Fit-for-purpose modelling to determine climate durability of wadies

Student: Christian van de Laarschot
6198309
c.j.vandelaarschot@students.uu.nl

University Supervisor: dr. Menno Straatsma

Organisation: Geonius
Supervisor: Kas Lange k.lange@geonius.nl Tom Van Es t.vanes@geonius.nl

Date: 30 September 2019

Words: 11.444

Acknowledgments

This thesis is the final product of the Master Water Science and Management at the Utrecht University. The realization of this report would not be possible without the valuable contributions of others.

Firstly, I want to thank my supervisor Menno Straatsma for his patience, detailed feedback and guidance.

In addition, I would like to thank Kas Lange and Tom van Es from Geonius, for their guidance and feedback during the process. I also want to express my gratitude towards Ali Mirrezai, a fellow intern at Geonius, for his help during the fieldwork.

Christian van de Laarschot, 30-09-2019

Summary

Growing urban areas and more paved surfaces results in an increasing runoff of rainfall to the sewer system. To reduce the stress of a rainfall event and delay the discharge to the sewage system, a wadi is constructed in urban areas. Because of climate change, more intense rainfall events and higher intensities are expected to occur. These rainfall events and higher intensities could exceed the amount to what wadies are designed. By modelling the infiltration capacity of wadies and scenario modelling climate change on wadies, the climate durability of wadies can be determined. However, it is unknown if current models are fit-for-purpose to determine whether a wadi is able to cope with more intense rainfall events and higher intensities. Moreover, the accuracy of infiltration modelling is unknown.

What current available model is fit-for-purpose to determine the climate durability of wadies is assessed by reviewing literature of seven different hydrological models and score these models with benchmark criteria on:

- model applicability and relevance;
- model uncertainty and sensitivity;
- model transparency, ease of understanding and ease of use.

From the literature review and benchmark criteria, the Storm Water Management Model is considered fit-for-purpose to determine the climate durability of a wadi.

To determine the accuracy of the Storm Water Management Model (SWMM) for infiltration modelling, reference data of three wadies in the city of Utrecht is collected and compared to simulated values. The accuracy of the model is described with the Nash-Sutcliffe Efficiency (NSE), were a NSE value of 0.65 represents an acceptable accuracy and a NSE value of 1.0 represents a perfect fit. Reference data on groundwater elevation, ponding depths and rainfall intensity is collected between the 16th of May 2019 and 28th of July 2019. Model input data on saturated conductivity and unsaturated conductivity is measured once by the Porchet method and double ring infiltrometer.

Calibration of two rainfall events, on the 4th of June 2019 and 13th of June, resulted in an average NSE value for groundwater elevations of 0.79 at the Klifrakplantsoen. The NSE value for maximum ponding depths for these two rainfall events is 0.99. Validation of the SWMM model on the rainfall event on the 19th of June resulted in a NSE value of 0.75 in groundwater elevation and 1.0 on maximum ponding depths. Calibration of the model at the Karel Doormanlaan is done by rainfall events on the 4th of June and the 12th of June. The simulated values on groundwater elevations resulted in average NSE value of 0.72. NSE values on maximum ponding depths are divided, the calibration event on the 4th of June resulted in a NSE value of -2.54 were the calibration event on the 12th of June resulted in a NSE value of 0.99. Validation of the model is done with the rainfall event on the 19th of June. This resulted in a NSE value of 0.75 on groundwater elevations and 0.78 in maximum ponding depths.

This research showed that the Storm Water Management Model is fit-for-purpose to determine the climate durability of wadies. The SWMM model is able to determine the groundwater elevations in two wadies with an acceptable to good accuracy. Determination of maximum ponding depths varies between extreme low accuracy values and almost perfect values. In general, the maximum ponding depths are determined with a good accuracy. However, the accuracy of ponding depths on complete timeseries are less satisfactory.

Table of contents

1. Introduction

Urban areas are growing and so are paved surfaces (United Nations, 2018). The consequence is an increasing runoff of rainfall from roofs and paved surfaces to the sewer system. In case of intense rainfall events, the sewer system is incapable to discharge the rainfall in time, resulting in flooded street surfaces. In many cities, a *wadi* is used to collect the rainfall and delay the discharge during the peak of a rainfall event. There is a distinction between two different definitions of a wadi. The first definition of a wadi is a valley, ravine or channel which is dry in most seasons except in the rain season. The second definition of a wadi is a Dutch abbreviation for "Water Afvoer Door Infiltratie", which translates in English to Water Discharge By Infiltration. The definition of a wadi in this research refers to the second definition.

A wadi is a ditch with a permeable soil. The top layer of the soil can be enhanced to increase the infiltration capacity of the soil. Beneath the top layer, a storage layer can be constructed. This storage layer can be made out of porous material, e.g. gravel, plastic crates or clay pellets (STOWA, 2003). Some wadies include a drain at the lowest point of the wadi to increase the discharge capacity of the wadi towards surface water. Figure 1 shows an example of a wadi in a residential area.

Figure 1 Example of a wadi in a residential area

In 2014, the Intergovernmental Panel on Climate Change (IPCC) presented a report with the latest state of knowledge regarding climate change. This report stated that western Europe has to cope with an increase in rainfall amount varying from 0-10% in the RCP2.6 scenario to 10-20% in the RCP8.5 scenario (Core Writing Team, Pachauri, & Meyers, 2014). The Koninklijk Nederlands Meteorologisch Instituut (KNMI) translated the results of the IPCC report to four climate scenarios for the Netherlands in which the increase in rainfall amount for the Netherlands varies from 2.5% to 5.5% in 2050 and 5% to 7% in 2085 (van den Hurk et al., 2014).

Requirements of existing wadies are based on the frequency of rainfall events developed in 2004. This frequency is based on rainfall data from the period 1906-2003, but is not representative for future climate conditions (STOWA, 2004a). This raises the question whether existing wadies are capable of accommodating future rainfall events within the lifespan of the wadi, resulting in possible floodings of surfaces next to the wadi. Wadies are designed with an expected hydraulic lifespan of 40 to 60 years, similar to sewage systems (Boogaard & Wentink, 2007). When a wadi is not able to accommodate future rainfall events within its lifespan, it is more cost-efficient for municipalities to take proactive measures during major maintenance activities than reacting to damage costs (Figure 2).

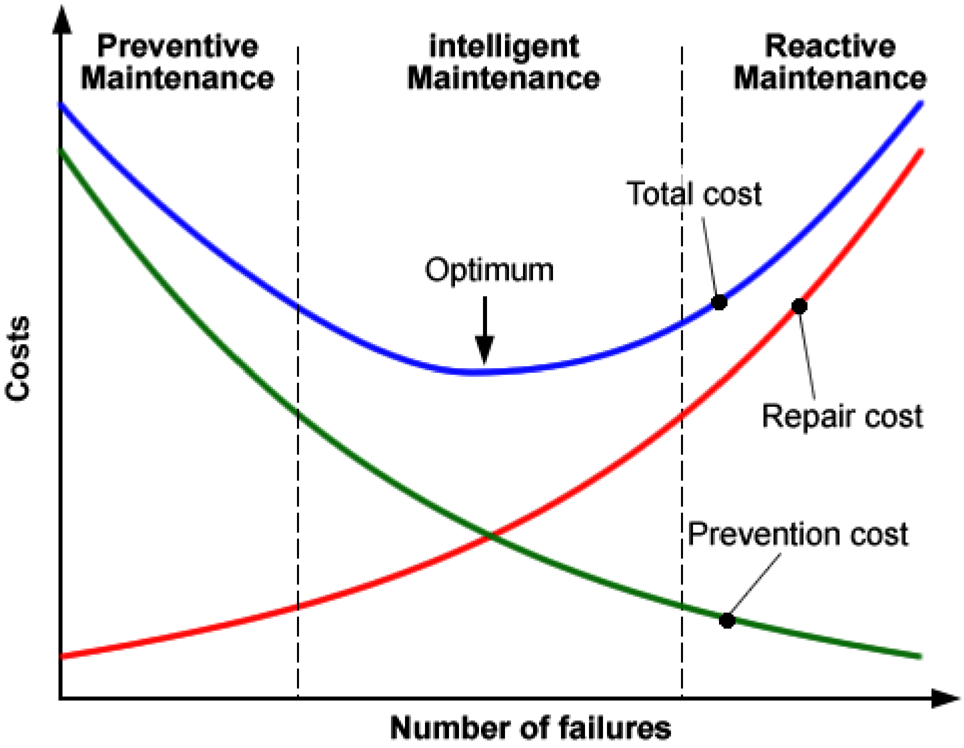


Figure 2 Traditional maintenance strategies (Tchakoua, Wamkeue, Slaoui-Hasnaoui, Tameghe, & Ekemb, 2013)

Current guidelines for wadies are based on stationary rainfall events, which means that the sum of rainfall amount over a certain time, for example, a rainfall event of 48 hours, is used (Wijngaard, Kok, Smits, & Talsma, 2005). However, the sum of rainfall is often a result of intense peaks interspersed with longer-lasting moderate rainfall. Moreover, the intense peaks of a rainfall event determine the ponding depth and possible flooding of surfaces next to a wadi. Therefore, it is more useful to determine whether a wadi is climate-resistant by using a non-stationary or dynamic model. Consultants currently use a model that assumes the infiltration rate equal to conductivity and rainfall as a stationary event. The accuracy of this model is not known and only used for designing new infiltration facilities, but the desire is to use a model for assessing existing infiltration facilities based on observed data instead of conductivity measurements.

Due to the increasing calculation power of computers, geographical information systems and digital terrain maps, hydrological models have been developed rapidly since the first model of Freeze and Harlan was published in 1969 (Beven K. , 1990). This rapid development results in a wide range of models for hydrological modelling. There are at least two motivations for the development of each of these models. Firstly, assist understanding of physical systems by providing a framework for hypothesis testing and secondly to provide a predictive tool. These two reasons are frequently confused (Grayson, Moore, & McMahon, 1992). The context in which the model is originally developed is often lost, so the model is applied in situations outside its scope or capabilities (Grayson, Moore, & McMahon, 1992). For example, a model which is developed for calculating outflow hydrographs does not necessarily imply that the output flow depths are comparable to observed values (Grayson, Moore, & McMahon, 1992). Similar, a model that is developed for a particular catchment area or climate region is not necessarily able to generate satisfying model outputs in a different catchment or climate region. For example, the HBV-model is designed for calculating flow rates, percolation rates and evaporation rates in particular for the Swedish catchment area but is also successfully applied in catchment areas in Zimbabwe, India and Colombia (Swedish Meteorological and Hydrological Institute, 2019). Other models, such as the Soil Water Assessment Tool (SWAT) and the Hydrologic Modeling System (HEC-HMS), are not developed for a specific catchment area but are developed for specified hydrological processes. The SWAT model primarily developed to predict impacts of land management practices on water and sediment in large watersheds with varying soils (Neitsch, Arnold, Kiniry, & Williams, 2011). The HEC-HMS model is developed for various tasks, including urban flooding studies and environmental studies in watersheds (U.S. Army Corps of Engineers, 2008). These models are fit-for-purpose to calculate an output of the model within the scope it is developed. However, it is unclear whether available models have the ability to determine the extent to which a wadi is climate proof. Therefore, it is necessary to determine when and what model is fit-for-purpose for assessing current and future infiltration capacity of an infiltration facility. Furthermore, the accuracy of infiltration modelling for this purpose and effect of climate change on the accuracy of infiltration modelling is unknown.

1.1 Research question

For this research, the following research questions are stated:

“What hydrological infiltration model is fit-for-purpose to assess the infiltration capacity of infiltration facilities (wadies) within the design process?”

“What is the accuracy of this hydrological infiltration model in relation to groundwater elevation and ponding depths?”

“What is the future infiltration capacity of infiltration facilities?”

The following minor objectives are stated to answer the main research questions:

- Define requirements that make an infiltration model fit-for-purpose;
- Compare seven hydrological models with the benchmark criteria and defined requirements;
- Selecting a fit-for-purpose model by the benchmark criteria;
- Set up of the model for three cases;
- Field campaign for collecting reference data on groundwater elevation, ponding depths, conductivity and porosity;
- Calibration and validation of the model with reference data on three cases;
- Sensitivity analysis of the model in three cases;
- Scenario analysis of climatological changes for three cases.

2. Literature review

This chapter describes the current literature on infiltration equations and how these infiltration equations are embedded in different hydrological models. Literature on hydrological model calibration and validation is reviewed, followed by sensitivity analysis. Afterwards, a description of wadies and its design requirements is given. Finally, climate change and scenario modelling are described before reviewing literature on fit-for-purpose modelling.

2.1 Infiltration

Infiltration of water into the soil is an important process in hydrology, agriculture and urban water management. The process of infiltration is influenced by many factors, such as soil properties, soil depth, geomorphology and rainfall (Morbidelli et al., 2018). The understanding of how these factors influence the infiltration process and how to mathematically describe this process have been developed over the last decades. This has resulted in a few point infiltration models. An inexhaustive list includes the Horton Empirical equation, Philip equation and the Green-Ampt equation. These equations differ in their approach of calculating the infiltration.

The Horton equation is an empirical equation which does not incorporate ponding. It describes the infiltration rate by the initial and final infiltration capacity and decreases exponentially. The final infiltration capacity is considered equal to the saturated hydraulic conductivity (Morbidelli et al., 2018). Figure 3 shows how the infiltration rate decreases over time, according to the Horton equation.

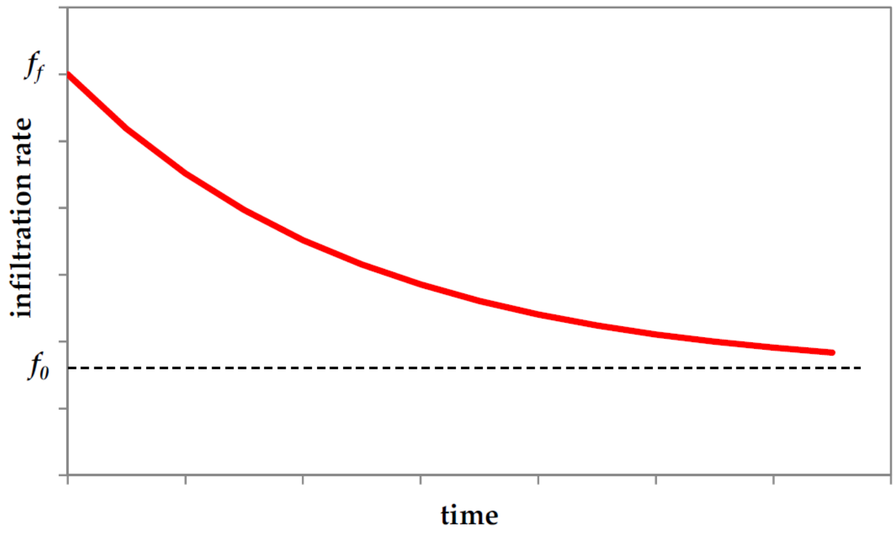


Figure 3 Graphical representation of the Horton empirical equation (Morbidelli, et al., 2018)

The Philip equation is an analytical solution that describes the infiltration rate by the sorptivity of the soil. The sorptivity is influenced by the soil properties and initial moisture content (Morbidelli et al., 2018). In this analytical solution, it is assumed that there is a saturated soil surface and immediate ponding. Philip's equation is extended for less restricted conditions in which no immediate ponding occurs. In this extended equation, it is assumed that the soil is completely saturated after the time of ponding. In the equation, ponding occurs when the constant rainfall is greater than the saturated conductivity (Morbidelli, et al., 2018).

The Green-Ampt model applies Darcy’s law and the principle of conservation of mass (Dingman, 2015). Within the Green-Ampt model, they use two situations. The first situation considers the water input by rain is smaller than the hydraulic conductivity of the homogenous soil. The second situation considers a higher water input by rain than the hydraulic conductivity of the homogenous soil (Dingman, 2015). In the first situation, the infiltration rate is considered the same as the water input by rainfall. The second situation starts when the water input by rain exceeds the hydraulic conductivity of the homogenous soil. The moment when the surface layer becomes saturated is called time of ponding. This approach is similar to the approach of the extended Philip equation. In the Green-Ampt equation, the infiltration rate after ponding is calculated with using the wetting front depth and height of ponding.

Another approach to mathematically describe infiltration of water in the soil is by assuming that the conductivity of the soil is equal to the infiltration rate of the soil. This equation is currently used by consultants for designing infiltration facilities and is too simplified for accurate estimations of infiltration rates.

For modelling and designing infiltration facilities, ponding depths are most relevant to calculate. When ponding depths become above set levels, the infiltration facility does not meet the requirements set for the infiltration facility. Equations which mathematically describe the infiltration process without ponding are presumably not fit-for-purpose for designing infiltration facilities. Furthermore, all equations assume that groundwater elevation is infinite below surface level or is not incorporated in the equation at all. When groundwater elevation is close to the surface, it can be expected that ponding occurs earlier and ponding depths become higher.

2.2 Hydrological models

In general, three different types of hydrological models can be distinguished. These types are empirical models, conceptual models and physically based models (Devi, Ganasri, & Dwarakish, 2015). Empirical models use mathematical equations which are derived from input and output time series. Conceptual models describe the hydrological processes by interconnected reservoirs and semi-empirical equations. The reservoirs in conceptual models represent physical elements and the model parameters are determined with field data and through calibration. For calibration, large records of meteorological and hydrological data is required (Devi, Ganasri, & Dwarakish, 2015). Physically-based models are an idealized representation of the real phenomenon. These models use state variables which are measurable and a function of time and space. Less data for calibration is needed, however, more data on physical characteristics of the catchment or study area is needed (Devi, Ganasri, & Dwarakish, 2015). Table 1 shows a summary of each type of model.

