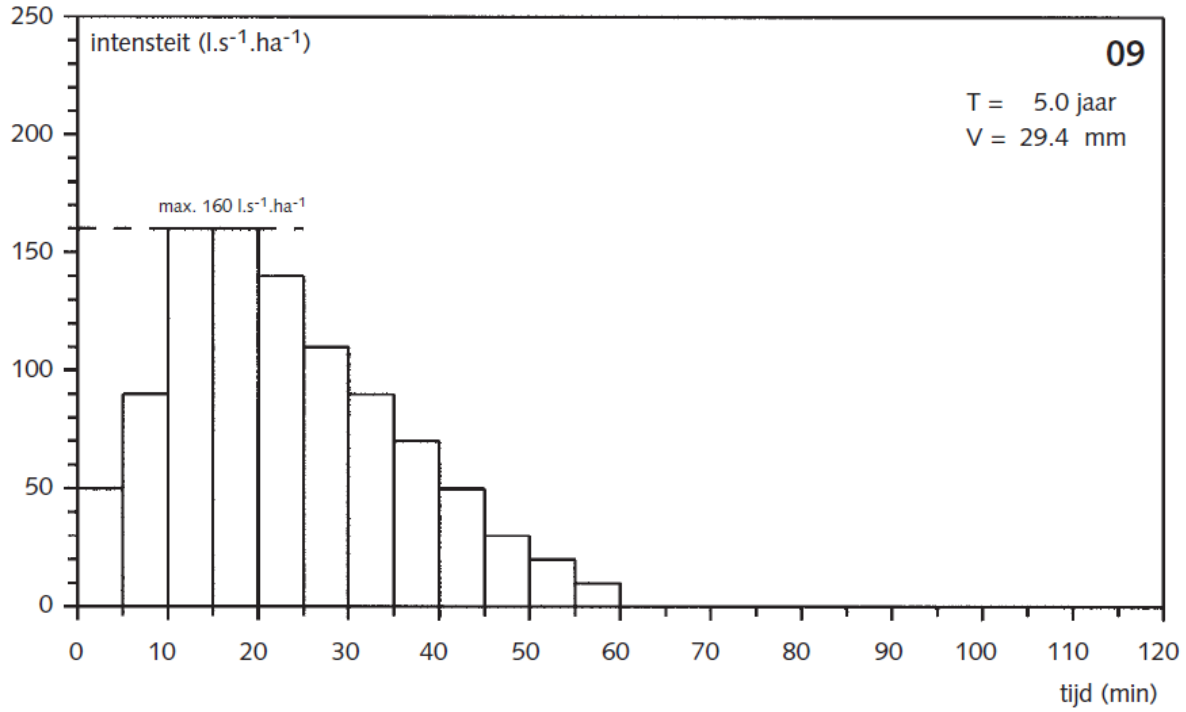


Figure 40 – Dutch Design Storm 9 or Bui 09. This storm represents a return period (T) of five years (Stichting RIONED, 2004)

G. Results of WOLK-model for Loenen

The municipality of Apeldoorn provided the results of a WOLK-model that was created to point out vulnerable areas. WOLK is a modelling software package that uses a Digital Elevation Model (DEM) to assess the flow paths of rain water. The model does not include equations for shallow water flow and should therefore not be classified as a hydrodynamic model (STOWA, 2017). The calculated water volumes are entirely based on the filling up of topographic depressions in the landscape. Figure 41 shows the results of the analysis.

As this research does not compare modelling packages and the models in this research generally take more processed into account, the WOLK modelling results are not used in this research. The model does provide an insight into locations that are vulnerable to pluvial flooding because of the topography. In theory, the results of the 2D_St model should mainly coincide with the flood prone locations, as both the WOLK model and 2D_St model focus on topographic depressions.

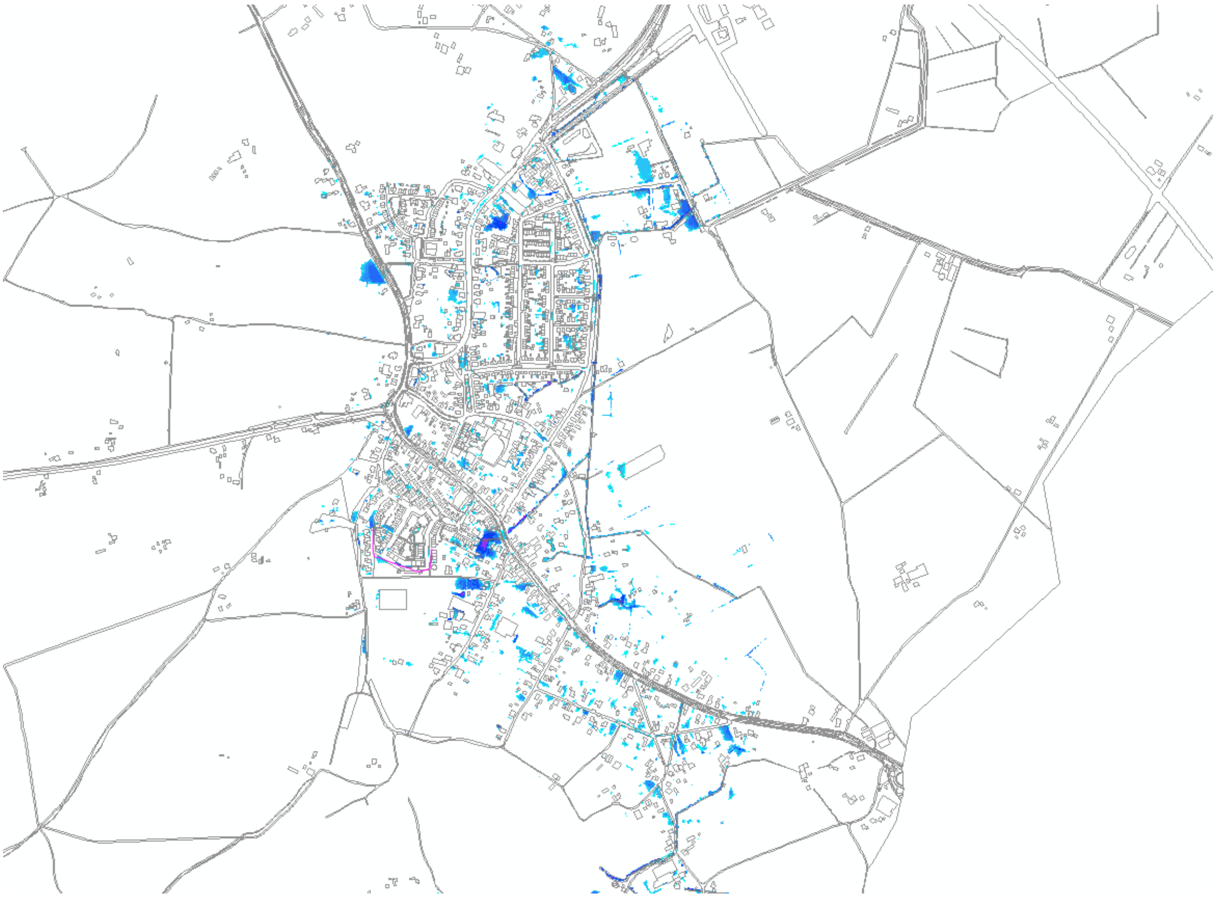


Figure 41 – Results of a WOLK model for test case Loenen, provided by the municipality of Apeldoorn

H. A Brief History of Sewers

The Etruscans, who settled in Italy in the fifth and sixth century before Christ, were amongst the first civilizations to have sewer systems to discharge waste water, as well as storm water. It is therefore likely that the sewer system underneath ancient Rome, *Cloaca Maxima*, was designed by Etruscan civil engineers (Bosselaar, 1940). By making use of gravity flow, the water was diverted from the city to the surrounding water system. With the downfall of the Roman Empire, progress in the field of sewer discharge came to a halt.

In the Middle Ages, most waste water was dumped directly into the nearest river or directed to the river by open gutters (Newman, 2001). This led to large scale pollution of rivers and epidemics, especially in larger cities. In the nineteenth century, when progression in the field of medicine revealed the need for proper sanitation, a new period of development in the construction of sewers started. In London, at that time the largest city in the world, experiments conducted by John Snow (1849-1953) on the relation between cholera and the source of drinking water played an important role in the understanding of water-borne diseases (see (Snow, 1855)). It is however not the outbreak of cholera, but stench that sparked the creation of the sewer system in that same city (London). In 1858, during a hot summer (known as the Great Stink), the stench from the extremely polluted river Themes led to the decision to start the design of an underground sewer system (Mann, 2016). Joseph Bazalgette was the engineer that designed the sewer system that transported waste water and storm water out of the city of London (see (Bazalgette & Forrest, 1865)). To prevent polluted water from entering the houses in the case of intense rainfall, overflow structures were created. Over the second half of the nineteenth century, more European cities built sewer systems to prevent stench and further deterioration of the water quality in city canals.