Table 1 Characteristics of three types of hydrological models (Devi, Ganasri, & Dwarakish, 2015)

Emperical model	**Conceptual model**	**Physically based model**
Data based or metric or black box model.	Parametric or grey box model.	Mechanistic or white box model.
Involve mathematical equations, derive value from available time series.	Based on modeling of reservoirs and include semi empirical equations with a physical basis.	Based on spatial distribution, evaluation of parameters describing physical characteristics.
Little consideration of features and processes of system.	Parameters are derived from field data and calibration.	Require data about initial state of model and morphology of catchment.
High predictive power, low explanatory depth.	Simple and can be easily implemented in computer code.	Complex model. Require human expertise and computation capability.
Cannot be generated to other catchments.	Require large hydrological and meteorological data.	Suffer from scale related problems.
ANN, unit hydrograph	HBV-model, TOPMODEL	SHE or MIKESHE model, SWAT
Valid within the boundary of given domain	Calibration involves curve fitting make difficult physical interpretation	Valid for wide range of situations

Because the modelling of infiltration facilities is done for different catchments, empirical models are most likely to be insufficient for this task. Examples of conceptual and physically based models which can determine the rate of water infiltrating into the soil are the Hydrologiska Byråns Vattenbalansavdelning model (HBV model), Hydrologic Modeling System (HEC-HMS model), ModFlow-2005, Sobek Urban, Storm Water Management Model (SWMM) and Soil Water Assesment Tool (SWAT).

The HBV model uses subbasins with area-elevation and crude classifications of land use (forest, open and lakes) as primary hydrological units. Within the HBV model there are three main components (Bergström, 1992):

- subroutines for snow accumulation and melt;
- subroutines for soil moisture accounting;
- response and river routing subroutines.

The subroutine for soil moisture consists of two reservoirs with parameters for recession coefficient, a threshold limit for quickest runoff component and a constant percolation rate (Bergström, 1992). The constant percolation rate is mainly responsible for the total infiltration rate.

The HEC-HMS model was developed by the US Army Corps of Engineers mainly for analysing urban flooding, flood frequency and flood-loss reduction measures (U.S. Army Corps of Engineers, 2008). Infiltration of water in the soil with the model can be determined by the initial and constant-rate loss method, the deficit and constant rate method, the SCS curve number method and the Green-Ampt method (U.S. Army Corps of Engineers, 2000).

Modflow was originally developed as a groundwater flow model. The authors of the model felt that additional related equations could be done in separate programs (Harbaugh, 2005). By the late 1990s, the authors decided to allow Modflow incorporate capabilities such as transport and parameter estimation. Since this incorporation, the percolation rate of water in the unsaturated zone is approximated by simplifying Richards equation (Niswonger, Prudic, & Regan, 2006).

Sobek is an integrated software package for river, urban and rural management developed by the Dutch research institute Deltares. The Sobek urban package can be used for analysing the performance of the urban drainage system. Infiltration in the Sobek urban package is based on the Horton equation (Deltares, 2018).

The SWMM model is a rainfall-runoff model, primarily for urban areas (Rossman & Huber, 2016). The model is functioning with four compartments. The atmosphere compartment which generates precipitation. The land surface compartment, which receives precipitation and generates outflow by converting it to evaporation, surface runoff and infiltration to the sub-surface compartment. The sub-surface compartment transforms the inflow by infiltration to groundwater interflow and the conveyance compartment contains a network of elements, for example, pipes, channels and pumps (Rossman & Huber, 2016). The infiltration processes which can be chosen in the SWMM model are the Horton method, modified Horton method, Green-Ampt method and the Curve Number method. Unique for the SWMM model is the possibility of implementing Low Impact Development (LID) measures. The LID measures can be used for reducing surface runoff. The types of LID measures in the SWMM model are rain gardens, bio-retention cells, green roofs, infiltration trenches, permeable pavements, rain barrels and vegetative swales.

The SWAT model is a physically-based model which is mainly used in agricultural and rural watersheds (Hunt, Kannan, Jeong, & Gassman, 2019). The model requires specific information on weather, soil properties, topography, vegetation and land management practices (Neitsch, Arnold, Kiniry, & Williams, 2011). Infiltration in the SWAT model is divided in percolation and recharge of the groundwater. Percolation is determined based on the soil water content and field capacity, recharge is determined by a function of Venetis (Neitsch, Arnold, Kiniry, & Williams, 2011).

2.3 Model calibration and validation

Model calibration is “the process of adjustment of the model parameters and forcing within the margins of the uncertainties (in model parameters and/or model forcing) to obtain a model representation of the processes of interest that satisfies pre-agreed criteria” (Vlaams Instituut voor de Zee, 2019a). In general, there are two basic methods for hydrological model calibration. The first and most simple method is a trial and error procedure in which the user changes the parameter values (Anderson, 2002). By comparing simulated values with observed values, decisions on parameter changes can be made. The calibration process with this method is finished when the user determines that the objectives of the model have been met (Anderson, 2002). The objectives of the model can be an accurate simulation of groundwater levels over time.

The second method to calibrate hydrological models is the automated model calibration method. Automated model calibration has been under development for over three decades, in which the degree of sophistication is parallel to improving computing power (Boyle, Gupta, & Sorooshian, 2000). With the automatic calibration method, the user is required to specify feasible upper and lower bounds for each parameter. Within these parameter boundaries, an algorithm determines the optimal fit. Most algorithms define the quality of reproduction by a statistical method such as the daily root mean square error (RMSE) of an output parameter (Boyle, Gupta, & Sorooshian, 2000). The RMSE shows how concentrated data is around the line of best fit. With a lower RMSE, simulated values are closer to the observed values. Higher RMSE values indicate that simulated values differ more from the observed values.

When model calibration is finished, the model is validated. Model validation “is the formal confirmation that the model meets the quality criteria achieved in model calibration. Validation of the model is done with a set of independent data and model parameters are fixed with the calibrated values" (Vlaams Instituut voor de Zee, 2019b). Different methods to quantify to what extent models perform within the set quality criteria are available. A model validation method that has received considerable attention is the Nash Sutcliffe Efficiency (NSE) coefficient (Ritter & Muñoz-Carpena, 2013). The NSE is a dimensionless goodness-of-fit indicator which represents the ratio between the RMSE of observed values versus predicted values and the variance of observations. An NSE value of 1 represents a perfect fit of model results compared to observed values. The NSE is needed to determine the accuracy of the fit-for-purpose model.

2.4 Sensitivity analysis

One of the steps in model development is the determination of which parameters are most influential on model results. This step is called "sensitivity analysis" and is done for several reasons. These reasons include:

- Determination of which parameters requires additional research to reduce uncertainty;
- Determination of which parameters are insignificant and can be eliminated from the model;
- Determination of which parameters contribute most to output variability;
- Determination of which parameters are most highly correlated with the output (Hamby, 1994).

Hamby (1994) has reviewed over a dozen methods such as differential analysis, one-at-a-time design, factorial design and the relative deviation method (Hamby, 1994). According to Hamby (1994), the consensus among all scientific literature on sensitivity analysis is that models are sensitive to input parameters in two different ways. The first consensus among literature is that the variability associated with a sensitive parameter is propagated through the model resulting in a large contribution to the overall output variability. The second consensus among literature is that model results can be highly correlated to an input parameter. This means that a small change in the input value can result in a significant change in the output value (Hamby, 1994).

Important with sensitivity analysis is the distinction between important and sensitive parameters. This distinction is reflected in the type of analysis conducted: uncertainty analysis is done to determine parameter importance and sensitivity analysis is done to determine parameter sensitivity (Hamby, 1994). A graphical comparison, sensitivity plot, of percentage change in output and percentage change in parameter can be used to examine the stability of a parameter to the optimum solution (McCuen, 1973). Derivation of these sensitivity plots is an iterative process in which the percentage change of output is computed for different percentage changes in parameter value. For multi-parameter models, this derivation is often time extensive (McCuen, 1973). The sensitivity analysis is relevant to determine which parameter is influencing the accuracy of the model the most.

2.5 Wadies

A wadi is the abbreviation of *Water Afvoer Door Infiltratie* in Dutch, which translates to Water Discharge by Infiltration in English. Wadies are used to infiltrate rainwater from rooftops and residential streets and reduce the pressure on sewer systems during extreme rainfall events. Usually, a wadi looks like a regular ditch, but the soil in a wadi is permeable so rainwater can infiltrate into the ground (STOWA, 2003). Some wadies contain an enhanced top layer of the soil, this is done to increase the infiltration capacity of the soil. The hydraulic lifespan of a wadi in the Netherlands is estimated at 40-60 years (Boogaard & Wentink, 2007). Figure 4 shows the profile of a regular wadi with an enhanced top layer.

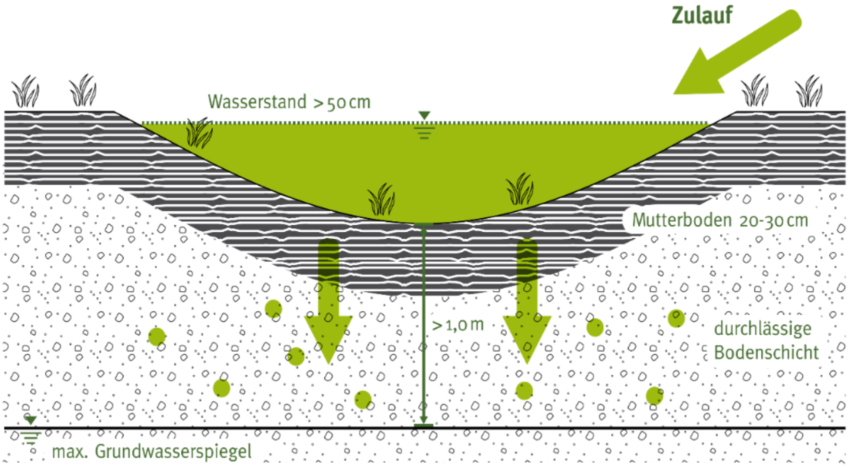


Figure 4 Cross-section of a regular wadi (Autonome Provinz Bozen - Südtirol, 2019)

When the soil is unable to infiltrate the precipitation and runoff from rooftops and paved surfaces, additional water storage is constructed. This additional storage is made from porous media such as gravel, clay granules or a plastic infiltration crate. In some situations, a drain is placed to discharge the infiltrated water in the storage to surface water. The storage is surrounded by geotextile to prevent clogging by soil particles (STOWA, 2003). Figure 5 shows a cross-section of a wadi with a storage and drain.

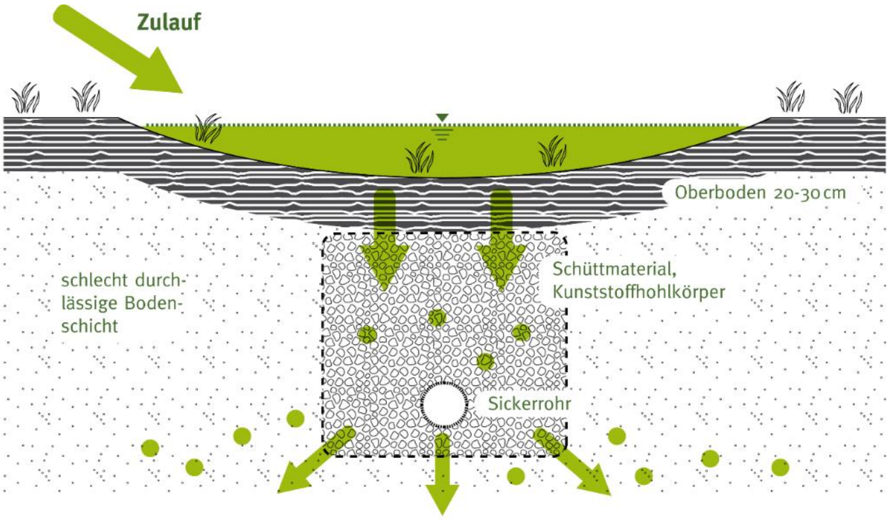


Figure 5 Cross-section of a wadi with infiltration crates (Autonome Provinz Bozen - Südtirol, 2019)

The most important function of a wadi is the infiltration of water and the storage of water (STOWA, 2003). Infiltration of water reduces the impact and pressure of heavy rainfall events on surface water and sewer systems. Water which usually flows directly into the sewer system is retained and eventually discharged by infiltration. This gives surface water the opportunity to distribute the discharge of the rainfall event over a longer period of time.

When precipitation is too intens to infiltrate in the soil it can be stored in the wadi. In regular wadies, the storage is done at the surface like regular ditches. In wadies with storage facilities, the storage first is done at the surface and then in the storage facility below the surface. This storage prevents direct overflow to surface water and sewer systems when the precipitation rate is higher than the infiltration rate (STOWA, 2003).

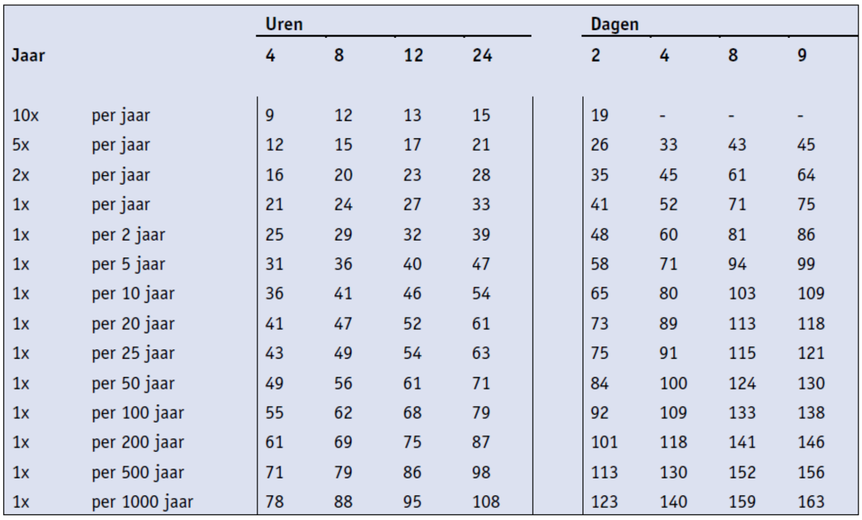
Wadies are designed with the ability to cope with extreme precipitation events. The ability to cope with extreme precipitation events are related to standards which describe the annual probability in which inundation occurs. In 2003, the Dutch government introduced standards that legally described the annual probability of surfaces due to precipitation. Table 2 shows the annual probability of five types of land-use according to the NBW standard (Hoes & Schuurmans, 2006). Depending on the municipality in which the wadi is located, different design regulations can occur.

Table 2 NBW-standards per land-use type

Standard related to land-use	**Annual probability (1/yr)**
Grassland	1/10
Agriculture	1/25
(Greenhouse) horticulture	1/50
Urban and industrial	1/100

The frequency in which inundation occurs is related to the amount of rainfall in a certain time. The frequency of rainfall events is determined by analyzing rainfall data of De Bilt from 1906 to 2003. For every year in this period, the ten maximum rainfall amounts for rainfall durations of 4, 8, 12, 24, 48, 96 and 192 hours are used to define the return time of rainfall events from once per year to once per 1000 years. To determine the rainfall amount of rainfall events which occur ten times a year to twice a year, peak-over-threshold (POT) values with an average threshold of ten times a year are used (Smits, Wijngaarden, Versteeg, & Kok, 2004). To get independent POT-values, a filter with an interval of 24 hours between the rainfall events is used. To get a good probability distribution or return time value, the method of generalized extreme value (GEV) is used.

The GEV method is chosen because of the extra shape parameter which proved to be an added value to compare De Bilt weather station to other stations (Smits, Wijngaarden, Versteeg, & Kok, 2004). The design precipitation is showed in Figure 6. For example, an urban area that may flood once per 100 years has to accommodate a rainfall amount of 55 millimetres in 4 hours.



		Uren				**Dagen**			
Jaar		**4**	**8**	**12**	**24**	**2**	**4**	**8**	**9**
10x	per jaar	9	12	13	15	19	-	-	-
5x	per jaar	12	15	17	21	26	33	43	45
2x	per jaar	16	20	23	28	35	45	61	64
1x	per jaar	21	24	27	33	41	52	71	75
1x	per 2 jaar	25	29	32	39	48	60	81	86
1x	per 5 jaar	31	36	40	47	58	71	94	99
1x	per 10 jaar	36	41	46	54	65	80	103	109
1x	per 20 jaar	41	47	52	61	73	89	113	118
1x	per 25 jaar	43	49	54	63	75	91	115	121
1x	per 50 jaar	49	56	61	71	84	100	124	130
1x	per 100 jaar	55	62	68	79	92	109	133	138
1x	per 200 jaar	61	69	75	87	101	118	141	146
1x	per 500 jaar	71	79	86	98	113	130	152	156
1x	per 1000 jaar	78	88	95	108	123	140	159	163

Figure 6 Design precipitation amounts in mm and return period (Smits, Wijngaarden, Versteeg, & Kok, 2004)

The water board Hoogheemraadschap de Stichtse Rijnlanden requires that an infiltration facility above the ground is empty after 24 hours from the end of the rainfall (Hoogheemraadschap de Stichtse Rijnlanden, 2015). A wadi is empty when there is no ponding water. The precipitation event to assess what time is needed to empty the wadi is derived from two different regulations. The first regulation refers to the requirements of the sewage system. Sewage systems are designed for precipitation events with a return time of once in two years (Gemeente Utrecht, 2011). The other regulation refers to the NBW-standards for urban areas. According to this standard, infiltration facilities are designed for precipitation events with a return time of once in 100 years (Gemeente Utrecht, 2011).

2.7 Climate change and scenario modelling

Since the late 20th century, more unusual changes in climate occur. These unusual changes in climate are caused by an increase in greenhouse gasses (GHG) such as carbon dioxide. Model projections of the "Business As Usual" scenario show extraordinary temperatures estimates for 2100, which is the warmest over the past 400.000 years (Crowley, 2000). The effects of climate change have consequences for humans all over the world, therefore anthropogenic climate change has appeared on the public agenda since the mid-to-late 1980s (Moser, 2010). The IPCC report of 2014 stated that western Europe has to deal with an increase in rainfall amounts and intensity (Core Writing Team, Pachauri, & Meyers, 2014). This increase in average rainfall amounts is estimated from 0-10% in the RCP2.6 scenario to 10-20% in the RCP8.5 scenario, see Figure 7 (Core Writing Team, Pachauri, & Meyers, 2014).

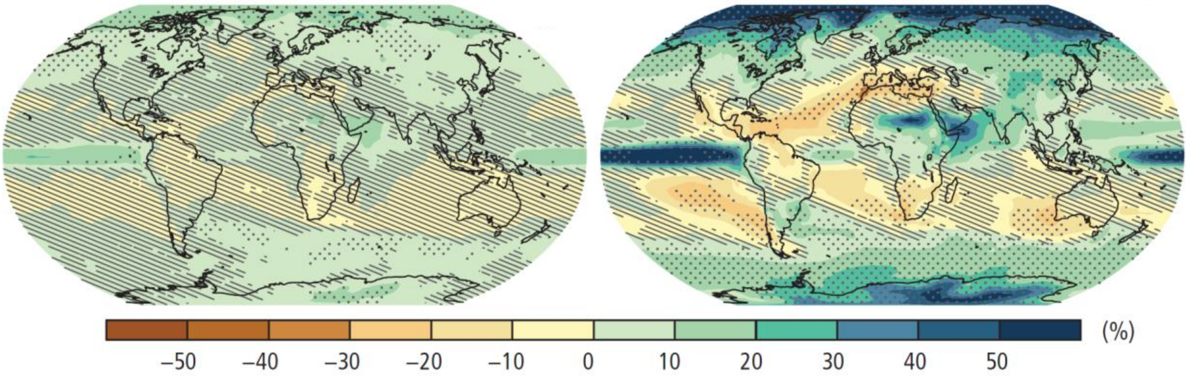


Figure 7 Change in average precipitation according to IPCC Report 2014, left RCP2.6 and right RCP8.5

The KNMI translated the results of the IPCC report of 2014 to four climate scenarios to determine the increase in rainfall for the Netherlands. Input for these climate scenarios are global temperature rise and change in airflow patterns (KNMI, 2015). For the increase in rainfall towards 2050 a global temperature rise of 1°C for a moderate scenario "G" is used and 1,5°C for a hot scenario "W" is used. For the increase in rainfall towards 2085 a temperature rise of 1,5°C for a moderate scenario and 3,5°C for a hot scenario is used. For change in airflow patterns, the values of a weak model response and strong model response are used. Weak response "L" results in small changes in precipitation for summer and winter. Strong response "H" results in wetter winters and drier summers (van den Hurk, et al., 2014). The combination of these input values results in the four scenarios G_L, G_H, W_L and W_H. The increase in average yearly rainfall for those four scenarios in 2050 and 2085 are showed in Table 3.

Table 3 Average increase in rainfall

Scenario	2050	2085
G_L	+4%	+5%
G_H	+2,5%	+5%
W_L	+5,5%	+7%
W_H	+5%	+7%

The effect of climate change can affect extreme precipitation events in two ways. The first is an increase in the amount of time an extreme precipitation event occurs. For example, a precipitation event which currently occurs once in 5 years, can occur once in 4 years due to climate change. The second way in which climate change can affect precipitation is by an increase in corresponding rainfall amount for an extreme precipitation event. For example, an extreme precipitation event which returns once in 50 years has a rainfall amount of 49 millimetres. This rainfall amount increases with 7% due to climate change, resulting in 52,4 millimetres of rainfall.

2.8 Fit-for-purpose modelling

Hydrological models are increasingly embedded in modelling systems that represent environmental processes. This is widely associated with increasing model complexity, lack of observational data and increasing number of model outputs (Wagener, et al., 2001). When developing a model, the complexity of a model should be a function of:

- The modelling purpose;
- The characteristics of the hydrological system;
- The available data.

The suitability of a model can be measured in terms of model performance of objective function values and the uncertainty of model parameters (Wagener, et al., 2001). To reduce the uncertainty of model parameters to a level which is acceptable, an ontological approach for the fitness of use of geospatial datasets is developed (Vasseur, Devillers, & Jeansoulin, 2003). This ontological approach for the fitness of use is based on the comparison of user requirements and data specifications using different criteria (Vasseur, Devillers, & Jeansoulin, 2003).

Compared to the long discussion among researchers on how to determine appropriate effective parameters, there is little discussion about what qualifies a model fit-for-purpose for different types of purpose (Beven, 2018). A few methods of model hypothesis are available, including a possible evaluation which allows the possibility of model falsification as not fit-for-purpose by the Generalized Likelihood Uncertainty Estimation (GLUE) (Beven, 2018). With this framework, it can be decided when model structure and parameter sets should be considered acceptable or rejected using what is known or can speculate about the nature of errors and what is needed to make a difference to a decision in the purpose of a model application (Beven, 2018). In general, the fit-for-purposeness of a model is determined by reducing uncertainties in model parameters. Benchmark criteria based on the concept of uncertainty management can help selecting a fit-for-purpose model (Saloranta, Kamari, Rekolainen, & Malve, 2003).

3. Methods

To answer the research question, a combination of field research and desk research is done. Figure 8 shows a schematization of the research in which the combination of field research and desk research is visualized. This chapter describes the different research locations and used methods to collect data.

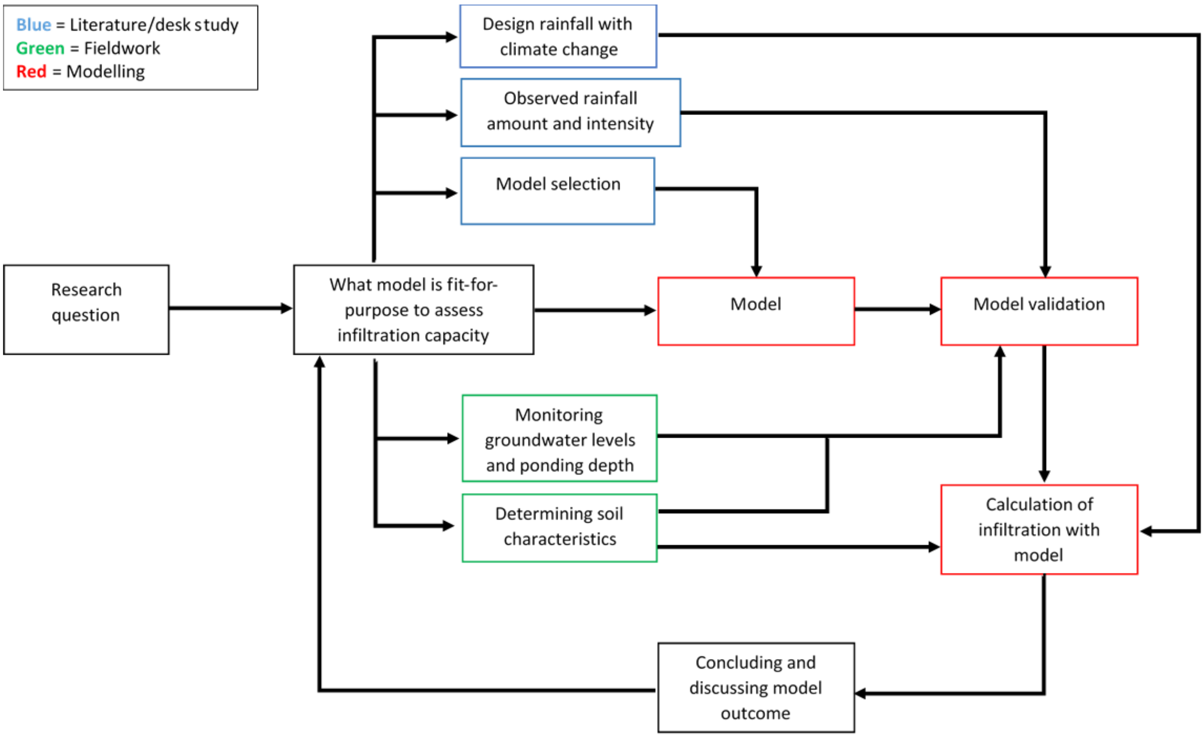


Figure 8 Schematization of research approach

3.1 Model selection

To determine which model is fit-for-purpose, seven different models are examined. The models which are examined are listed below:

- Hydrologiska Byråns Vattenbalansavdelning model (HBV-model)
- Hydrologic Modeling System (HEC-HMS model)
- ModFlow-2005
- Sobek Urban
- Storm Water Management Model (SWMM)
- Soil Water Assesment Tool (SWAT)
- Geonius Excel model

To answer the questions of the benchmark criteria for the HBV-model, the studies by Uhlenbrook (2009), Das (2008) and Lindström (1997) are used. The HEC-HMS model is reviewed by the literature of Zhang (2013) and Cunderlik (2004). ModFlow-2005 is reviewed by literature of Harbough (2005) and the SWMM model by McCutcheon (2013). By reviewing literature of Vergroesen (2014) and Bruni (2015) the questions of the benchmark criteria for the Sobek model can be answered. For the SWAT model, literature of Green (2007) and Van Griensven (2005) are reviewed to answer the questions in the benchmark criteria of Saloranta (2003).

The benchmark criteria contains questions in three categories:

- model applicability and relevance for the management task;
- model uncertainty and sensitivity;
- model transparency, ease of understanding and ease of use.

The management task, to which the benchmark criteria is referred, is a list of requirements which is made by consults with internship advisors at Geonius. The complete list of requirements is added in Annex I.

Every question can be answered with “Good”, “Adequate” or “Inadequate”. The answer “Good” is scored with 2 points, “Adequate” is scored with 1 point and “Inadequate” is scored with 0 points. Boundary conditions for each answer determine whether the question can be scored with “Good”, “Adequate” or “Inadequate”. The boundary conditions for each answer can be found in Annex II. The model which has the highest overall score is selected to carry out three case studies.

3.2 Site description

The three study areas were located in the city of Utrecht. Measurement instruments are placed on three different locations in the city. These locations are the Leuvenlaan, Klifrakplantsoen and Karel Doormanlaan. In the following paragraphs, each study site is described.

Leuvenlaan

The Leuvenlaan is located at the Utrecht Science Park (Figure 9). The wadi at the Leuvenlaan was not primarily designed to infiltrate water but to buffer water from nearby rooftops. Therefore, the soil is not enhanced, no drainage is constructed and no overflow to the sewage system is possible. The wadi is 75 meters long and 10 meters wide. The Koningsberger building, north of the wadi, is the only surface that discharges on the wadi. The total surface of this building is 3.650 square meters.

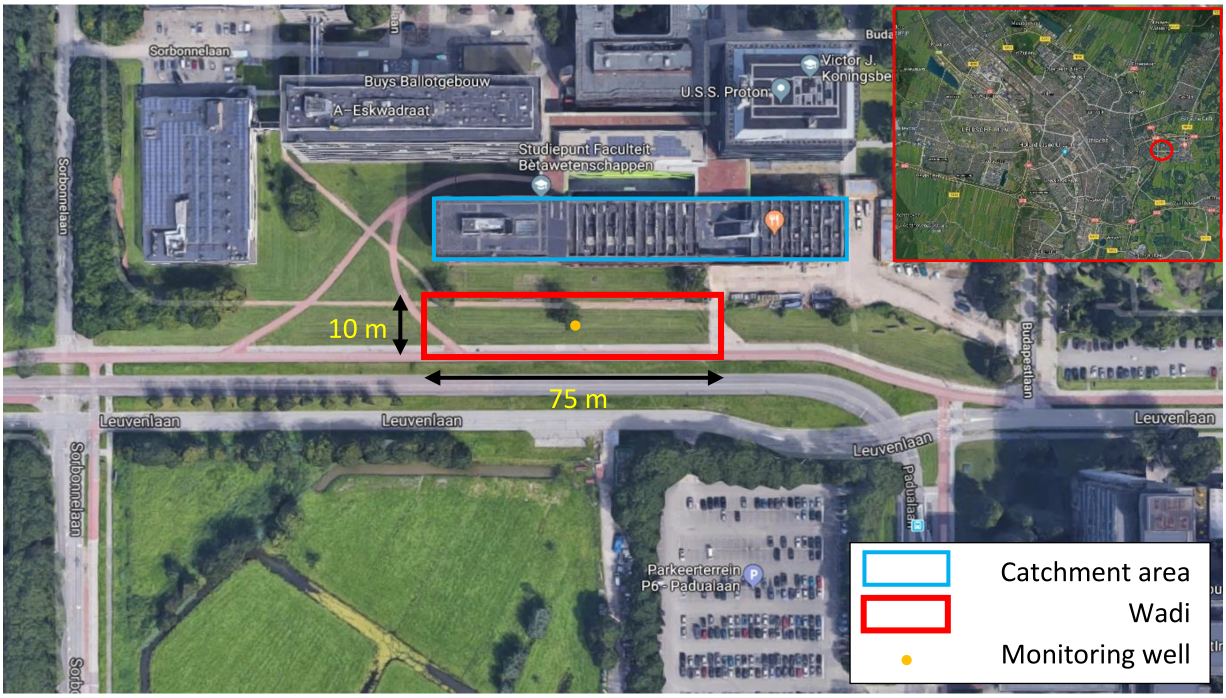


Figure 9 Location of wadi at the Leuvenlaan

Klifrakplantsoen

The Klifrakplantsoen is located in the neighbourhood Langerak, in the western part of Utrecht (Figure 10). This neighbourhood is built around 2000. The wadi is around 325 meters long and is divided into three equal parts of around 108 meters. Each part has a width of around 6 meters. The part of the wadi which is monitored is designed to collect water from the surrounding surface with a total of 6.600 square meters. The wadi is equipped with a drainage pipe which is constructed one meter below surface level. This drainage pipe discharges the infiltrated water on the "Langeraksingel" north of the wadi. When ponding depths are above a maximum of 30 centimetres, the water is able to overflow to the sewage system.

Figure 10 Location of wadi at the Klifrakplantsoen

Karel Doormanlaan

The construction of the wadi at the Karel Doormanlaan, in the northeast of Utrecht, was finished at the end of 2018 (Figure 11). This wadi has a total length of around 230 meters and a width of 4,5 meters in which 8.430 square meters of the surrounding area is collected. The wadi at the Karel Doormanlaan is, like the wadi at the Klifrakplantsoen, equipped with a drainage pipe. This drainage pipe is constructed 90 centimetres below the surface and is aimed to increase infiltration to the first aquifer layer, which is 80 centimetres below the drainage pipe. When ponding depths become above the maximum of 30 centimetres, an overflow to the sewage system occurs. A design drawing of the wadi is added in Annex III.

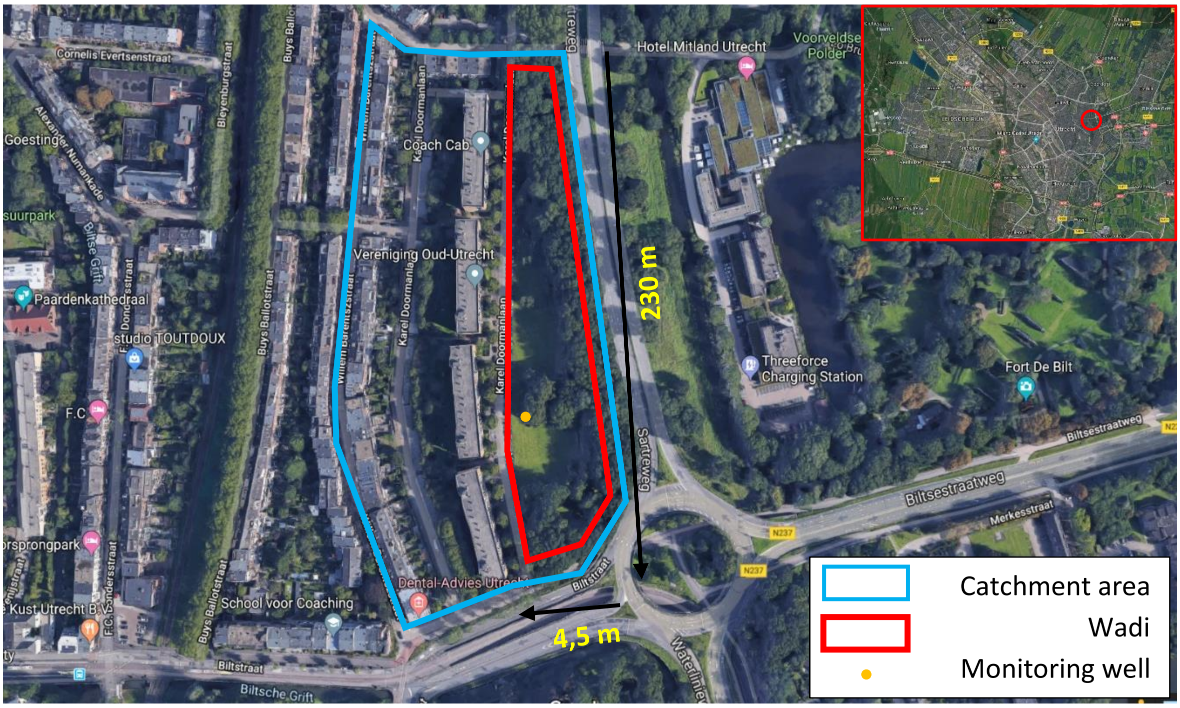


Figure 11 Location of wadi at the Karel Doormanlaan

3.3 Data collection

The following reference data is collected:

- Rainfall intensity per hour;
- Groundwater level below the wadi;
- Porosity;
- Hydraulic conductivity;
    - Saturated conductivity;
    - Unsaturated conductivity.

The following paragraphs explain the methodology of collecting reference data.

Rainfall amount and intensity

Timeseries of rainfall intensities in millimetres per hour a temporal resolution of 10 minutes is collected by data of weather stations from the website WOW-KNMI. On the Weather Observations Website (WOW) KNMI, data of private and governmental weather stations are shown with time intervals of ten to five minutes (KNMI, 2019). Due to the spatial variability of rainfall, only weather stations within a radius of 1500 meters of the wadi are selected. Rainfall data for the Klifrakplantsoen is downloaded from weather station “Hoge Weide”. For the Karel Doormanlaan, weather station “Goedeweer” is selected and the rainfall data for the Leuvenlaan comes from weather station “WnR Utrecht”. Locations of each weather station are shown in Figure 12. Each week, an Excel file with precipitation data of the week before is downloaded from the website WOW-KNMI. For example, on 15 April 2019 the data of 8 April 2019 from 00:00 to 14 April 23:59 is downloaded.

Figure 12 Locations of weather stations Hoge Weide (red), Goedeweer (blue) and WnR Utrecht (orange)

The downloaded data from the weather stations is compared with measurements of the KNMI to check on major deviations. This is done with a corrected and validated climatological radar dataset with 1 hourly precipitation values on a grid of one square kilometre of the KNMI (KNMI, 2008).