In the Netherlands, The Hague was amongst the first cities in the Netherlands to build a modern, underground sewer system. In 1889, a project to increase circulation through the canals only had a minor effect on the stench and water quality. It was therefore decided in 1895 that the discharge of waste water into the canals and surrounding polders should be stopped, after which the construction of a sewer system started (Bosselaar, 1940). Over the course of the nineteenth and twentieth century, all Dutch cities were connected to a sewer system. Nowadays, only at locations where the construction of a sewer system would be unrealistically expensive, sceptic tanks are still in use.

In the twentieth century, the quality of the receiving water body played an increasing role. Most sewer systems discharged waste water (and storm water) directly into a river or sea without further treatment. The enormous amount of biodegradable material caused oxygen depletion in the receiving surface water, leading to fish mortality and a water quality unsafe for human recreation. One of the earliest forms of biological waste water treatment was in the form of an irrigation field (around the 1870's): waste water was pumped onto a large field from where it could infiltrate. Porous drainage pipes caught the water and discharged it onto the nearest surface water (Wiesmann, et al., 2006). Over the course of the twentieth century, waste water treatment plants (WWTP's) in their modern form took shape and where built on a large scale.

I. Flood Maps

This Appendix shows flood maps where all flooded areas with at least 5 cm of water. Higher water depths are represented by a darker shade of blue. Only one model is depicted in a figure.