Groundwater level

Groundwater levels are collected with a monitoring well and a pressure sensor. The monitoring well is placed according to the protocol of Stichting Toegepast Onderzoek Water (STOWA). First, a borehole is made with an Edelman auger to a depth of 1 meter below the groundwater level. In this borehole, a PVC monitoring well with a diameter of 32 millimetres is placed. This monitoring well contains a filter at the bottom so water can reach the automatic pressure sensor (Stichting Toegepast Onderzoek Water, 2012). The automatic pressure sensor is placed 20 centimetres above the filter to prevent influence of dirt at the bottom of the borehole on the measurements. The automatic pressure sensor is set to measure the water pressure every 5 minutes.

After installing the pressure sensor, the following four values are measured:

- Depth of the monitoring well (D) in centimetre;
- Depth of the diver or cable length (K) in centimetre;
- Groundwater level (W) in centimetre;
- Height between top of the monitoring well and surface level (H) in centimetre.

A visual schematization of the set-up for the monitoring well and corresponding values can be found in Figure 13.

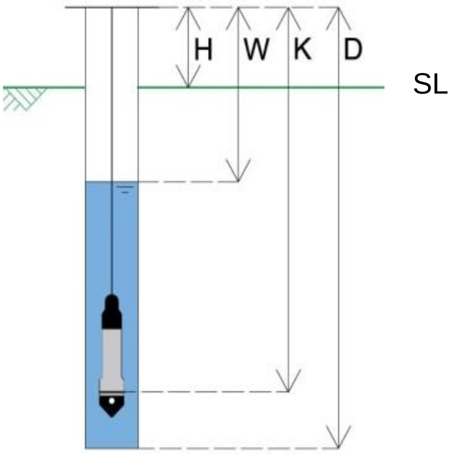


Figure 13 Schematic visualization of monitoring well

After installing the pressure sensor, the monitoring well is closed with a protective cover to prevent damage or theft. To determine the surface level relative to NAP, data of the Actueel Hoogtebestand Nederland (AHN) is used. The AHN is a detailed elevation map with an average of eight measurements per square meter. (Actueel Hoogtebestand Nederland, 2019)

Converting the pressure of the sensor to a groundwater level in meters is done by the equation:

$$GWL = SL + H - K + WC$$

In which:

SL = surface level in meters;
H = height between surface level and top of monitoring well;
K = distance between top of monitoring well and pressure sensor;
WC = water column in meters.

The water column above the pressure sensor can be calculated with the equation below:

$$WC = \rho * g * \frac{P_{diver} - P_{air}}{\rho * g}$$

In which:

ρ = density of groundwater, can be assumed as 1000 kg/dm^3;
g = gravitational acceleration, can be assumed as 9,81 m/s^2;
P_{diver} = Measured pressure by pressure sensor;
P_{air} = Pressure at surface level from the weather station.

To validate if the pressure sensor is functioning as expected, the observed groundwater level is compared with the groundwater level calculated by the pressure sensor. The observed groundwater level is measured by a measuring tape with a soil moisture sensor. The soil moisture sensor produces a high tone when it has contact with the groundwater. The distance from the top of the monitoring well to the groundwater level is then read from the measuring tape.

Ponding depth

The ponding depth in the wadi is monitored with a similar pressure sensor for groundwater level observations. The pressure sensor is put in the middle of the wadi at the surface. The ponding depth is calculated with the same equation as the groundwater level. To prevent damage or theft of the pressure sensor, a small protective cover that is attached to the monitoring well is used.

Porosity

During the drilling of the boreholes for the monitoring wells, soil is extracted from the ground. Based on the extracted soil, an estimation of the soil type can be made. The soil type can be estimated with the help of the USDA textural triangle, see Figure 14.

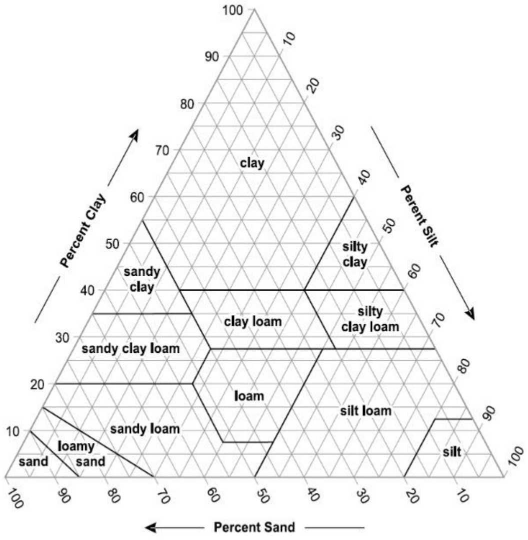
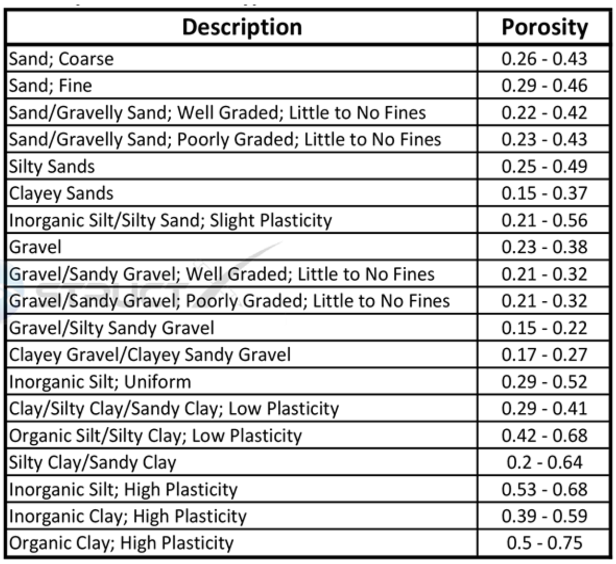


Figure 14 USDA Textural triangle to identify soil type (United States Department of Agriculture, 2017)

Table 4 Typical porosity values for various soil types (StructX, 2019)



Description	Porosity
Sand; Coarse	0.26 - 0.43
Sand; Fine	0.29 - 0.46
Sand/Gravelly Sand; Well Graded; Little to No Fines	0.22 - 0.42
Sand/Gravelly Sand; Poorly Graded; Little to No Fines	0.23 - 0.43
Silty Sands	0.25 - 0.49
Clayey Sands	0.15 - 0.37
Inorganic Silt/Silty Sand; Slight Plasticity	0.21 - 0.56
Gravel	0.23 - 0.38
Gravel/Sandy Gravel; Well Graded; Little to No Fines	0.21 - 0.32
Gravel/Sandy Gravel; Poorly Graded; Little to No Fines	0.21 - 0.32
Gravel/Silty Sandy Gravel	0.15 - 0.22
Clayey Gravel/Clayey Sandy Gravel	0.17 - 0.27
Inorganic Silt; Uniform	0.29 - 0.52
Clay/Silty Clay/Sandy Clay; Low Plasticity	0.29 - 0.41
Organic Silt/Silty Clay; Low Plasticity	0.42 - 0.68
Silty Clay/Sandy Clay	0.2 - 0.64
Inorganic Silt; High Plasticity	0.53 - 0.68
Inorganic Clay; High Plasticity	0.39 - 0.59
Organic Clay; High Plasticity	0.5 - 0.75

With the soil type, the porosity is determined with the values in Table 4. This table shows the typical minimum and maximum values for different soil types.

Saturated hydraulic conductivity

The saturated hydraulic conductivity is determined with the inversed auger-hole method (also known as the “Porchet method”) as described by Ritzema (2006). With the inversed auger-hole method, a borehole with a diameter of 25 millimetres is made to a drilling depth of about 60 or 70 centimetres below the water table (Ritzema, 2006). A schematization of the inversed auger-hole method setup can be found in Figure 15.

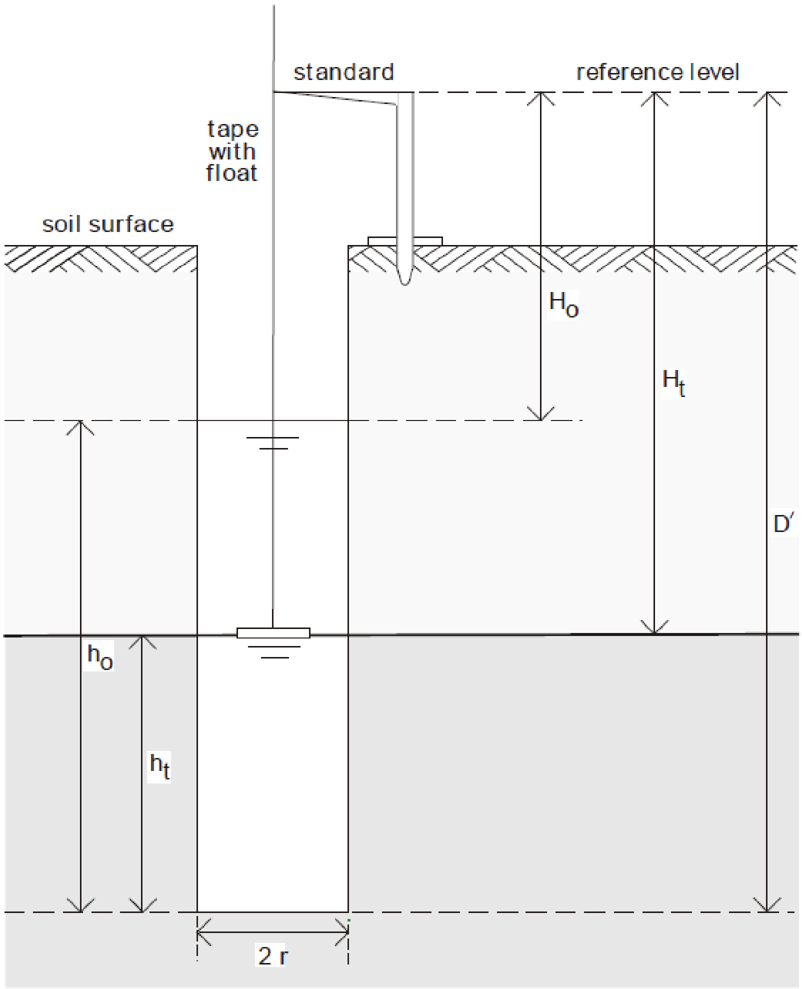


Figure 15 Schematization of Auger-method test setup (Ritzema, 2006)

After adding 20 to 40 centimetres of water to the borehole, the lowering of the water table is measured with a pressure sensor and time intervals of 5 seconds. The saturated conductivity (K_{sat}) is calculated with the equation (Ritzema, 2006):

$$K_{sat} = 1{,}15 * r * \frac{\log\left(h_0 + \frac{1}{2}r\right) - \log\left(h_t + \frac{1}{2}r\right)}{t - t_0}$$

In which:
h_0 = Water level at start of measurements [cm];
h_t = Water level at a certain time of measurements [cm];
t = time of measurements [sec];
t_0 = time at the start of measurement [sec];
r = radius of borehole [cm].

Unsaturated hydraulic conductivity

The unsaturated hydraulic conductivity is determined by a double ring infiltrometer test. A double-ring infiltrometer contains two rings with a different diameter, a bigger ring on the outside and a smaller ring inside (Figure 16).

Figure 16 Example of a double ring infiltrometer

By filling the outer ring with water the effect of capillary pressure and gravity forces on the infiltration rate is reduced (Dingman, 2015). The infiltrometer is installed 5 to 10 centimetres into the ground with the top horizontal. Then a pressure sensor is installed a few centimetres above the bottom of the inner ring and the outer ring is filled completely with water. Next, the inner ring is filled with water and the time and level of the pressure sensor are notated. When the water in the inner ring is drained, the outer and inner ring are again filled with water and the time is notated. A schematization of the double-ring infiltrometer test is shown in Figure 17.

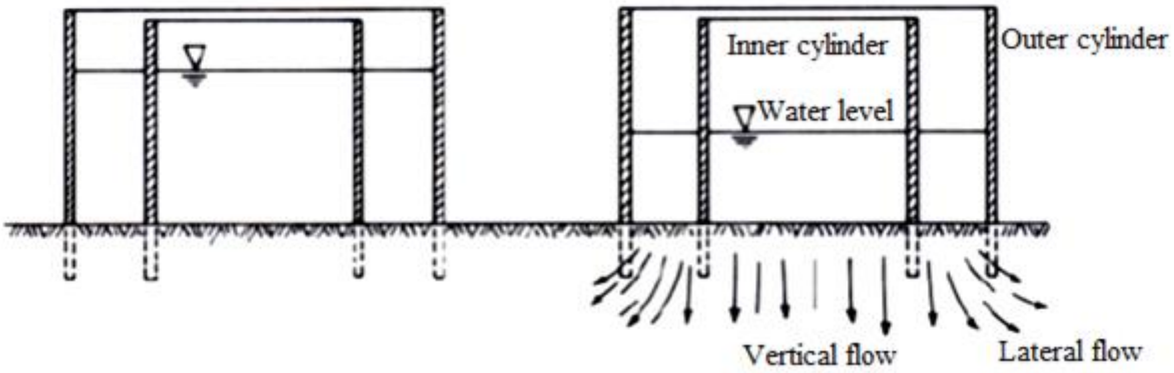


Figure 17 Schematization of double-ring infiltrometer test. Retrieved from http://ecoursesonline.iasri.res.in/mod/page/view.php?id=1994

With the data of the pressure sensor, the hydraulic conductivity of the unsaturated zone (K_{unsat}) in centimetres can derived. This is done by the equation:

$$K_{unsat} = \frac{(h_0 - h_1)}{(t_1 - t_0)}$$

In which:
h_0 = Water level at start of measurements [cm];
h_1 = Water level at the end of measurements [cm];
t_1 = time at the end of measurements [sec];
t_0 = time at the start of measurement [sec].

3.4 Model calibration

After collecting data on rainfall intensity, groundwater levels, porosity and hydraulic conductivity, the model is calibrated. The model is calibrated by using the trial and error method. In this method, parameters are changed and the output is then compared to the observed values of groundwater elevation and ponding depths. The values of saturated and unsaturated conductivity which are measured on the research locations are considered as reliable input values for the model and are not changed during model calibration. Because there is some uncertainty on the porosity, this parameter is changed within the range for each soil type. Other parameters that were not measured in the field are values for wilting point and field capacity. These values are changed within a range which is, according to scientific literature, likely to occur at the research locations. The process of changing parameters is repeated until the output of the model is close to the observed values.

3.5 Model validation

To determine whether the computed output of the model is sufficiently close to the observed output, the systematic approach of Nash and Sutcliffe is used (Nash & Sutcliffe, 1970). The approach of Nash & Sutcliffe to evaluate the relative efficiency of a model is done by the following three equations. The equation below is used to determine the index of disagreement.

$$F^2 = \sum (S - O)^2$$

In which:

F^2 = index of disagreement;
S = computed values at corresponding times;
O = observed values at corresponding times.

F^2 in the equation above is analogous to the residual variance of regression analysis.

The initial variance is defined by the equation:

$$F_0^2 = \sum (O - \bar{O})^2$$

In this equation are:

O = observed value at corresponding times;
$\bar{O}$ = mean observed value.

Combining the index of disagreement with the initial variance results in a new equation. This equation defines the relative efficiency of a model. The equation is stated as:

$$R^2 = \frac{F_0^2 - F^2}{F_0^2}$$

When R^2 is 1, the validation implies that the model has a perfect fit compared to the observed values. When $0,65 < R^2 < 0,8$ the model is considered acceptable. Other values of the Nash-Sutcliffe model efficiency can be found in Table 5 (Ritter & Muñoz-Carpena, 2013).

Table 5 Rating of model efficiency

Rating	**NSE**
Very good	> 0,9
Good	0,8 – 0,9
Acceptable	0,65 – 0,8
Unsatisfactory	< 0,65

The Nash-Sutcliffe model efficiency will be used to determine the accuracy of the infiltration model of a wadi. Groundwater elevations and ponding depths of the wadi will be compared to evaluate the model efficiency. Validation of the model is done with an independent rainfall event, in the same order size of the design precipitation.

3.6 Sensitivity analysis

After model validation, a sensitivity analysis of the individual parameters of the model is done. With this sensitivity analysis, the parameters of porosity, saturated hydraulic conductivity, unsaturated conductivity and field capacity were analyzed to what extent these parameters influence the model outcome and which values are essential for a reliable model outcome. This is useful because values such as soil type and porosity are usually determined by the interpretation of the field workers.

The sensitivity analysis is done by the one-at-a-time method. With this method, repeatedly one parameter is changed while the other parameters kept the same (Hamby, 1994). The parameters of porosity, saturated conductivity, unsaturated conductivity and field capacity are changed by a percentage ranging from -30% to 30% from the calibrated value. The model output of these changes are visualized in a graph in which the percentage of change of the maximum groundwater level is shown on the vertical axis, the change in the parameter value is shown on the horizontal axis. The same type of graph is made for the change in mean groundwater level.

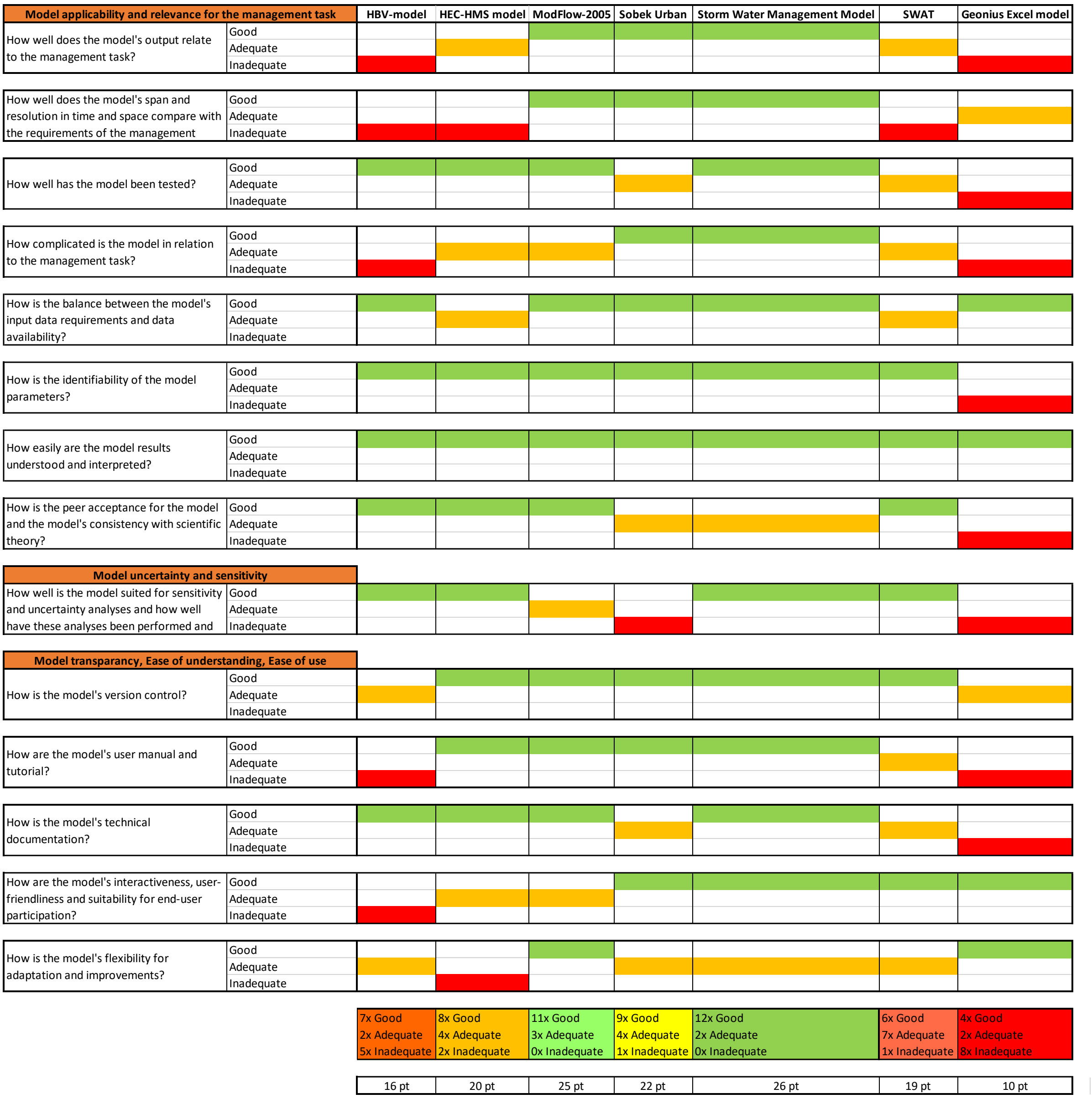
3.7 Climate durability

Regulations in the municipality of Utrecht assume the increase in precipitation according to climate scenario G for the year 2050 (Gemeente Utrecht, 2011). The climate durability of the wadies are assessed for a precipitation event with a return time of once in 100 years. The precipitation event for climate scenario G in 2050 has a duration of 48 hours in which 109 millimetres of rainfall occurs (STOWA, 2015). Design precipitation for timesteps of ten minutes is derived from the percentile statistics of the STOWA (STOWA, 2004b). A design precipitation event with a percentile of 87,5% is used to assess climate durability. The percentile statistics of different characteristic precipitation patterns are added in Annex III. By multiplying the percentile of each timestep with the total rainfall amount, the amount of rainfall per 10 minutes can be derived. To assess the climate durability of the wadies at the research locations, the validated model is used with the precipitation data of the percentile statistics.

4. Results

4.1 Selected models

To determine which model is fit-for-purpose, seven different models are examined. The requirements of the model to meet the management task are specified in Annex I. Figure 18 shows the results of the benchmark criteria for the different models and the total amount of points per model. Extensive answers and explanations of each question for the different models can be found in Annex V.



Model applicability and relevance for the management task		HBV-model	HEC-HMS model	ModFlow-2005	Sobek Urban	Storm Water Management Model	SWAT	Geonius Excel model
How well does the model's output relate to the management task?	Good			🟩	🟩	🟩		
	Adequate		🟧				🟧	
	Inadequate	🟥						🟥
How well does the model's span and resolution in time and space compare with the requirements of the management	Good			🟩	🟩	🟩		
	Adequate							🟧
	Inadequate	🟥	🟥				🟥	
How well has the model been tested?	Good	🟩	🟩	🟩		🟩		
	Adequate				🟧		🟧	
	Inadequate							🟥
How complicated is the model in relation to the management task?	Good				🟩	🟩		
	Adequate		🟧	🟧			🟧	
	Inadequate	🟥						🟥
How is the balance between the model's input data requirements and data availability?	Good	🟩		🟩	🟩	🟩		🟩
	Adequate		🟧				🟧	
	Inadequate							
How is the identifiability of the model parameters?	Good	🟩	🟩	🟩	🟩	🟩	🟩	
	Adequate							
	Inadequate							🟥
How easily are the model results understood and interpreted?	Good	🟩	🟩	🟩	🟩	🟩	🟩	🟩
	Adequate							
	Inadequate							
How is the peer acceptance for the model and the model's consistency with scientific theory?	Good	🟩	🟩	🟩			🟩	
	Adequate				🟧	🟧		
	Inadequate							🟥
Model uncertainty and sensitivity								
How well is the model suited for sensitivity and uncertainty analyses and how well have these analyses been performed and	Good	🟩	🟩			🟩	🟩	
	Adequate			🟧				
	Inadequate				🟥			🟥
Model transparancy, Ease of understanding, Ease of use								
How is the model's version control?	Good		🟩	🟩	🟩	🟩	🟩	
	Adequate	🟧						🟧
	Inadequate							
How are the model's user manual and tutorial?	Good		🟩	🟩	🟩	🟩		
	Adequate						🟧	
	Inadequate	🟥						🟥
How is the model's technical documentation?	Good	🟩	🟩	🟩		🟩		
	Adequate				🟧		🟧	
	Inadequate							🟥
How are the model's interactiveness, user-friendliness and suitability for end-user participation?	Good				🟩	🟩	🟩	🟩
	Adequate		🟧	🟧				
	Inadequate	🟥						
How is the model's flexibility for adaptation and improvements?	Good			🟩				🟩
	Adequate	🟧			🟧	🟧	🟧	
	Inadequate		🟥					
		7x Good 2x Adequate 5x Inadequate	8x Good 4x Adequate 2x Inadequate	11x Good 3x Adequate 0x Inadequate	9x Good 4x Adequate 1x Inadequate	12x Good 2x Adequate 0x Inadequate	6x Good 7x Adequate 1x Inadequate	4x Good 2x Adequate 8x Inadequate
		16 pt	20 pt	25 pt	22 pt	26 pt	19 pt	10 pt

Figure 18 Result of benchmark criteria for different models

The Storm Water Management Model (SWMM) from the Environmental Protection Agency is considered most fit-for-purpose by existing literature. To determine the accuracy of the Storm Water Management Model in relation to groundwater elevation and ponding depths, this model is used for three case studies.

4.2 Field measurements

Field reference data on groundwater elevation levels, ponding depths and rainfall intensity were collected between the 16th of May 2019 and 28th of July 2019. Model input values for saturated conductivity and unsaturated conductivity are measured with the Porchet method and a double ring infiltrometer once.

Leuvenlaan

The groundwater levels and rainfall intensity at the Leuvenlaan are monitored from the 24th of May till the 23rd of June (Figure 19). The first thing which can be noticed at the groundwater level is a certain regularity in which higher and lower groundwater levels occur. The average groundwater level at the Leuvenlaan is 70 cm above NAP (Gemeente Utrecht, 2013). This suggests that the groundwater level is lowered instead of sudden increases. About every two days, the groundwater is lowered between 40 cm above NAP and 50 cm above NAP. It is unknown what causes the lowering of these groundwater levels.

When looking at the peak groundwater levels, it is noticeable that rainfall increases the groundwater level. For the rainfall events on 27th of May, 4th of June and 12th of June the groundwater level increases with around 10 cm compared to the average groundwater level. The rainfall intensity on the 19th of June causes even a higher increase in groundwater level, which resulted in a groundwater level of 105 cm NAP.

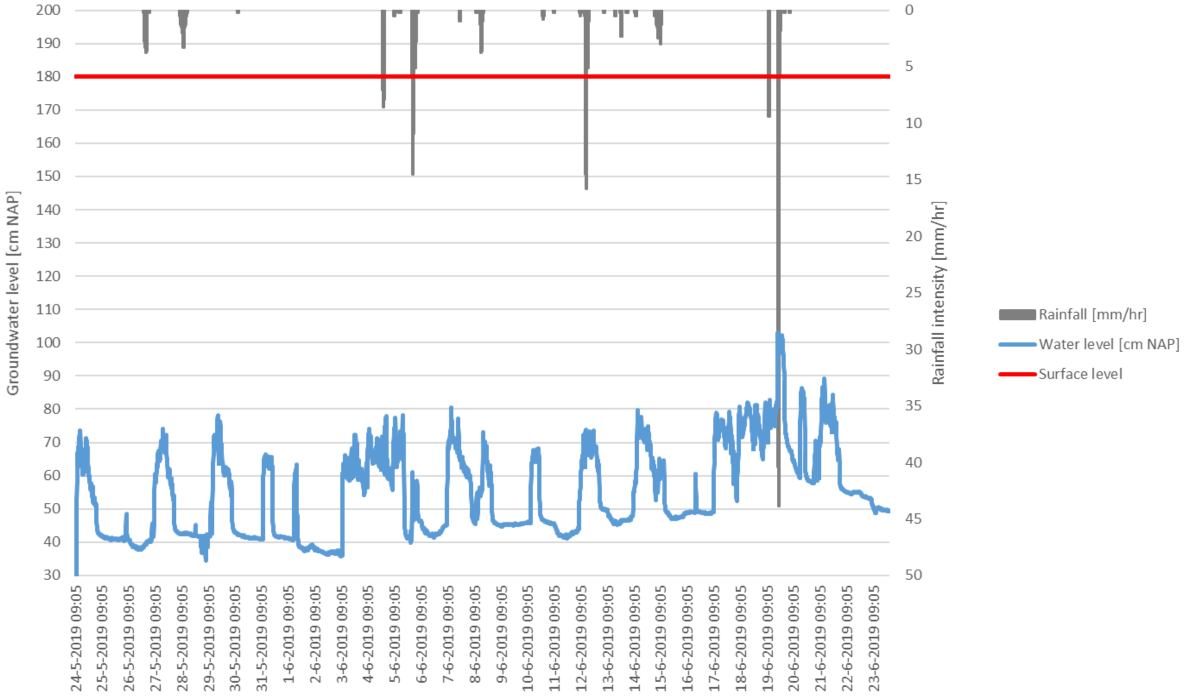


Figure 19 Groundwater levels and rainfall intensity in the period 24-05-2019 to 23-06-2019 at the Leuvenlaan

The ponding depths at the Leuvenlaan over the monitoring period are shown in Figure 20. Only the intense rainfall on the 19th of June resulted in ponding of about 3.5 centimetres.

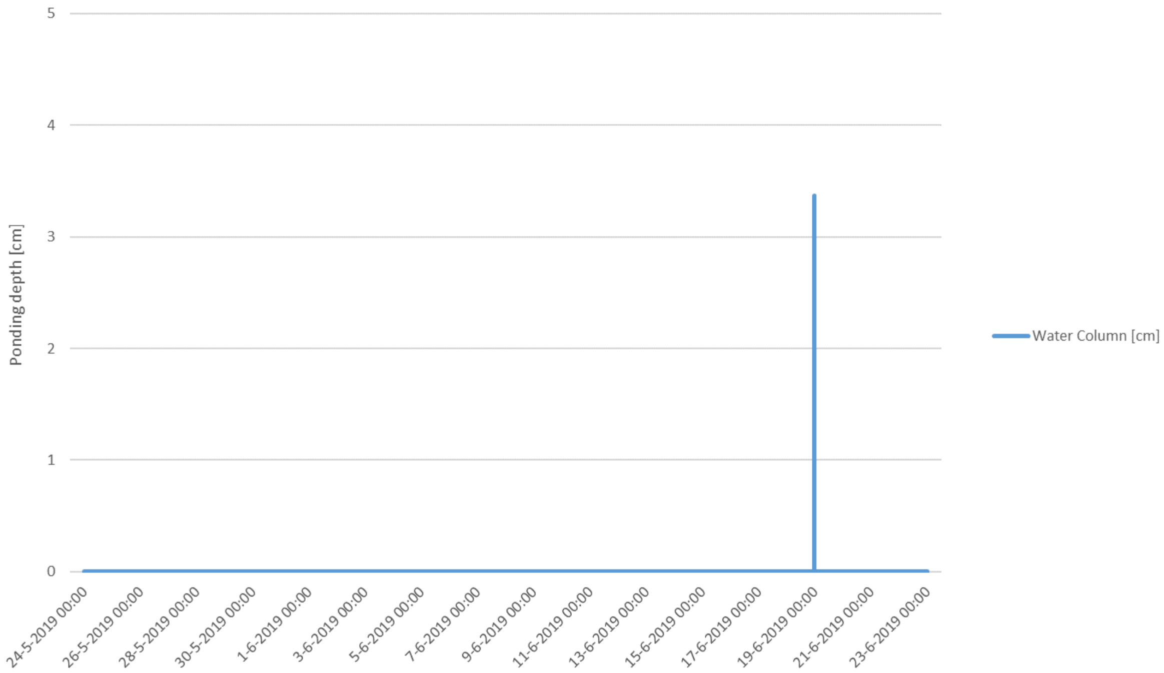


Figure 20 Ponding depths in the period 24-05-2019 to 23-06-2019 at the Leuvenlaan

Because of external influences in groundwater levels at the Leuvenlaan, it is impossible to make a model within the time of this research. Therefore, no values for saturated and unsaturated conductivity were measured. Determining what external forces influence these changes in groundwater levels is outside the scope of this research.

Klifrakplantsoen

The groundwater levels and rainfall intensity at the Klifrakplantsoen are monitored from the 16th of May till the 28th of July (Figure 21). The average groundwater level at the Klifrakplantsoen is between -30 cm NAP and -20 cm NAP (Gemeente Utrecht, 2013).

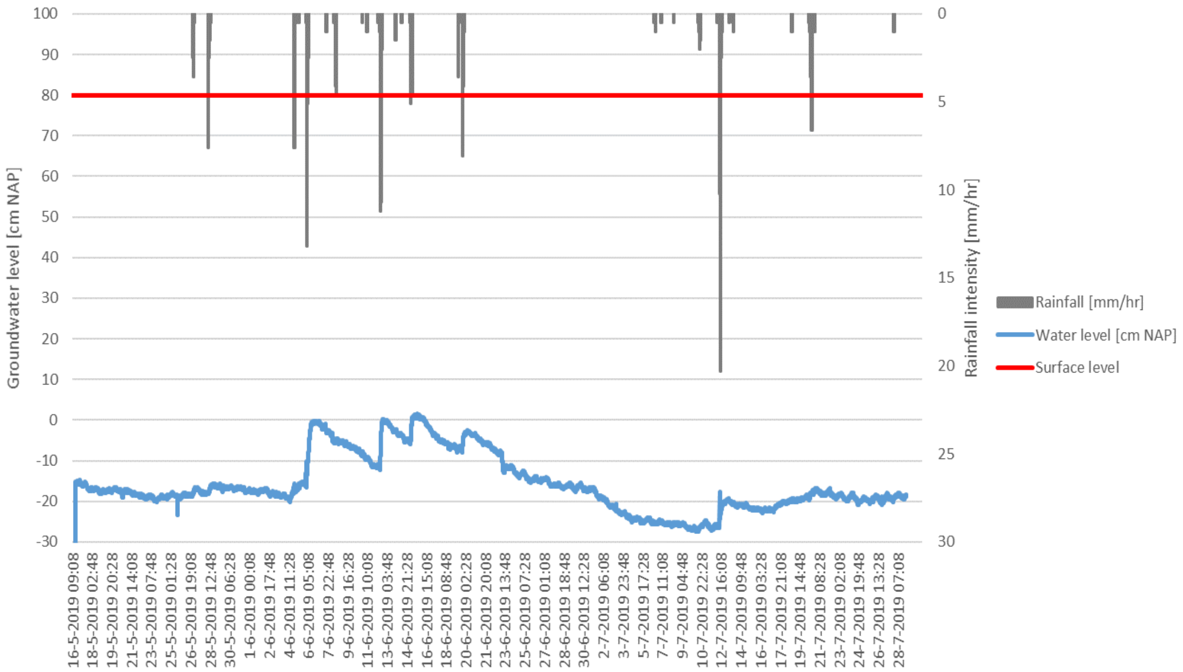


Figure 21 Groundwater level and rainfall intensity in the period 16-05-2019 to 28-07-2019 at the Klifrakplantsoen

Figure 22 shows the ponding depth at the Klifrakplantsoen over the period from 16th of May till the 28th of July. It can be noticed that during this period, only on the 5th of June the surface was inundated with 0.4 cm of water.