Ulvenhout

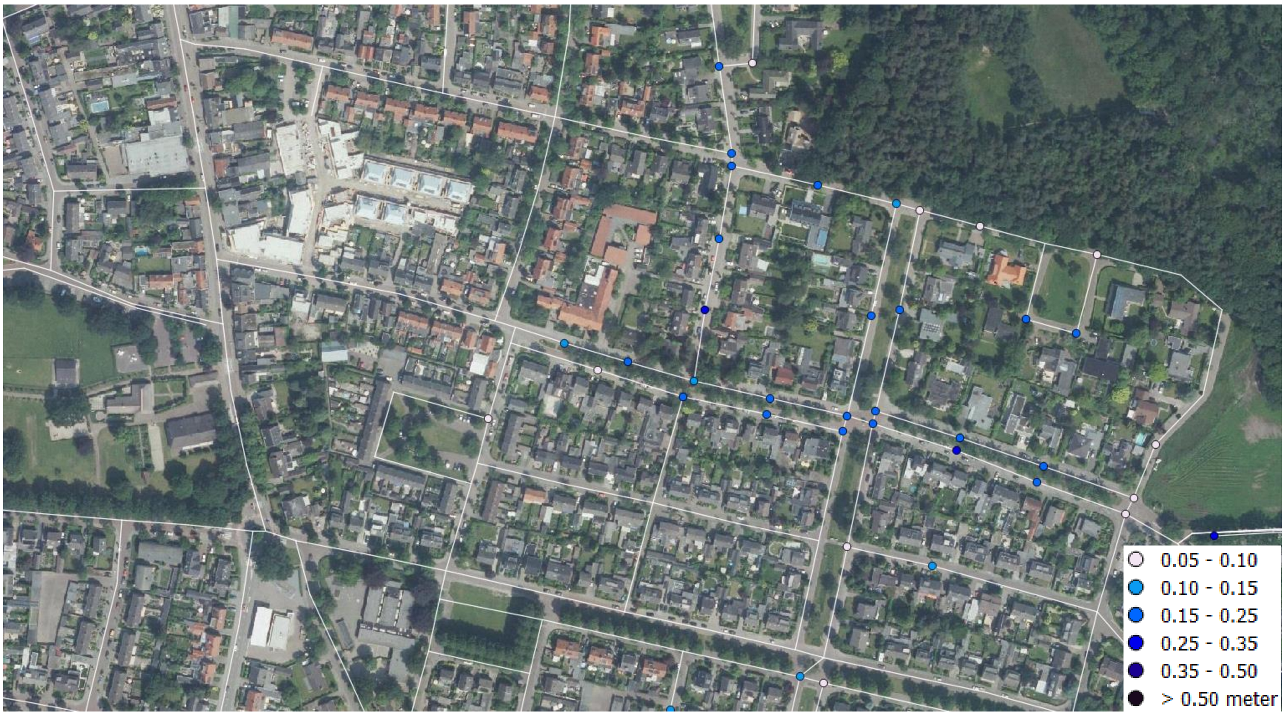


Figure 42 – Flood map Ulvenhout – Ulvenhout 2016 storm event – 1D_St model

Figure 43 – Flood map Ulvenhout – Ulvenhout 2016 storm event – 1D/2D_St model

Figure 44 – Flood map Ulvenhout – Ulvenhout 2016 storm event – 1D/2D_P model

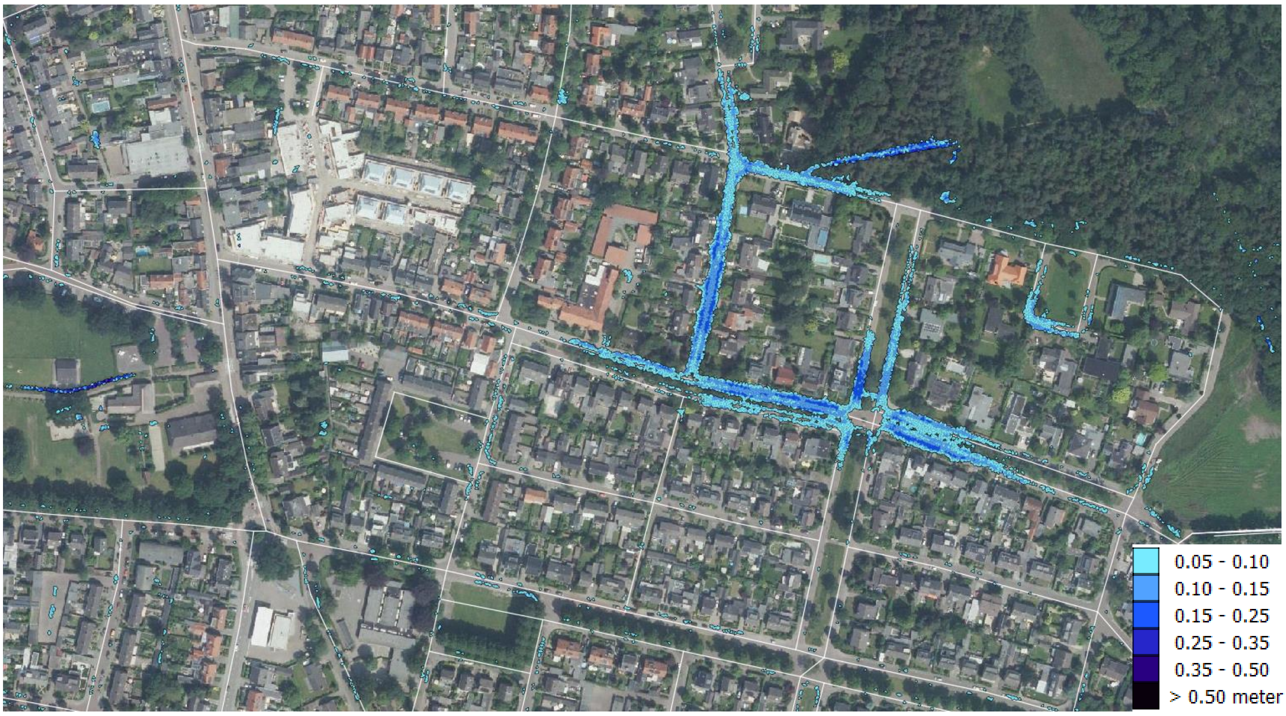


Figure 45 – Flood map Ulvenhout – Ulvenhout 2016 storm event – 1D/2D+_H model

Figure 46 – Flood map Ulvenhout – Ulvenhout 2016 storm event – 1D/2D+_Y model

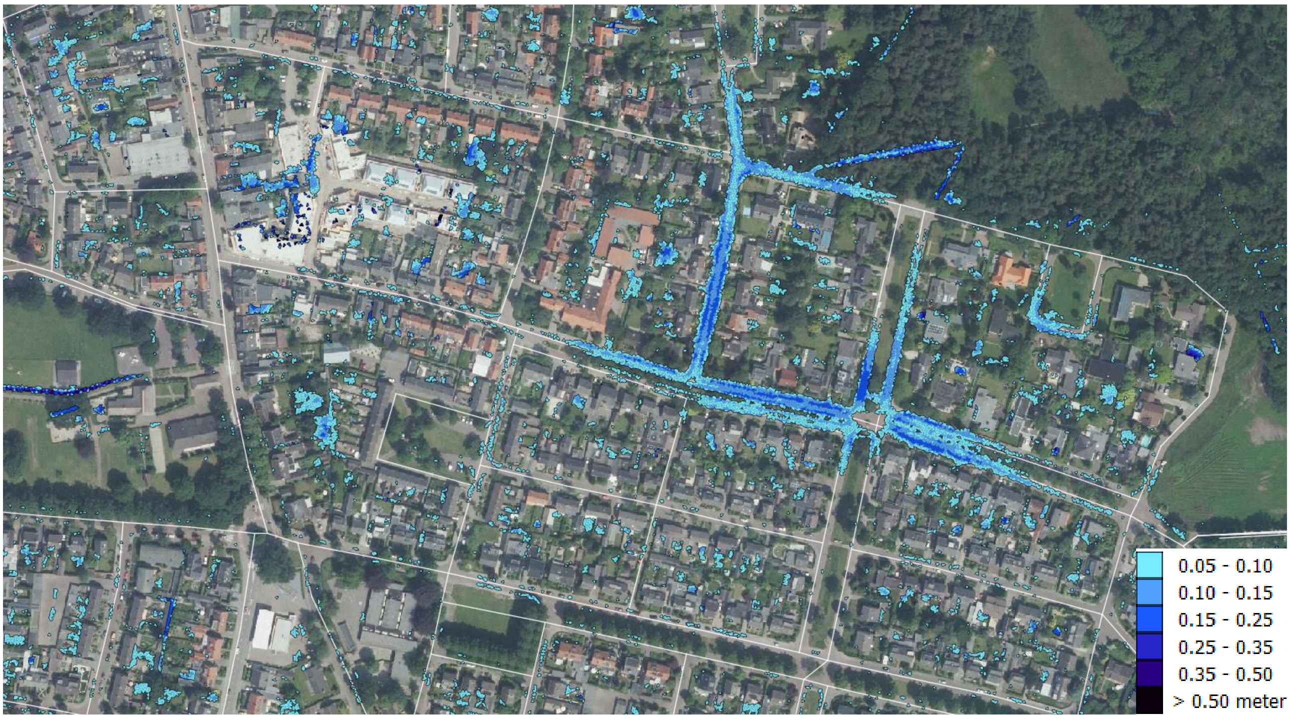


Figure 47 – Flood map Ulvenhout – Ulvenhout 2016 storm event – 1D/2D+_HY model

Figure 48 – Flood map Ulvenhout – Ulvenhout 2016 storm event – 2D_St model

Figure 49 – Flood map Ulvenhout – Ulvenhout 2016 storm event – 2D_I model

Tuindorp

Figure 50 – Flood map Tuindorp – Tuindorp 2013 storm event – 1D_St model

Figure 51 – Flood map Tuindorp – Tuindorp 2013 storm event – 1D/2D_St model

Figure 52 – Flood map Tuindorp – Tuindorp 2013 storm event – 1D/2D_P model

Figure 53 – Flood map Tuindorp – Tuindorp 2013 storm event – 1D/2D+_H model

Figure 54 – Flood map Tuindorp – Tuindorp 2013 storm event – 1D/2D+_Y model

Figure 55 – Flood map Tuindorp – Tuindorp 2013 storm event – 1D/2D+_HY model

Figure 56 – Flood map Tuindorp – Tuindorp 2013 storm event – 2D_St model

Figure 57 – Flood map Tuindorp – Tuindorp 2013 storm event – 2D_I model