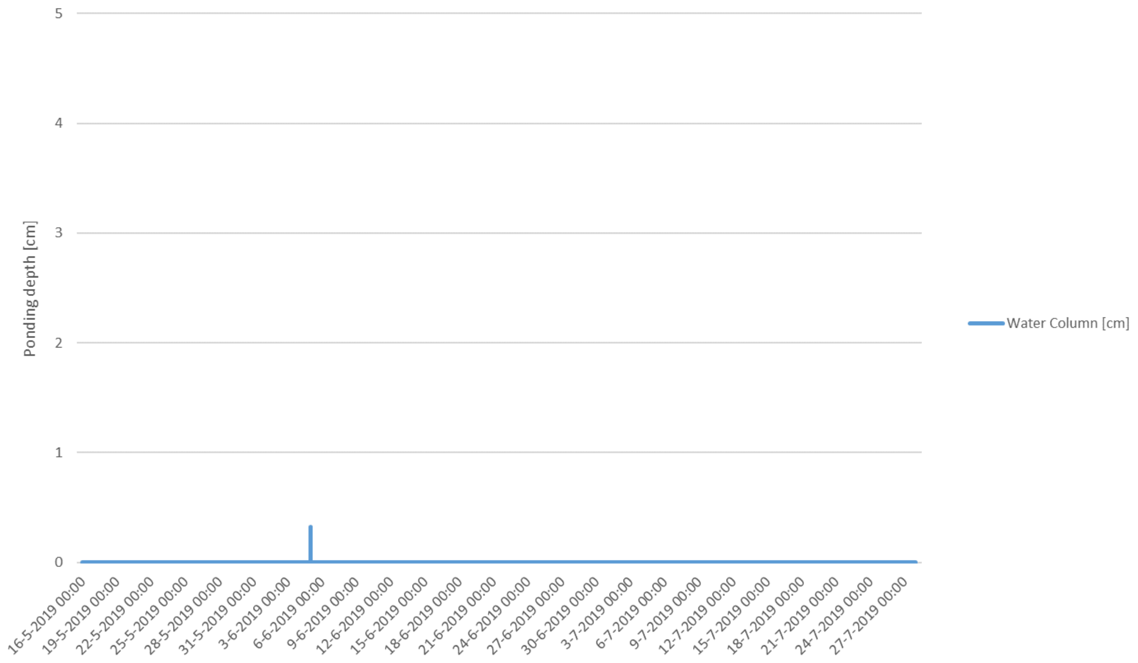


Figure 22 Ponding depths in the period 16-05-2019 to 28-07-2019 at the Klifrakplantsoen

From the groundwater levels and rainfall intensity in Figure 21, it is possible to recognize multiple heavy rainfall events. The rainfall on the 4th of June was measured with an amount of rainfall of 32 millimetres and the rainfall on the 13th of June was measured 19 millimetres. These two rainfall events are selected for model calibration. The rainfall of the 19th of June had a measured amount of 16 millimetres and is selected for model validation.

The results of the field measurements for saturated conductivity, unsaturated conductivity and porosity are shown in Table 6. The complete measurement for saturated conductivity can be found in Annex VI and the measurement of the unsaturated conductivity can be found in Annex VII. Based on the extracted soil during the drilling of the monitoring well, a layer with the soil type of fine sand and a layer with a soil type of sandy clay were found. Table 6 shows the minimum, mean and maximum values for the two layers.

Table 6 Minimum, mean and maximum values field measurements Klifrakplantsoen

	Minimum	**Mean**	**Maximum**
Saturated conductivity [m/d]	-	0.54	-
Unsaturated conductivity [m/d]	-	17.18	-
Porosity sand layer [-]	0.22	0.33	0.46
Porosity clay layer [-]	0.29	0.35	0.41

Karel Doormanlaan

Groundwater levels are monitored from the 22th of May till the 28th of July (Figure 23). The average groundwater level at the Karel Doormanlaan is between 30 cm NAP and 40 cm NAP (Gemeente Utrecht, 2013).

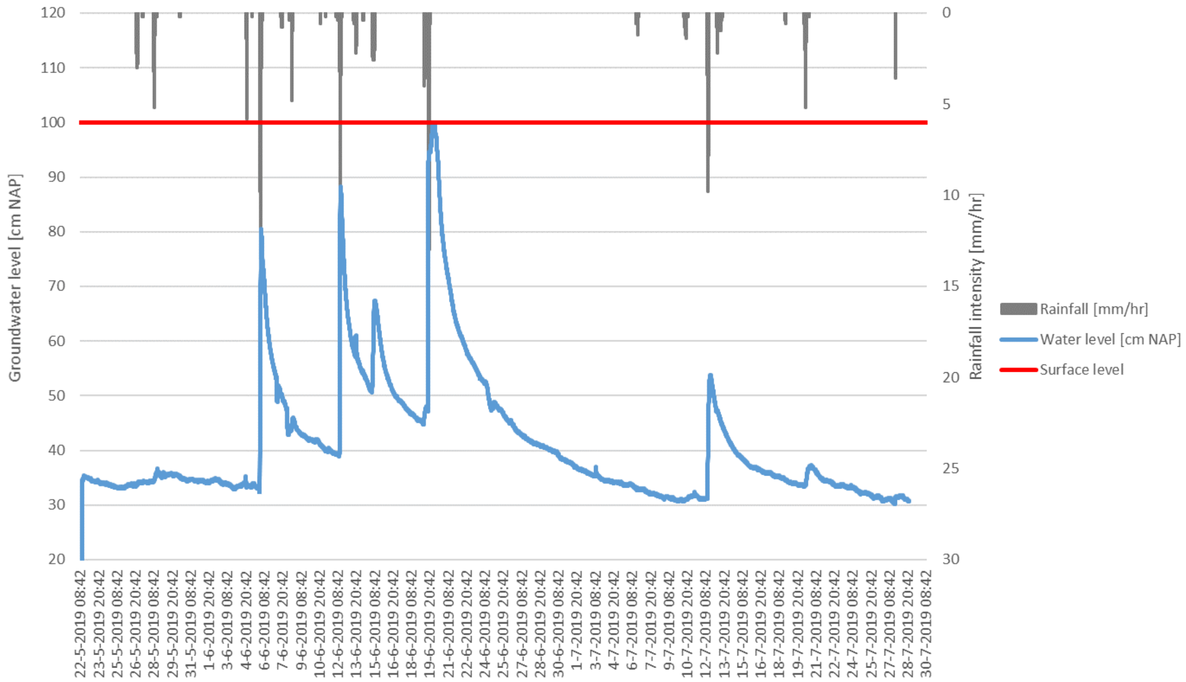


Figure 23 Groundwater level and rainfall intensity in the period 22-05-2019 to 28-07-2019 at the Karel Doormanlaan

Figure 24 shows the ponding depth at the Karel Doormanlaan over the period 22nd of May untill the 28th of July. During three rainfall events, the surface of the wadi was inundated. The maximum water depth occurred on the 19th of June and is around 20 cm.

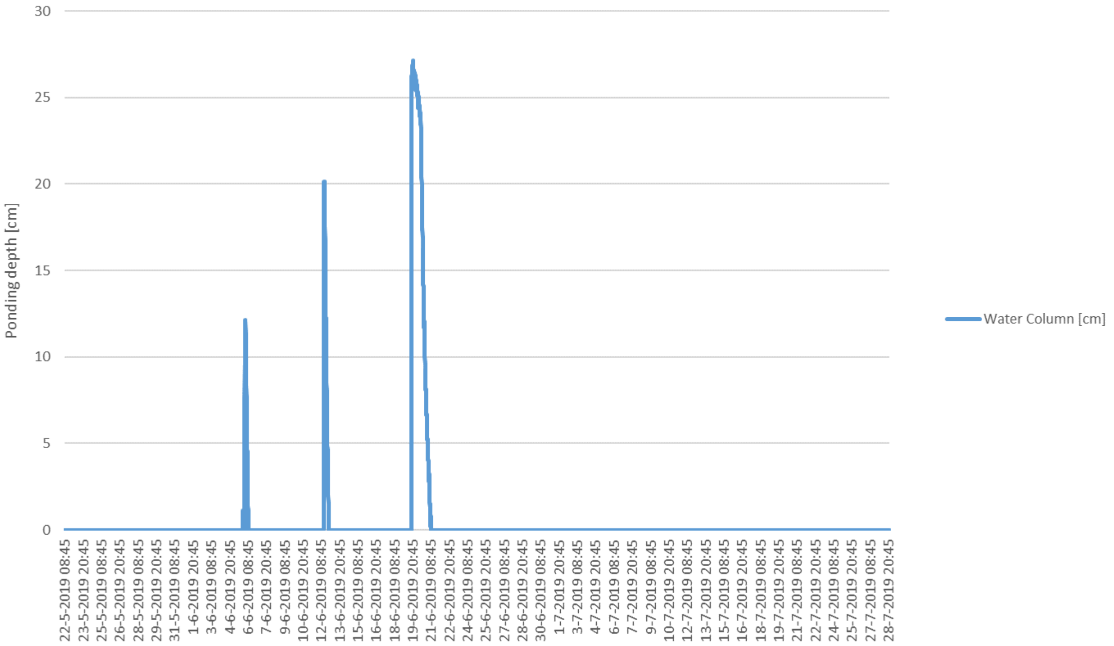


Figure 24 Ponding depth in the period 22-05-2019 to 28-07-2019 at the Karel Doormanlaan

From Figure 23, it is possible to recognize heavy rainfall events at the Karel Doormanlaan. The rainfall events which are selected for calibration at the Karel Doormanlaan occurred on the 4th of June and the 12th of June. The rainfall amount on the 4th of June was measured a total of 25 millimetres, the rainfall amount on the 12th of June was measured 17 millimetres. The rainfall event which is used for model validation occurred at the 19th of June and measured 21 millimetres.

The results of the field measurements for saturated conductivity, unsaturated conductivity and porosity are shown in Table 7. The complete measurement for saturated conductivity can be found in Annex VIII and the measurement of the unsaturated conductivity can be found in Annex IX. Based on the extracted soil during the drilling of the monitoring well, two layers of soil were found. The first layer is a unique kind of sand, specially made for infiltration purposes. This layer of sand has a big grain size and contains about 5% organic material. The precise porosity is unknown, however relative high porosity is assumed. For this research the porosity of silty sand is used. The second layer is a layer of inorganic clay with high plasticity. Table 7 shows the minimum, mean and maximum values for the two layers.

Table 7 Minimum, mean and maximum values field measurements Karel Doormanlaan

	Minimum	**Mean**	**Maximum**
Saturated conductivity [m/d]	-	0.64	-
Unsaturated conductivity [m/d]	-	35.81	-
Porosity sand layer [-]	0.25	0.37	0.49
Porosity clay layer [-]	0.39	0.49	0.59

4.3 Model calibration and validation

The SWMM model for the case studies contains two subcatchments and one aquifer. One represents the catchment area where precipitation flows into the wadi. This subcatchment is completely impermeable, so all the rainfall flows into the second subcatchment. The other subcatchment is the size of the wadi and is configured as an low impact development (LID) measure. Below both subcatchments, an aquifer is located. The LID is configured with two layers where the bottom layer has the characteristics of the aquifer. The top layer is configured with values for unsaturated conductivity and porosity of the top layer. The aquifer is configured with saturated conductivity and bottom layer porosity. The presence of a drain is indicated in the bottom layer. Field capacity and wilting point is configured the same for both layers. The model is configured to determine infiltration rates by the equation of Green-Ampt.

Klifrakplantsoen

Figure 25 shows the observed and modelled groundwater elevation for the precipitation event on the 4th of June. The calibration of this rainfall event resulted in a NSE of 0.80 on groundwater levels.

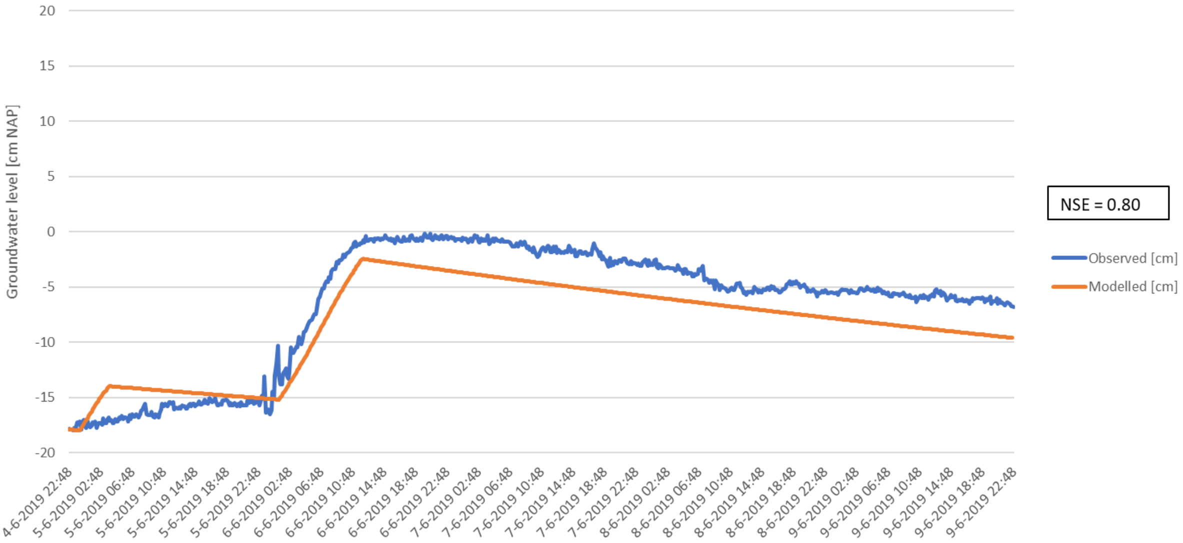


Figure 25 Observed and modelled groundwater levels Klifrakplantsoen

The ponding depth for this timeserie resulted in a NSE of -2.20 (Figure 26), which is below the required 0.65 for acceptable models. When the maximum modelled ponding depth is compared with the maximum observed ponding depth, a NSE of 0.98 is achieved.

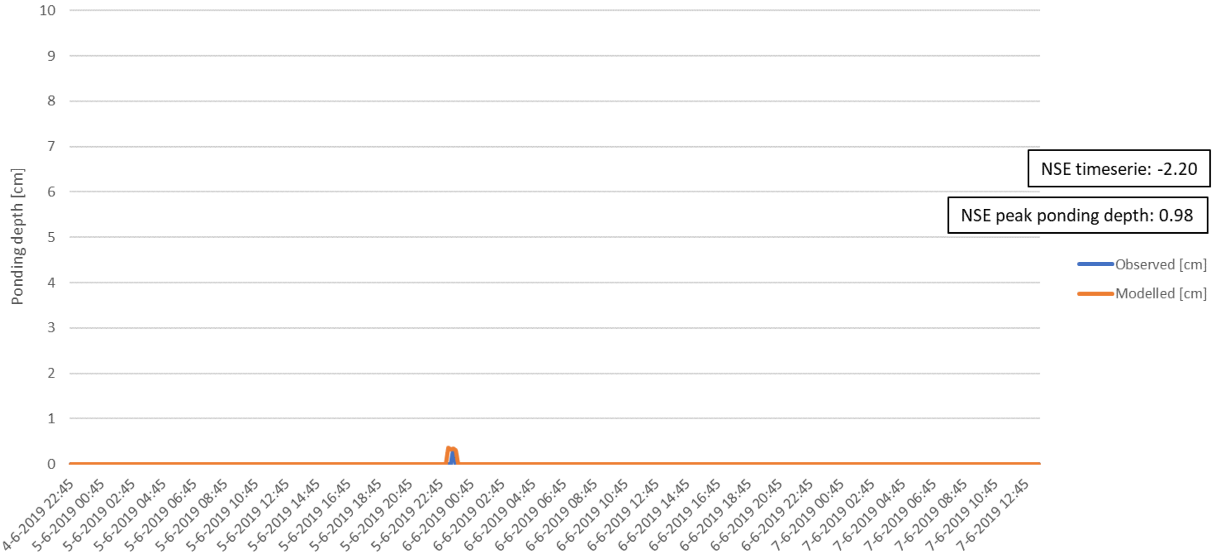


Figure 26 Observed and modelled ponding depth Klifrakplantsoen

The second rainfall event for calibration, on the 13th of June, resulted in a NSE of 0.77 in groundwater levels (Figure 27).

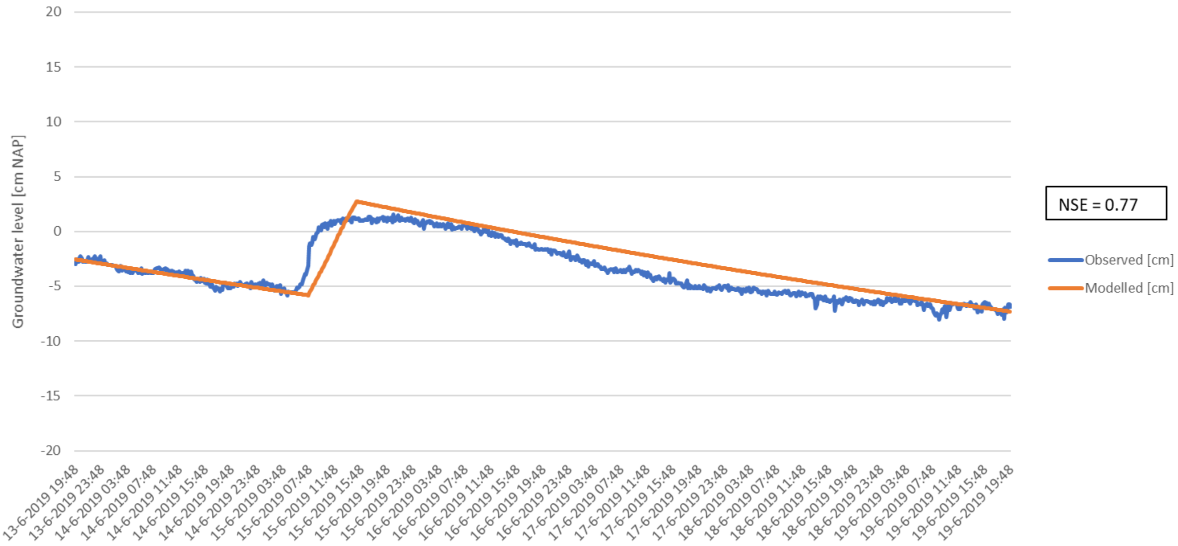


Figure 27 Observed and modelled groundwater levels Klifrakplantsoen

The NSE of ponding depth for this event is 1.0 (Figure 28). It is difficult to compare these values because no ponding was observed and modelled during this rainfall event.