Loenen

Figure 58 – Flood map Loenen – Bui 09 – 1D_St model

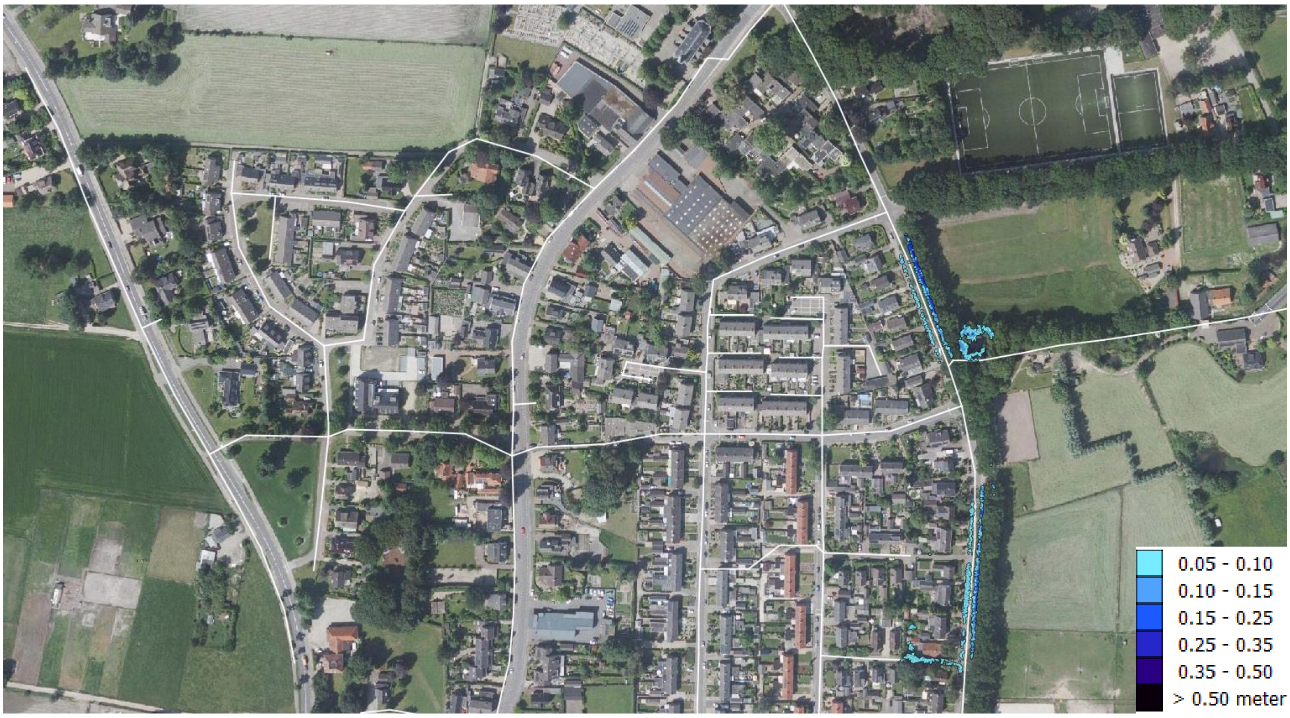


Figure 59 – Flood map Loenen – Bui 09 – 1D/2D_St model

Figure 60– Flood map Loenen – Bui 09 – 1D/2D_P model

Figure 61– Flood map Loenen – Bui 09 – 1D/2D+_H model

Figure 62– Flood map Loenen – Bui 09 – 1D/2D+_Y model

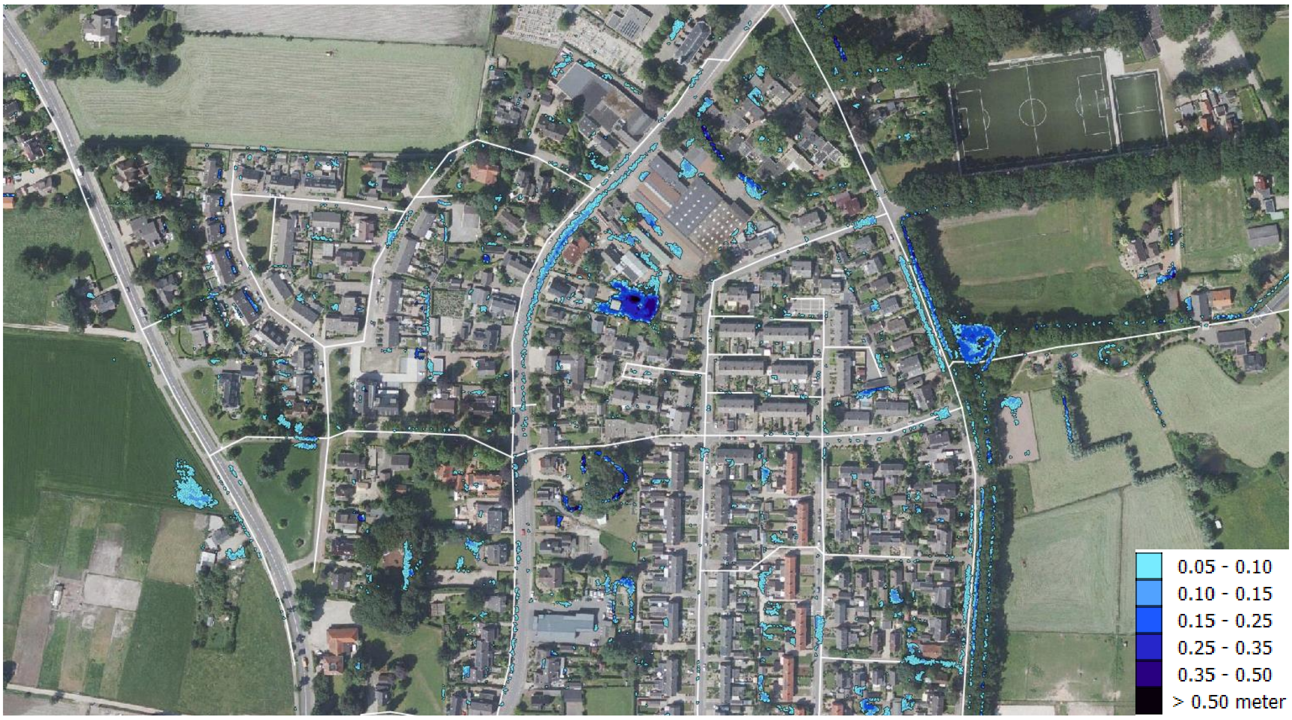


Figure 63– Flood map Loenen – Bui 09 – 1D/2D+_HY model

Figure 64– Flood map Loenen – Bui 09 – 2D_St model

Figure 65– Flood map Loenen – Bui 09 – 2D_I model

J. Flood Contour Maps

This Appendix shows flood contour maps. Water depths of more than 5 cm are taken into account. All models are depicted in the same figure.

Figure 66 – Flood Contour Map Ulvenhout – Ulvenhout 2016 storm event

Figure 67 – Flood Contour Map Tuindorp – Tuindorp 2013 storm event

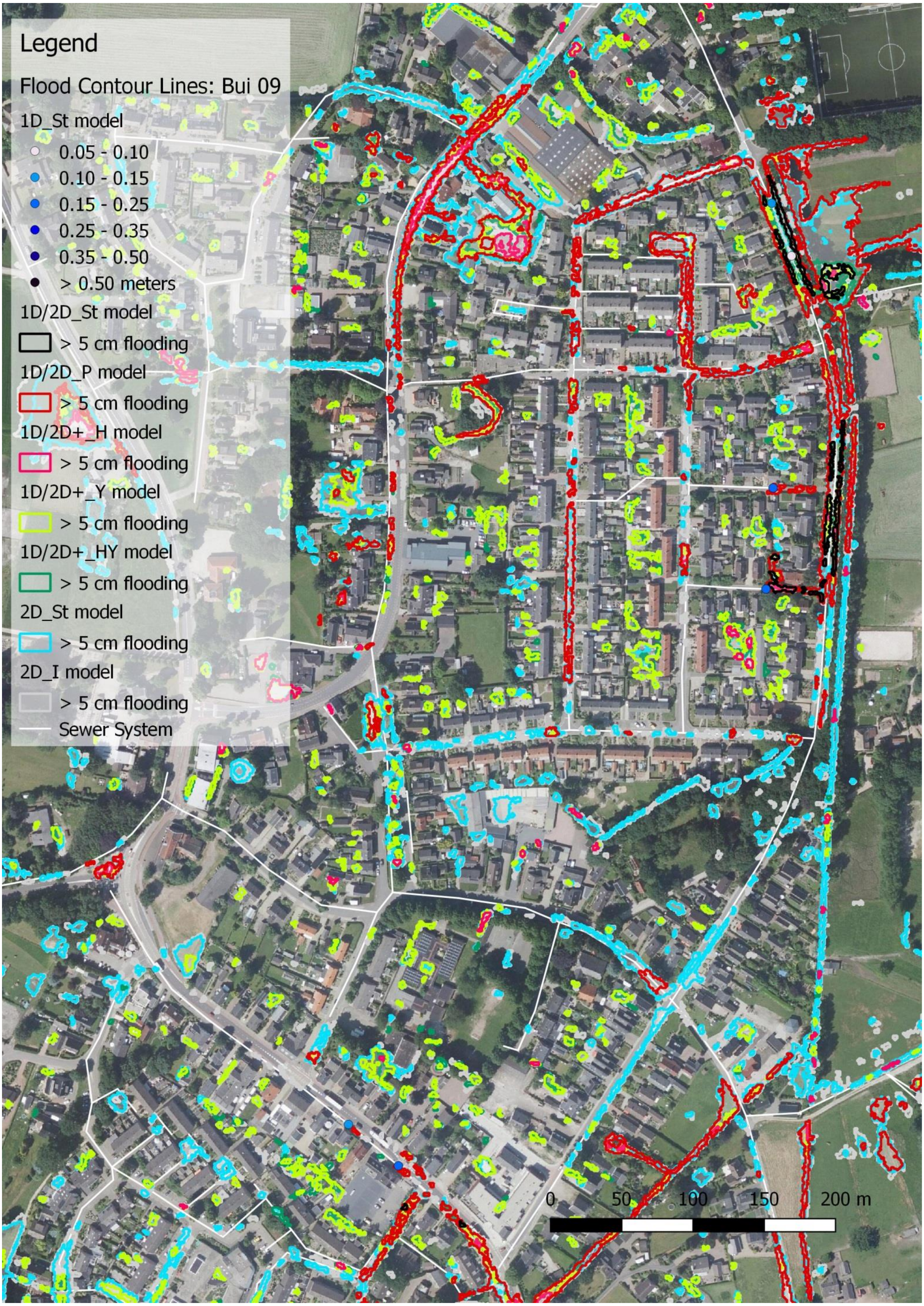


Figure 68 – Flood Contour Map Loenen – Bui 09

K. Flood Difference Maps

This Appendix shows flood difference maps for the four additional aspects that were presented in § 4.3 and discussed in § 8.1. The maps show the difference in water depth between two models.

Ulvenhout

Figure 69 – 1D/2D_St vs. 1D/2D_P model – Ulvenhout – Ulvenhout 2016 storm event

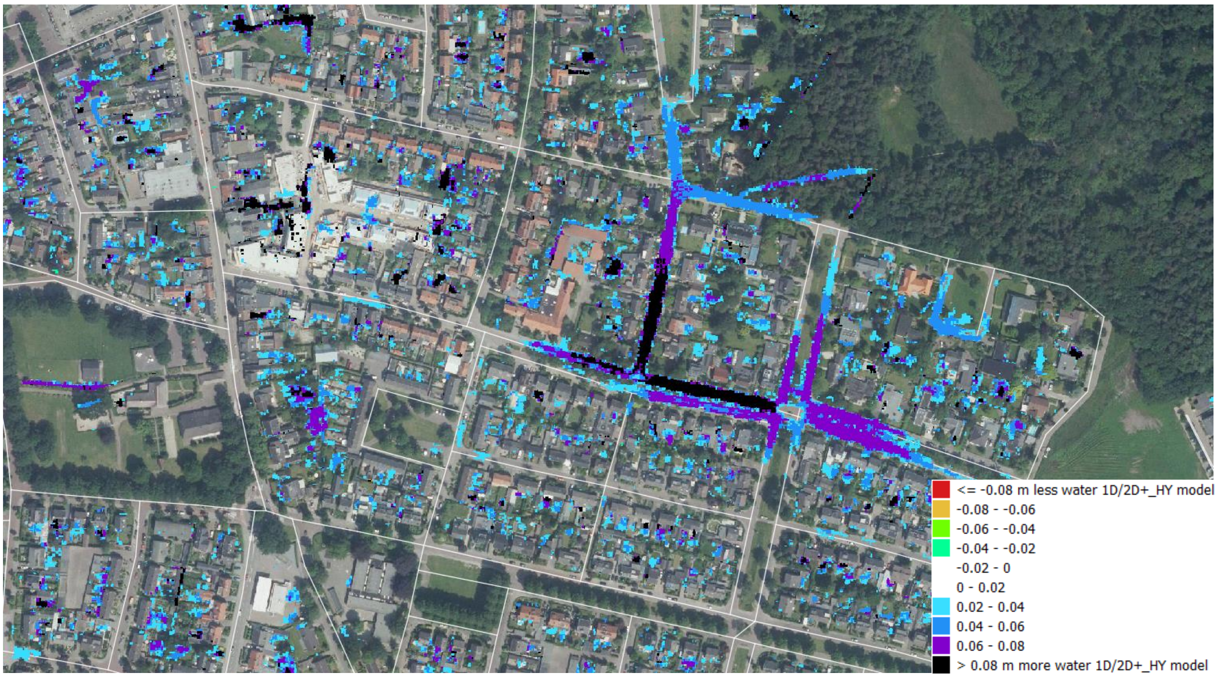


Figure 70 – 1D/2D+_H vs. 1D/2D+_HY model – Ulvenhout – Ulvenhout 2016 storm event