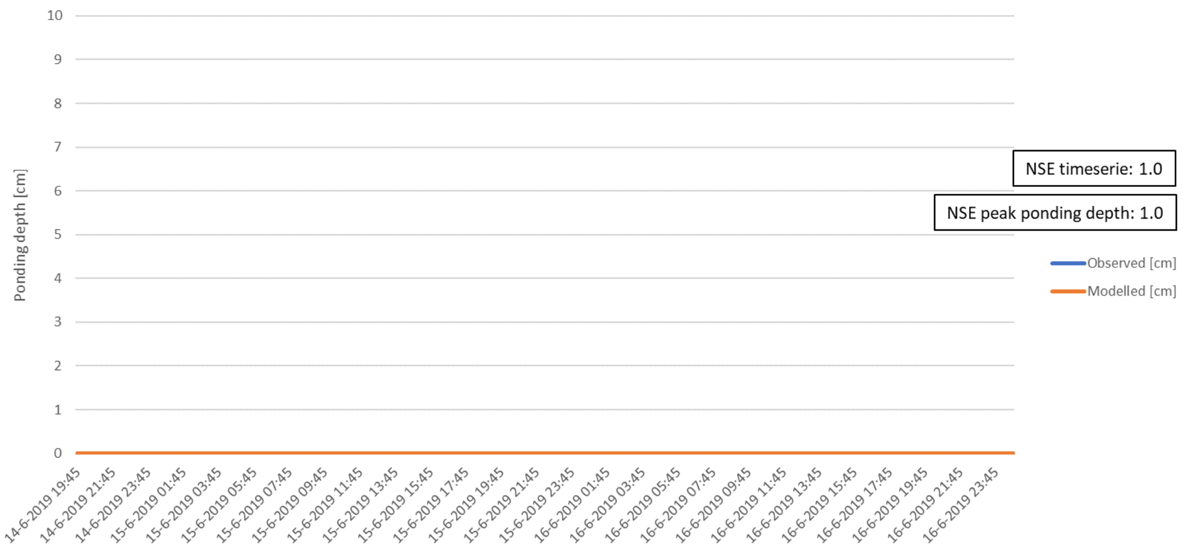


Figure 28 Observed and modelled ponding depth Klifrakplantsoen

During the calibration, the parameters of porosity, field capacity and wilting point were varied between the minimum and maximum values of Table 8. The column "Value" of Table 8 shows the values which resulted in the best accuracy in groundwater elevation and ponding depth.

Table 8 Calibrated input values Klifrakplantsoen

Parameter	**Minimum**	**Value**	**Maximum**
Porosity top [-]	0.22	**0.45**	0.46
Porosity bottom [-]	0.29	**0.4**	0.41
Saturated conductivity [mm/hr]	-	**22.5**	-
Unsaturated conductivity [mm/hr]	-	**712**	-
Field capacity [-]	0.3	**0.35**	0.4
Wilting point [-]	0.1	**0.15**	0.15

With the calibrated values from Table 8, the model is validated on the rainfall event of the 19th of June. The validation resulted in a NSE of 0.75 for groundwater levels (Figure 29).

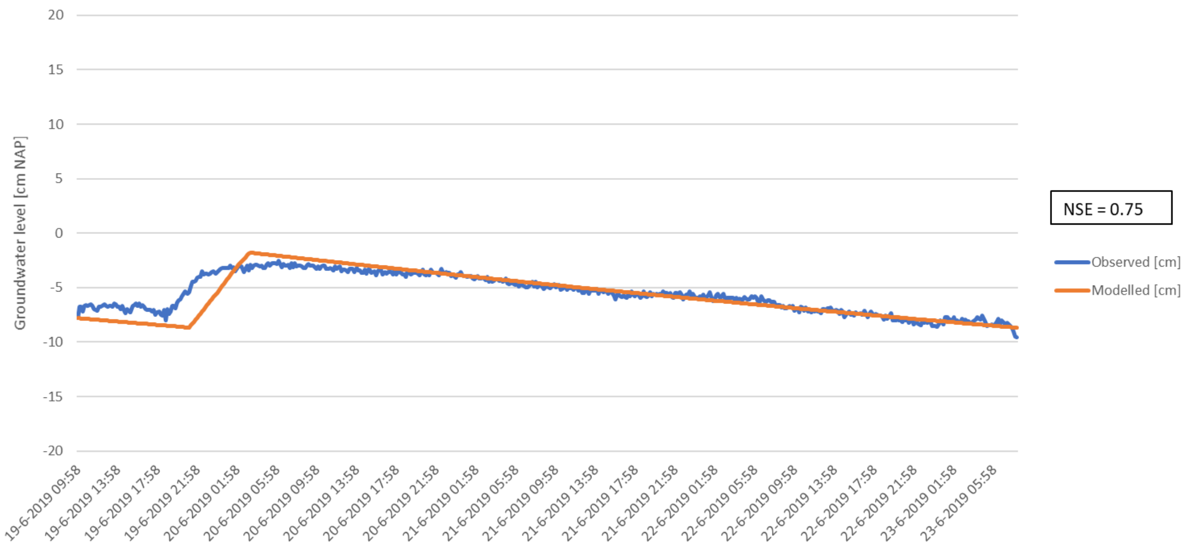


Figure 29 Observed and modelled groundwater levels Klifrakplantsoen

The validation rainfall event resulted in a NSE of 1.0 for ponding depths (Figure 30). Similar to the result of ponding depths at the rainfall event of the 13th of June, the observed and modelled ponding depths are difficult to compare because no ponding has occurred.

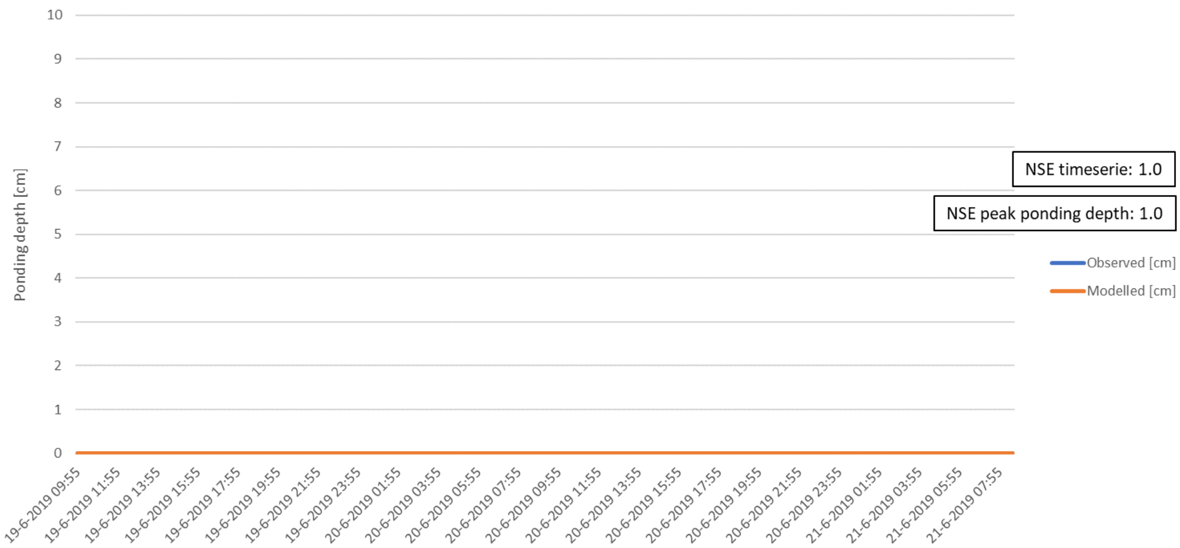


Figure 30 Observed and modelled ponding depth Klifrakplantsoen

The average NSE on groundwater levels is 0.77. With this NSE, the SWMM model is able to generate acceptable to good outputs on groundwater elevations. Based on the rainfall event on the 4th of June, a good output on maximum ponding depths can be simulated.

Karel Doormanlaan

Figure 31 shows the observed and modelled groundwater elevation for the precipitation event on the 4th of June. The calibration of this rainfall event resulted in a NSE of 0.76 on groundwater levels.

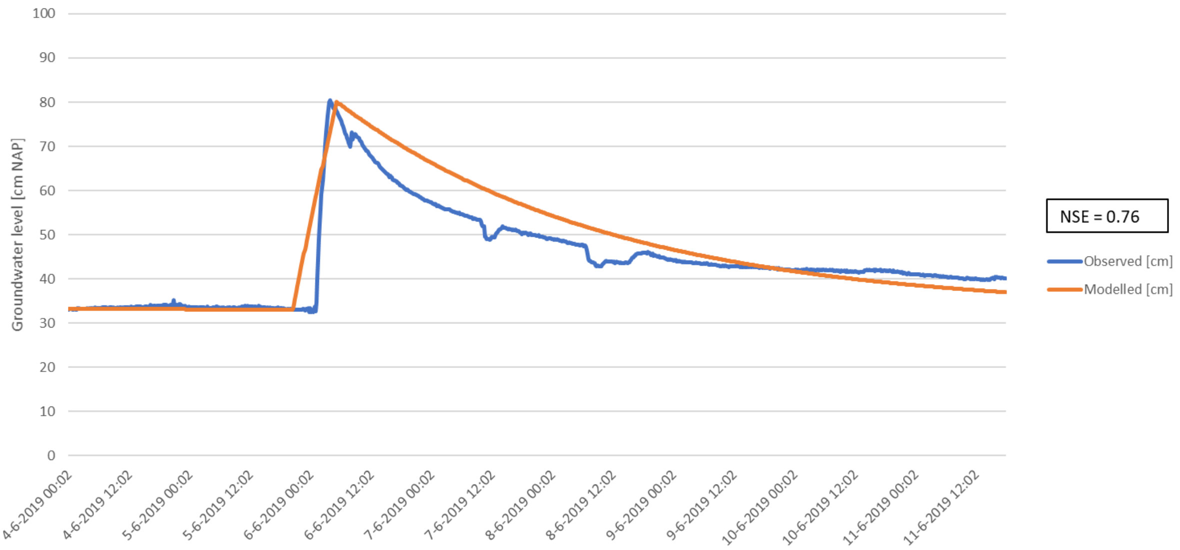


Figure 31 Observed and modelled groundwater levels Karel Doormanlaan

The ponding depth for this rainfall event resulted in a NSE of -8.47 (Figure 32), which is below the required 0.65 for acceptable models. When the maximum modelled ponding depth is compared with the maximum observed ponding depth, a NSE of -2.54 is achieved.

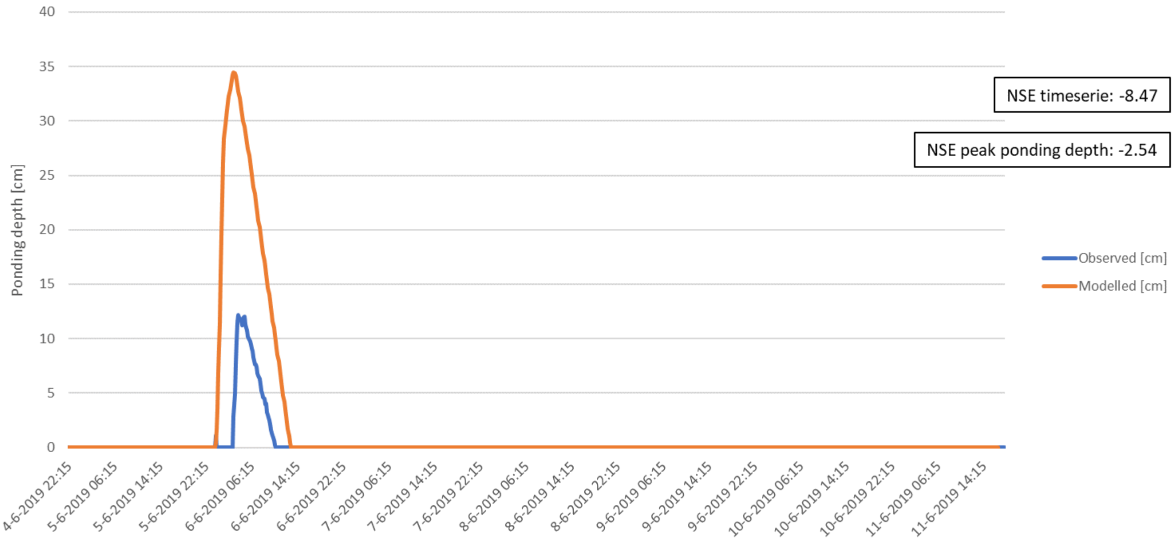


Figure 32 Observed and modelled ponding depth Karel Doormanlaan

The calibration rainfall event on the 12th of June resulted in a NSE of 0.68 in groundwater levels (Figure 33).

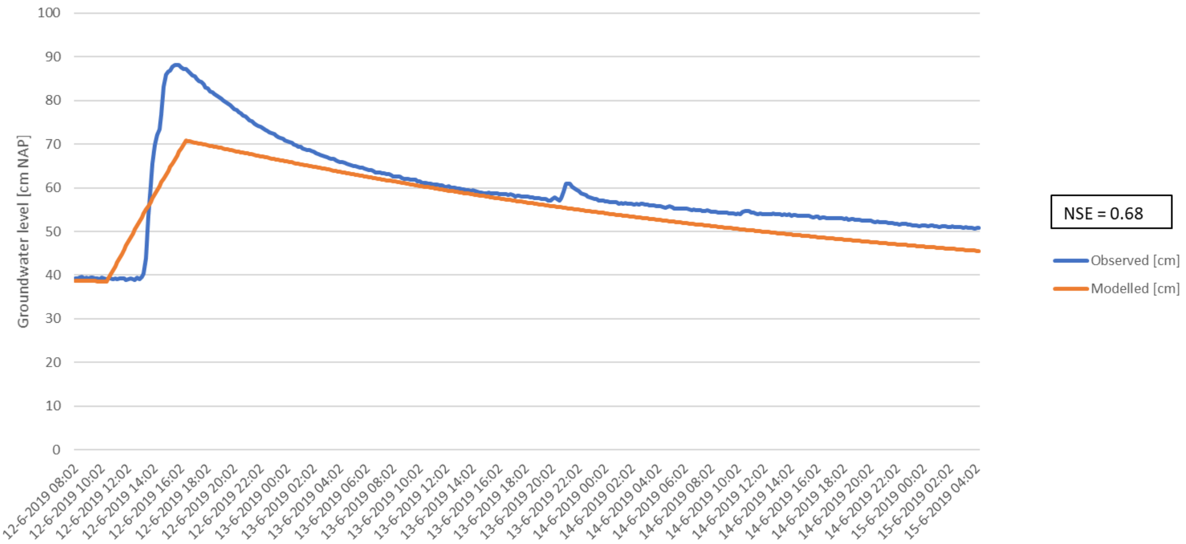


Figure 33 Observed and modelled groundwater levels Karel Doormanlaan

The NSE for the ponding depth for this rainfall event is 0.96 on the complete timeserie and 0.99 on the maximum ponding depth (Figure 34).

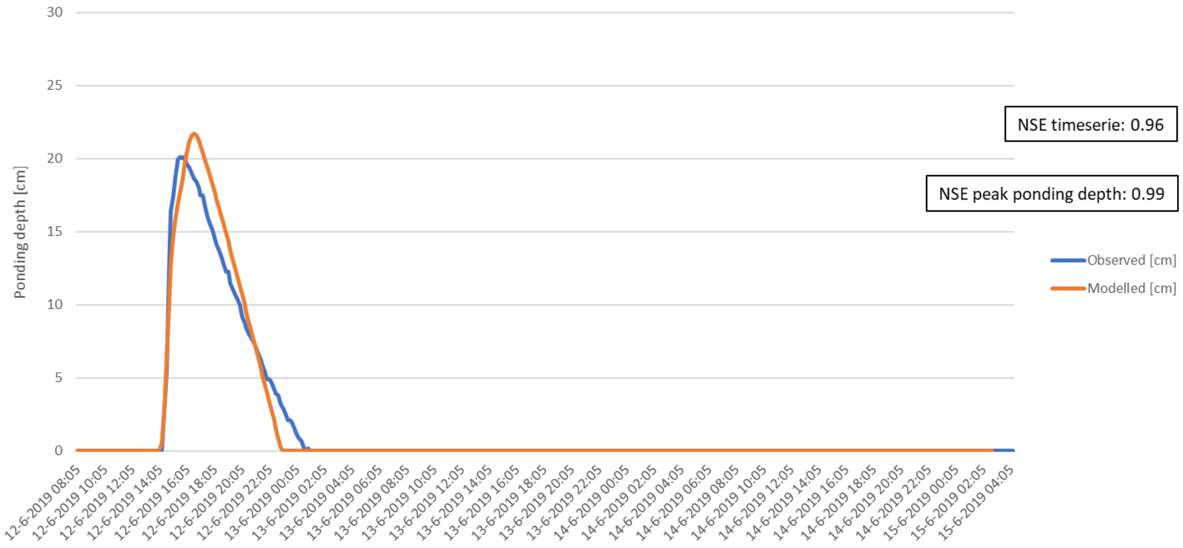


Figure 34 Observed and modelled ponding depth Karel Doormanlaan

During the calibration, the parameters of porosity, field capacity and wilting point were varied between the minimum and maximum values of Table 9. The values which resulted in the best accuracy are shown in the column “Value”.

Table 9 Calibrated input values Karel Doormanlaan

Parameter	Minimum	Value	Maximum
Porosity top [-]	0.25	**0.36**	0.49
Porosity bottom [-]	0.39	**0.47**	0.59
Saturated conductivity [mm/hr]	-	**26.6**	-
Unsaturated conductivity [mm/hr]	-	**1491**	-
Field capacity [-]	0.3	**0.35**	0.4
Wilting point [-]	0.1	**0.15**	0.15

With the values from calibration (Table 9), the model is validated on the rainfall event of the 19th of June. This validation resulted in a NSE of 0.75 on groundwater levels (Figure 35).