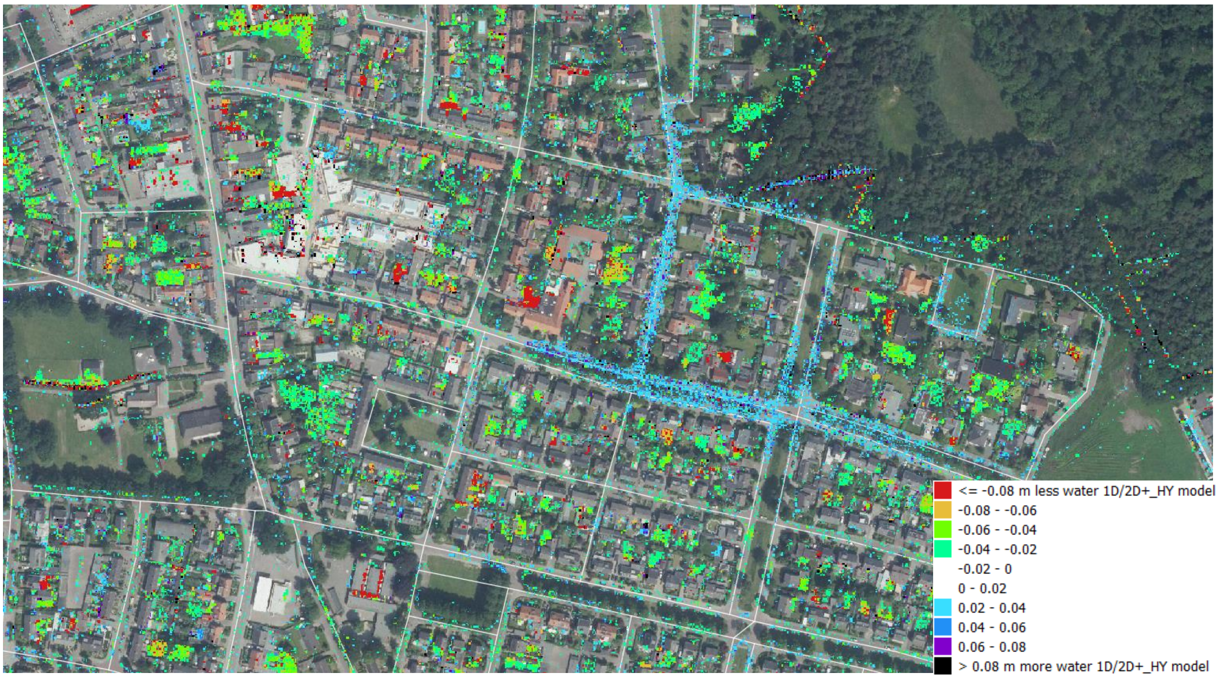


Figure 71 – 1D/2D+_H vs. 1D/2D+_Y model – Ulvenhout – Ulvenhout 2016 storm event

Figure 72 – 2D_St vs. 2D_I model – Ulvenhout – Ulvenhout 2016 storm event

Tuindorp

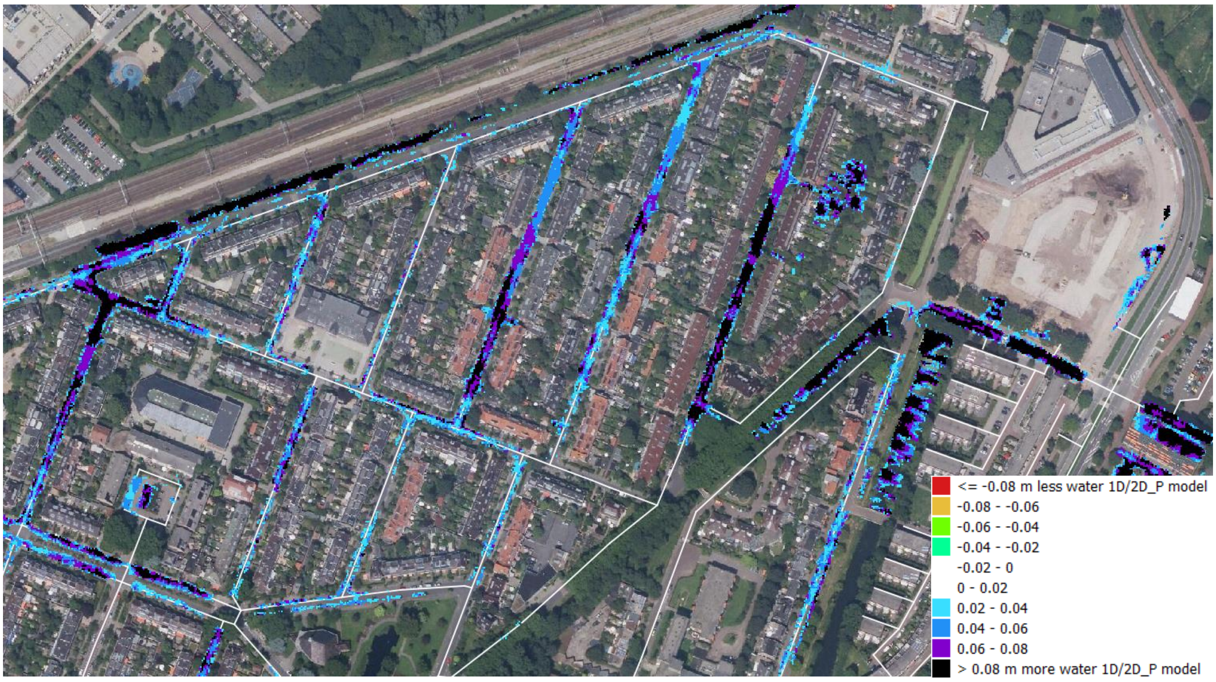


Figure 73 – 1D/2D_St vs. 1D/2D_P model – Tuindorp – Tuindorp 2013 storm event

Figure 74 – 1D/2D+_H vs. 1D/2D+_HY model – Tuindorp – Tuindorp 2013 storm event

Figure 75 – 1D/2D+_H vs. 1D/2D+_Y model – Tuindorp – Tuindorp 2013 storm event

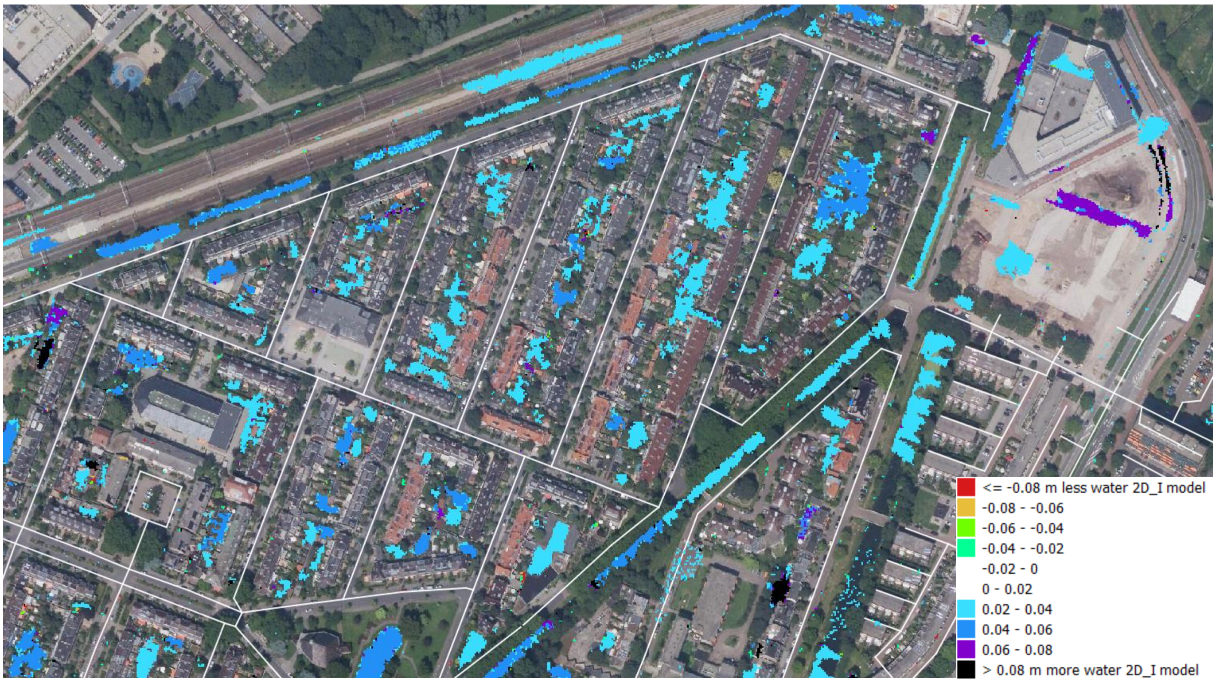


Figure 76 – 2D_St vs. 2D_I model – Tuindorp – Tuindorp 2013 storm event

Loenen

Figure 77 – 1D/2D_St vs. 1D/2D_P model – Loenen – Bui 09

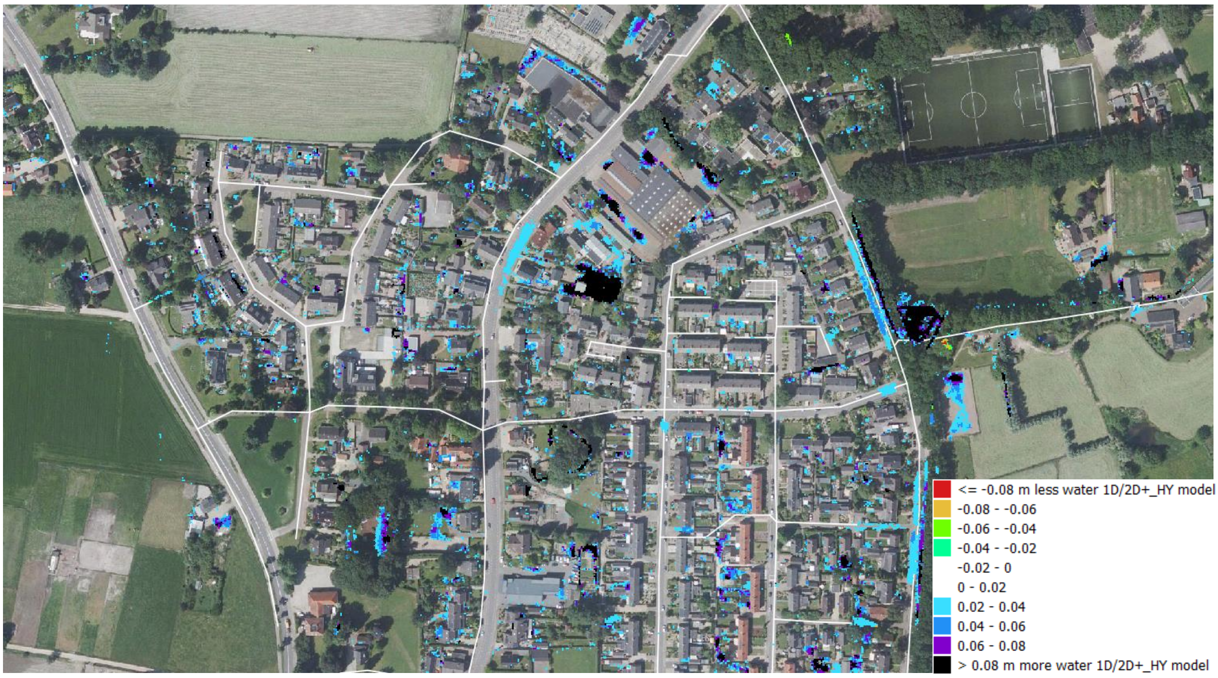


Figure 78 – 1D/2D+_H vs. 1D/2D+_HY model – Loenen – Bui 09

Figure 79 – 1D/2D+_H vs. 1D/2D+_Y model – Loenen – Bui 09

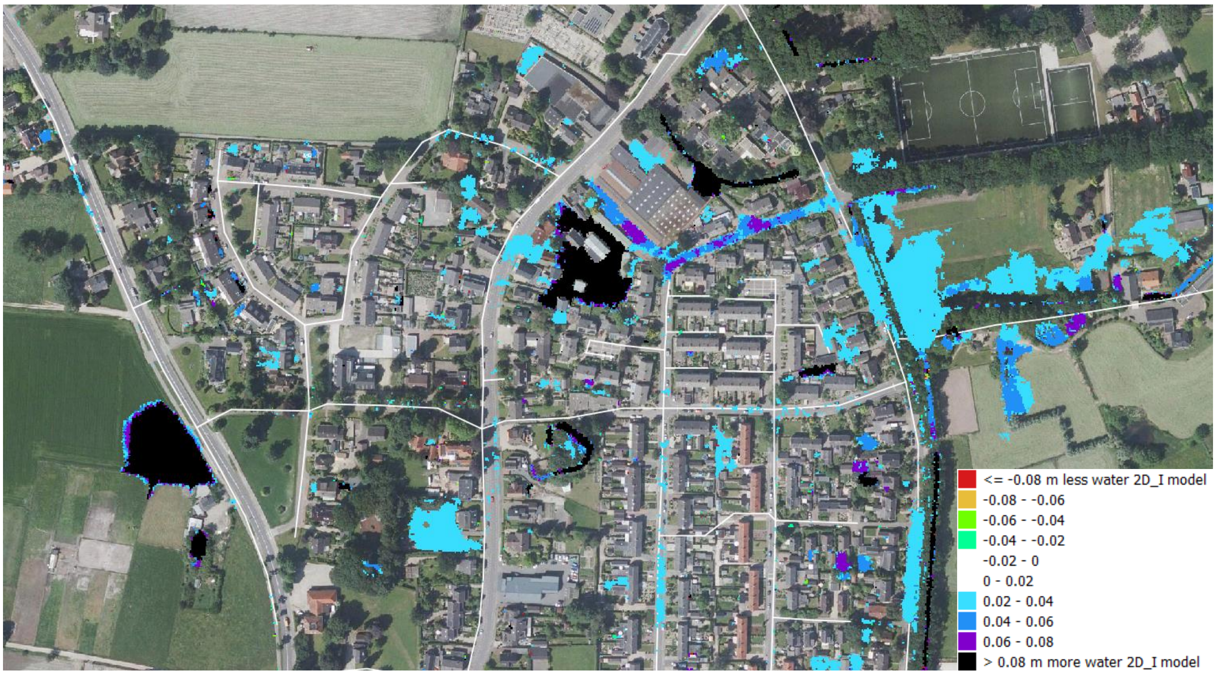


Figure 80 – 2D_St vs. 2D_I model – Loenen – Bui 09

L. Information on Locations corresponding to monitored Manholes in Test 2

Figure 28 - Figure 31 show the monitoring results of two monitored manholes: the monitored manhole in Ulvenhout and manhole/node 90037 in Tuindorp. The appendix provides background information on the locations of these manholes.

Monitoring Location Ulvenhout

The monitored manhole in Ulvenhout is a pump sump situated in the southern part of Ulvenhout. This implies that during dry weather flow, a saw tooth pattern is visible in the monitoring of water levels. The pump sump fills up with waste water until the level at which the pump turns on. The pump then empties the sump up to the level where the pump switches off. In the model, the manhole drains a small 'internal' combined sewer system. The pump is used to pump the waste water and storm water into the larger system up north. A weir (combined sewer overflow) is situated close to the pump to discharge water when street flooding is otherwise imminent.

Upon further investigation, a deviation was observed between the provided sewer model and the municipal database (stored in a software application called *Kikker*). In the database, a small conduit was present (200 mm) between the smaller 'internal' combined system and the main sewer system. Figure 81 shows the location of this deviation and the other important elements.

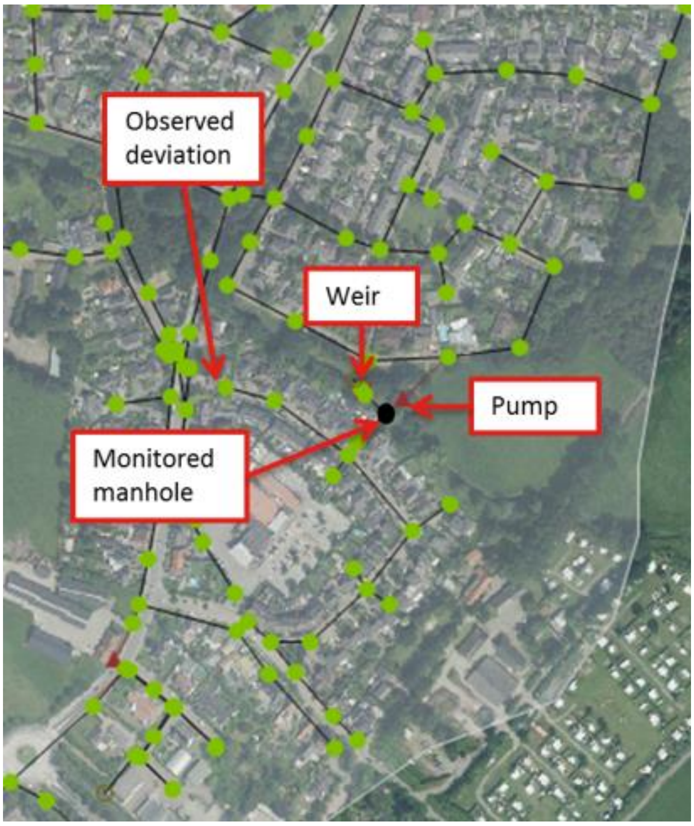


Figure 81 – Monitored location in Ulvenhout. The weir is a combined sewer overflow (CSO). At the 'observed deviation', an additional manhole was found in another municipal database.

Node 90037 in Tuindorp

In Tuindorp, monitoring results from one manhole (node 90037) are shown in Figure 30 and Figure 31. The monitored manhole is situated in the western part of Tuindorp. The connected conduits transport the waste water and storm water out of the western part of Tuindorp to a larger conveying conduit.

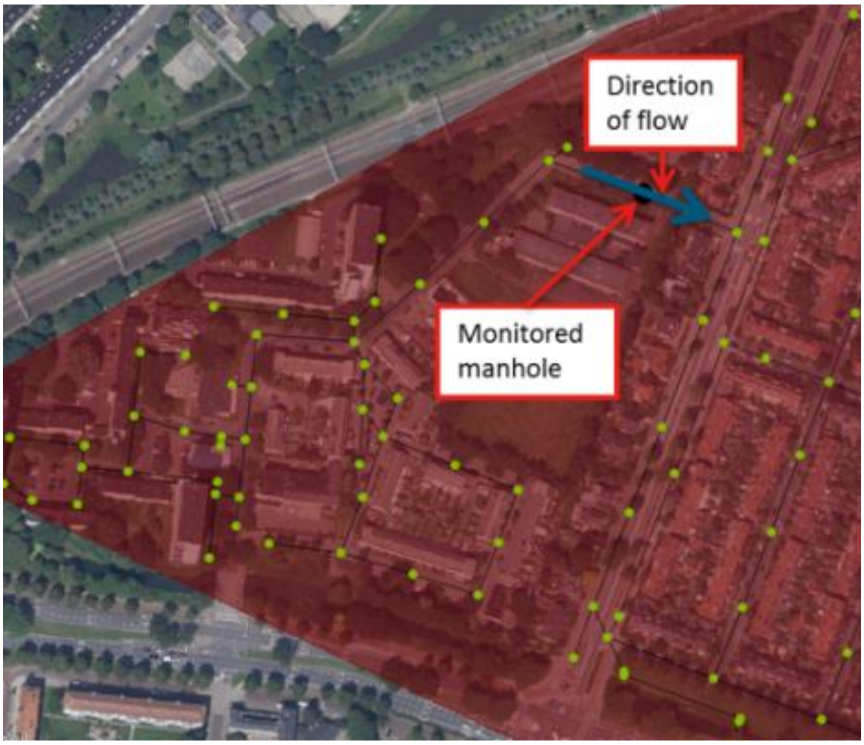


Figure 82 – Monitored manhole 90037 in Tuindorp

M. Tuindorp Monitoring Data

In the case of Tuindorp, the provided 1D_St model shows the same deviations from the measurements in a large amount of manholes:

- The initial water level of the monitoring data is higher and more constant than the initial water level in the models.
- The peak water level of the monitoring data is lower than the peak water level in the models, especially during the first peak.

Figure 82 and Figure 83 show two examples of these two observed deviations.

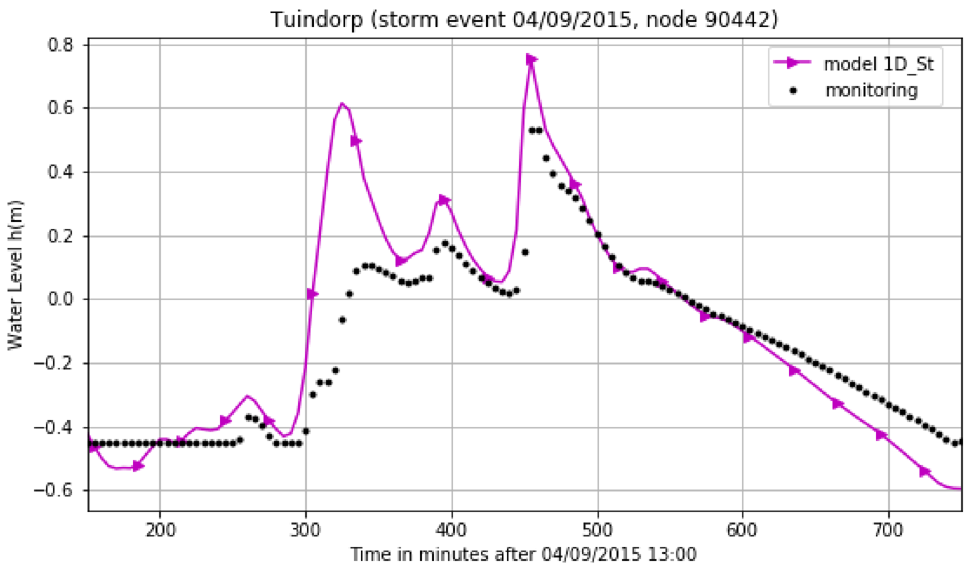


Figure 83 – Measured and modelled water level in Tuindorp, node 90442

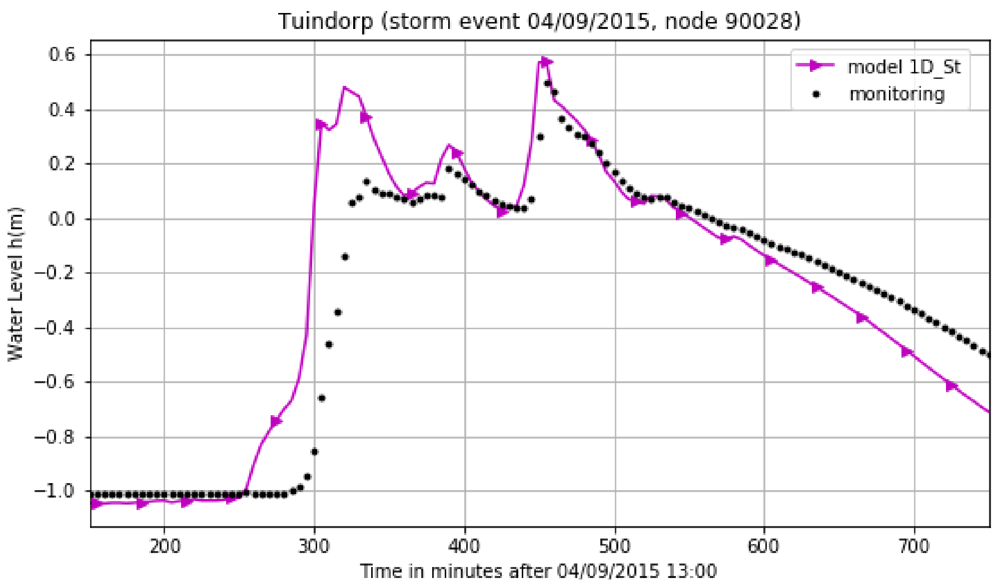


Figure 84 – Measured and modelled water level in Tuindorp, node 90028

N. Sensitivity of the Infiltration Capacity: a Test Case

Multiple times, it has been stated that the choice for parameters significantly influence the test results. Especially infiltration capacity of pervious surface types has a large range in parameter values, as urban soils are complex and non-uniform. In this Appendix, the influence of the choice for infiltration parameters is investigated by comparing four setups with a different infiltration capacity for pervious surface types. This Appendix should not be considered as a calibration of the model, as the altering of multiple parameters could give the same result.

Chosen Setup and Test Case

For the infiltration, the Horton infiltration model is used (see Equation 1). The Horton model uses three basic parameters (except time): f_0, f_c and k. In InfoWorks ICM, another parameter is added for the recovery of the infiltration capacity. Two of these parameters are altered in this case: f_0 (initial infiltration capacity) and f_c (infiltration capacity for the saturated soil).

The storm event in Ulvenhout in 2016 was the only storm event in this research where both observed data and monitoring data were available. Therefore, this storm event is selected. From the eight models, the 1D/2D+_H model is selected. The model is a 1D/2D+ model where roofs are modelled with semi-distributed runoff. There are three reasons for choosing this model:

- It is the model with the least variation in infiltration surfaces: only pervious and impervious surfaces are distinguished.
- The model scores lowest in test 1 for the 2016 storm event in Ulvenhout, both in the default configuration (68%) and the average of all configurations (66%).
- The results from both test showed that the model underestimates the amount of street flooding. The scores for true negatives (39% in the default setup) were higher than for true positives (97% in the default setup) and the β-term of the KGE was below zero (0.953).

The default parameter set for the model is based on (US Environmental Protection Agency, 1999)(see Appendix C). In the setups, both the parameter f_0 and f_c are lowered by 30%. Lowering the infiltration capacity will increase the amount of street flooding, and scores should therefore increase in theory. The parameters sets that are tested are shown in Table 23. Both test 1 and test 2 are executed. For test 1, the default test parameter settings are used: a grid size of 30 m, depth threshold of 10 cm and areal threshold of 5 m².

Parameter Values		f_0: default	f_0: - 30%
f_c: default	f_0	*127 mm/h*	*88.9 mm/h*
	f_c	*12.7 mm/h*	*12.7 mm/h*
f_c: - 30%	f_0	*127 mm/h*	*88.9 mm/h*
	f_c	*8.89 mm/h*	*8.89 mm/h*

Table 23 – Parameter sets that are tested in this Appendix

Results

Table 24 and Table 25 show the results for the four different setups. The scores for the default setup are equal to the scores given for the 1D/2D+_H model in §7.1. The hypothesis that both parameters influence the amount of street flooding is confirmed by the results. The scores for true positives significantly improve by lowering parameter values. By lowering both f_0 and f_c, the score improves from 39% to 83%. The score for false negatives slightly decreases (97% to 91% when both parameter values are lowered). The combined score for test 1 increases from 68% to 87% and the scores for true negatives and true positives are in better balance. The better balance (83% and 91%) implies that the underestimation of street flooding by the 1D/2D+_H model is partly cancelled by lowering the parameter values. Test 2 confirms the statements made. The RMSE lowers and the β-term increases. However, differences are less significant compared to test 1.

Test 1 Scores		f_0: default	f_0: - 30%
f_c: default	*TP*	*39%*	*53%*
	TN	*97%*	*97%*
	Both	*68%*	*75%*
f_c: - 30%	*TP*	*78%*	*83%*
	TN	*94%*	*91%*
	Both	*86%*	*87%*

Table 24 – Results of test 1 for the four setups. 'TP' stands for the score, considering only true positives. 'TN' stands for the score, considering only true negatives. 'Both' implies the combined score.

Test 2 scores		f_0: default	f_0: - 30%
f_c: default	*RMSE*	*26.34 cm*	*26.32 cm*
	β	*0.9537*	*0.9538*
f_c: - 30%	*RMSE*	*26.26 cm*	*26.13 cm*
	β	*0.9540*	*0.9546*

Table 25 – Results of test 2 for the four setups. 'RMSE' stands for the root mean squraed error. β is the bias-ratiom which is part of the KGE index and a simple measure for underestimation or overestimation.

Discussion

The difference in scores for test 1 (by changing the infiltration parameter values) is larger than the difference for test 2. As explanation for this is that in test 1, data is used from two areas in Ulvenhout (see Figure 15), while test 2 is performed on a single manhole in the south of Ulvenhout (see Appendix L). The area around the location used for test 2 is mainly impervious and a nearby weir is likely to level off most of the difference in water level.

For the test case selected, lowering f_c has a larger influence on the results than lowering f_0 by the same percentage. A possible explanation for this is the length of the storm event. The storm event lasts five hours and the maximum street flooding is reached after 3-4 hours. The initial infiltration capacity (127 mm/h in the default setting) decreases with a factor e^{-6t} to the final infiltration capacity (12.7 mm/h in the default setting), as a k value of 6.0 was selected for pervious surfaces. With the selected k-value, 90% of the difference between initial and final infiltration capacity is vanished after 46 minutes of rainfall (calculated with Equation 1). This implies that for most of the storm event, the final infiltration capacity is 'used'.

Conclusion

Lowering the infiltration parameters f_0 and f_c influences the test scores of the two tests that are proposed in chapter 6. The amount of true positives in test 1 increases from 39% to 83% if both parameters values are decreased by 30%. The combined score, an indicator for the model's ability to estimate the extent of surface flooding, increases from 68% to 87%. The scores for test 2 increase slightly, indicating a somewhat better estimation of in-sewer processes. For the selected test case, lowering f_c has a larger influence on the results than lowering f_0. The most likely reason for this is the length of the storm event.

Glossary

AHN: Algemeen Hoogtebestand Nederland. Digital elevation model of the Netherlands.

BGT: Basisregistratie Grootschalige Topografie. A database with information on land use in the Netherlands.

DEM: Digital Elevation Model. A height map of the area *(Dutch: hoogtekaart, hoogtemodel)*.

False Negative: in the context of this report: the model identifies an observed location of flooding as not flooded.

False Positive: in the context of this report: the model identifies an observed location of no flooding as flooded.

Fluvial Flooding: flooding due to overflowing rivers *(Dutch: rivieroverstroming).*

Fully distributed runoff: a type of runoff modelling where rain falls equally onto each grid cell of the surface model, and flows towards gullies according to 2D Shallow Water Equations *(Dutch: instroommodel aan de hand van 2D maaiveld stroming)*.

GIS: Geographical Informarion System. In the context of this report: a tool in which interactive maps can be created.

Gully pot: entry point for storm water into the underground sewer system *(Dutch: kolk).*

Impervious surface area: surface area of which the infiltration capacity is small, such as concrete, asphalt or roofs *(Dutch: ondoorlatend oppervlak)*.

Kennisbank Stedelijk Water: Dutch database with all kinds of information on sewers, including modules on design principles, maintenance, laws, research, finances and sewer models.

KGE: Kling-Gupta Efficiency. A measure for the agreement between a model and the corresponding observations. The ideal value is at unity. Used in test 2.

Leidraad Riolering: former name of the Kennisbank Stedelijk Water. In Module C2100, parameters and guidelines for the creation of an urban drainage model are provided.

Manhole: top opening to the underground sewer system. It can be used to enter, inspect or maintain the sewer system *(Dutch: riool put).*

Modelling concept: the method of representing the urban water system. In the context of this report, a modelling concept is the conceptual basis on which a sewer model is based *(Dutch: modelconcept, modeloptie)*.

Node: term in InfoWorks ICM used for the representation of certain elements in a sewer system, such as manholes, gully pots and outfalls. A conduit (link) is the connection between two nodes *(Dutch: knoop, knooppunt)*.

NDVI: Normalized Difference Vegetation Index. The normalized difference between the reflection of red and near infrared light. In this report used as a method to make a division between pervious and impervious surface area.

NSE: Nash-Sutcliffe Efficiency. A measure for the agreement between a model and the corresponding observations. The ideal value is at unity. Used in test 2.

Percent point: arithmetic difference of two percentages. If model A scores 30% and model B scores 90%, the difference is 300 percent or 60 percent point *(Dutch: procentpunt)*.

Pervious surface area: surface area of which the infiltration capacity is rather large, such as grass, farm land or parks *(Dutch: doorlatend oppervlak)*.

Pluvial flooding: flooding due to heavy rainfall *(Dutch: overstroming door regenwater)*.

Pump sump: a sewer manhole in which a pump is installed that discharges water to a higher location or to a pressurized pipe *(Dutch: pompput)*.

Raster: a map covering a certain area. Technical jargon used in GIS-applications.

RMSE: Root Mean Squared Error. A measure for the agreement between a model and the corresponding observations. The ideal value is zero and the unit is in meters or centimeters. Used in test 2.

Semi-distributed runoff: the 'classic' way of modelling runoff. Each manhole is assigned an amount of surface area, a so-called sub catchment. The surface area inside this sub catchment is divided into predefined land use categories (for example, 40% pavement, 30% roofs, 30% yards, 0% parks). Each of these land use categories has its own parameters for runoff delay, infiltration and storage. Sewer inflow is calculated for every time step and forms the input for the hydraulic sewer model *(Dutch: instroommodel aan de hand van deelstroomgebieden)*.

STOWA: Stichting Toegepast Onderzoek Waterbeheer **.** Dutch institute for applied research concerning water management.

Stichting RIONED: Dutch institute for applied research on sewer systems.

Sub catchment*:* An area from which rain water flows towards a certain point. In the case of urban drainage, this point is often assumed to be a manhole. *(Dutch: deelstroomgebied)*

True Negative: in the context of this report: the model identifies an observed location of no flooding correctly.

True Positive: in the context of this report: the model identifies an observed location of flooding correctly.

Unsaturated Zone: the part of the subsurface between surface level and the groundwater table *(Dutch: onverzadigde zone).*

WWTP: Waste Water Treatment Plant. Facility where waste water (and storm water in case of a combined sewer system) is treated before discharged into a surface water body *(Dutch: rioolwaterzuiveringsinstallatie).*