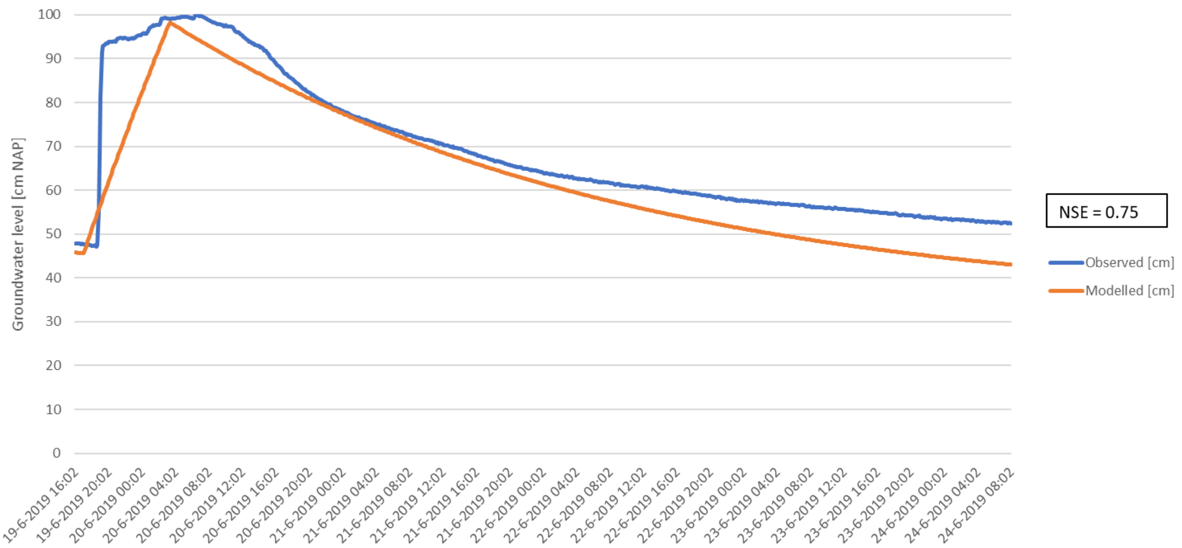


Figure 35 Observed and modelled groundwater levels Karel Doormanlaan

Validation of ponding depth resulted in a NSE value of 0.60 for the timeseries. The NSE value for the maximum ponding depth is 0.78 (Figure 36).

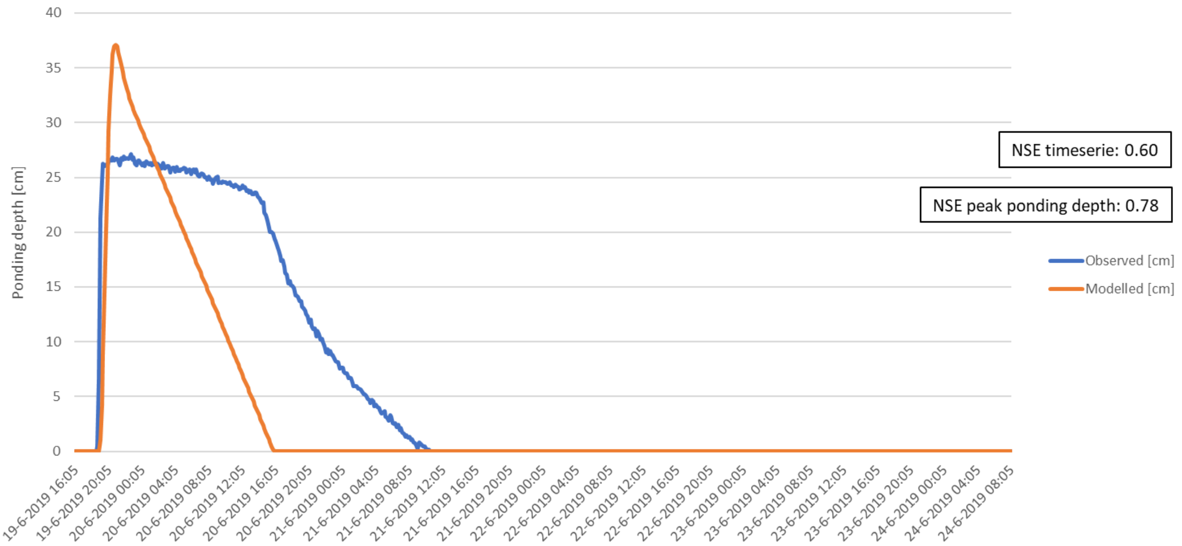


Figure 36 Observed and modelled ponding depth Karel Doormanlaan

The average NSE on groundwater levels is 0.73. With this NSE, the SWMM model is able to generate acceptable to good outputs on groundwater elevations. Based on the different rainfall events, the model is able to simulate acceptable maximum ponding depths. The simulation of complete timeseries results in less satisfactory ponding depths.

4.4 Sensitivity analysis

Figure 37 shows the change in mean groundwater levels in centimetres for different parameter values at the Klifrakplantsoen and Karel Doormanlaan. In these figures, it can be noticed that the change in mean groundwater levels is most significant when porosity is changed. The change in conductivity on the mean groundwater levels is almost negligible. It also can be noticed that the change in mean groundwater level is decreasing when changes in field capacity are higher.

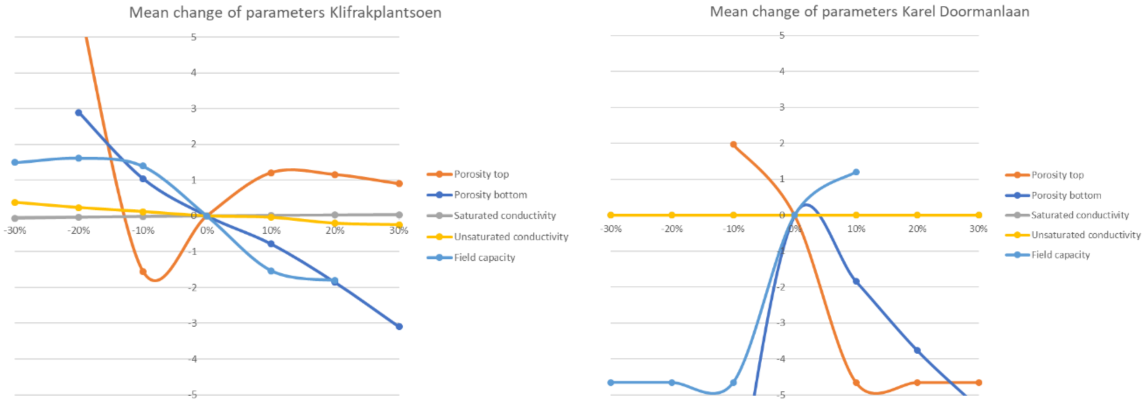


Figure 37 Change in mean groundwater level at the Klifrakplantsoen (left) and the Karel Doormanlaan (right)

Figure 38 shows the change in the maximum groundwater level in centimetres when different parameter values are used. Similar to the mean change in mean groundwater elevation, it can be noticed that the change in porosity values results in more significant changes in maximum groundwater levels. Timeseries of the sensitivity analysis for both locations are added in Annex X and Annex XI.

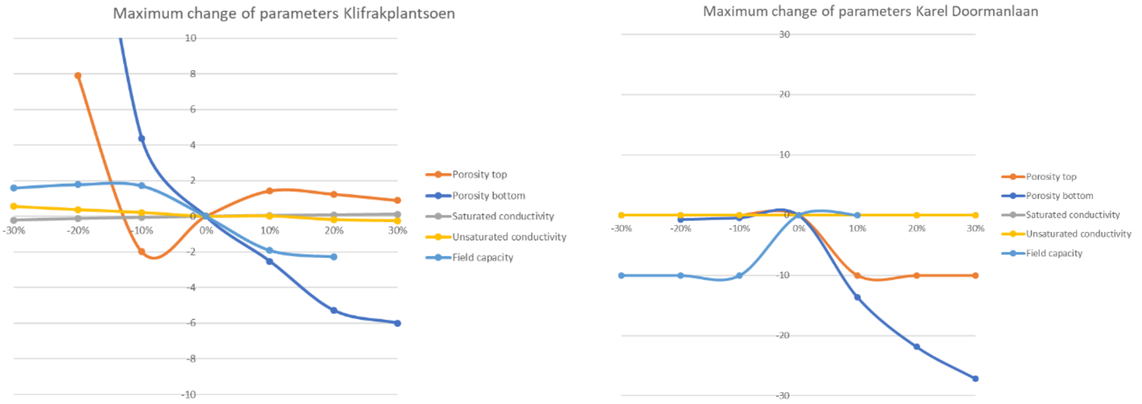


Figure 38 Change in maximum groundwater level at the Klifrakplantsoen (left) and the Karel Doormanlaan (right)

The results of the sensitivity analysis show that porosity is the most sensitive parameter in the Storm Water Management Model. The saturated and unsaturated conductivity has the least influence on the groundwater levels. When porosity values of the bottom soil layer are higher, the influence of different conductivity values disappears.

4.5 Climate durability

The municipality of Utrecht is planning to asses every project with changes in spatial planning on climate durability (Gemeente Utrecht, 2015). The climate durability of the wadi is assessed by a precipitation event of once in 100 years with a total amount of 109 millimetres in 48 hours.

Klifrakplantsoen

Figure 39 shows the increase in groundwater level for a precipitation event with a return time of once in 100 years for the Klifrakplantsoen. The groundwater level increases with around 25 centimetres and after 96 hours, the groundwater level is lowered with around 10 centimetres.

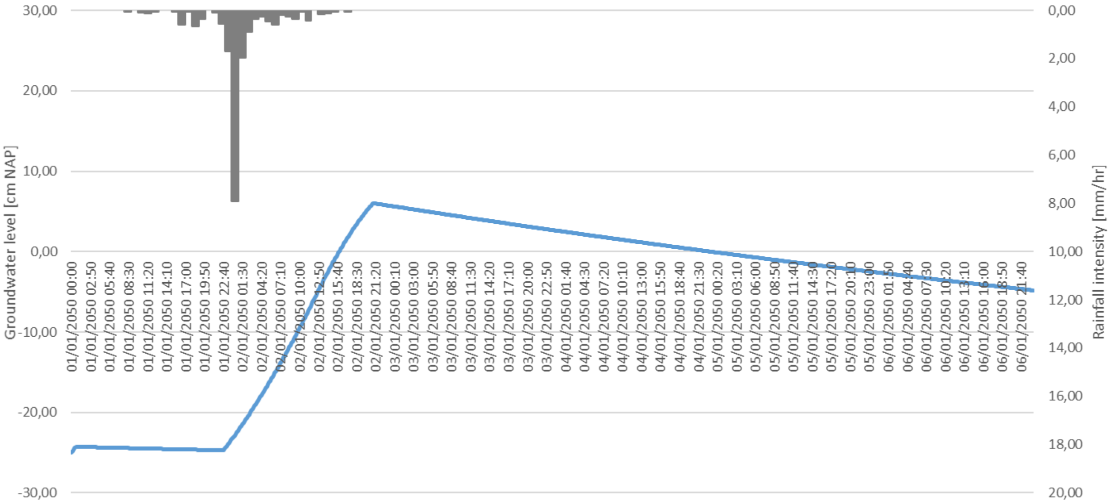


Figure 39 Modelled groundwater level and rainfall intensity for 1/100 year rainfall event at the Klifrakplantsoen

Figure 40 shows the ponding depth for the wadi at the Klifrakplantsoen for a rainfall event with a return time of once in 100 years. As can be noticed, no standing water occurs during the rainfall event. Therefore, this wadi is considered climate durable for future climate conditions.

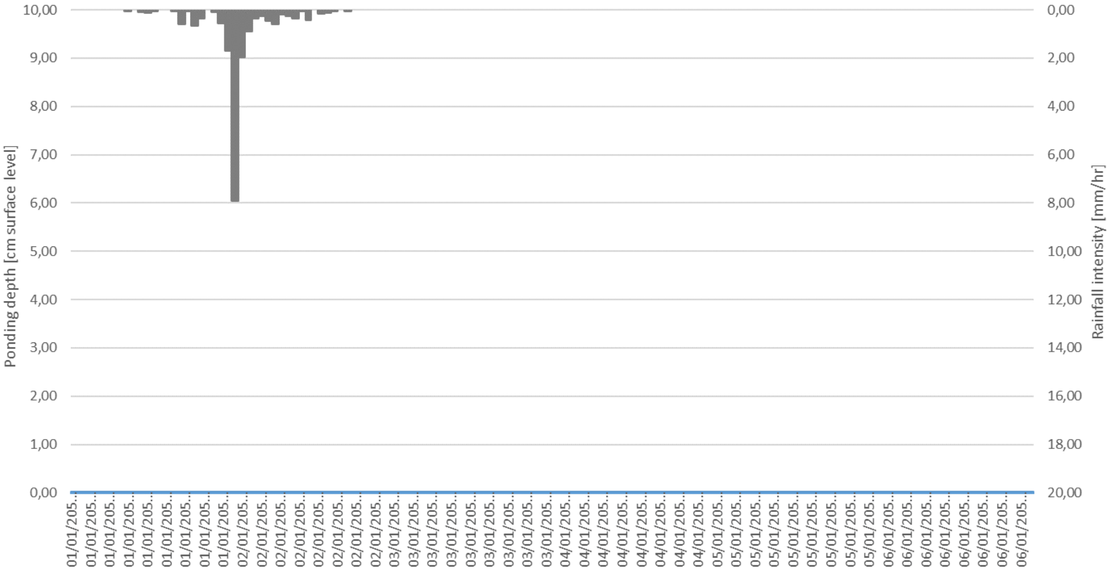


Figure 40 Modelled ponding depth and rainfall intensity for 1/100 year rainfall event at the Klifrakplantsoen

Karel Doormanlaan

Figure 41 shows the increase in groundwater level for a precipitation event with a return time of once in 100 years for the Karel Doormanlaan. The precipitation causes an increase in groundwater level of 65 centimetres where groundwater reaches the surface level. After around 96 hours, the groundwater level has returned to around 10 centimetres above the average groundwater level.

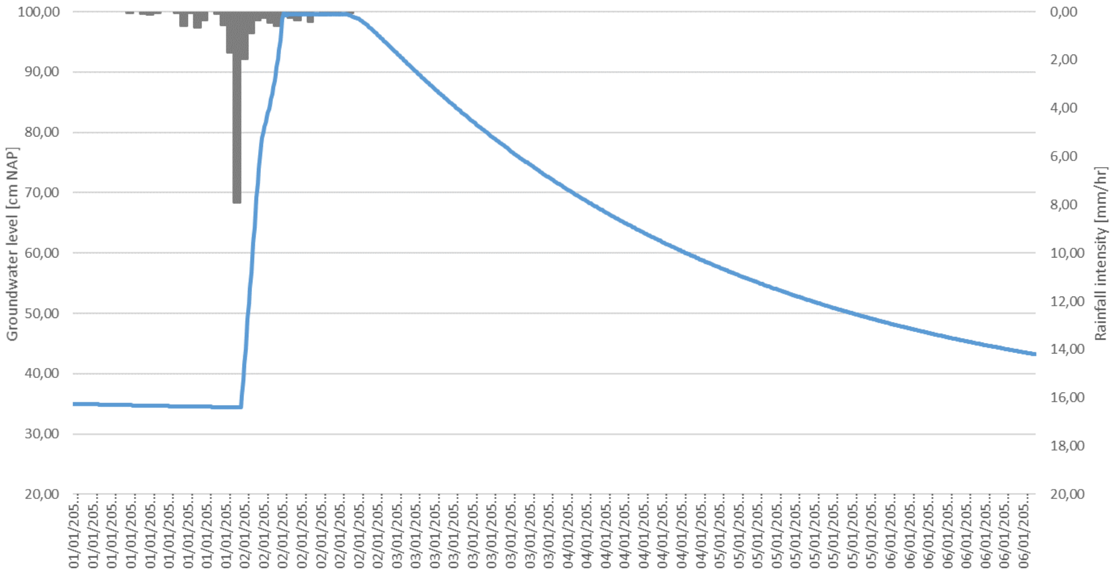


Figure 41 Modelled groundwater level and rainfall intensity for 1/100 year rainfall event at the Karel Doormanlaan

Figure 42 shows that during the precipitation, the surface is inundated with 6 centimetres of water. After 1.5 hours the standing water is completely infiltrated in the wadi. This is within the required 24 hours and therefore the infiltration facility is climate durable for precipitation events with a return time of once in 100 years.

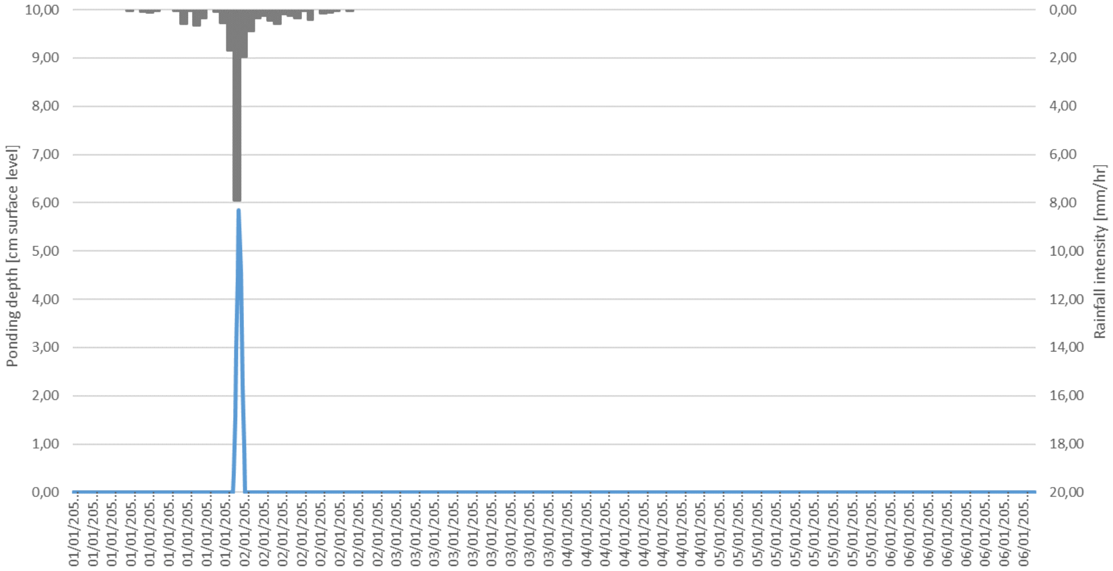


Figure 42 Modelled ponding depth and rainfall intensity for 1/100 year rainfall event at the Karel Doormanlaan

5. Discussion

5.1 Spatial variability field measurements

Spatial variability of soil parameters can be present at all scales because of variability in geologic controls (Gómez-Hernández & Gorelick, 1989). Determination of values for porosity and conductivity were measured at one place in the infiltration facility. However, there are many problems associated with spatial variability and measurement scale (Seyfried & Wilcox, 1995). Large scale measurements would produce stochastic average parameters for small-scale models. Even when the small-scale modelling is quite accurate, the impact of small errors in parameters on overall simulation accuracy is not known (Seyfried & Wilcox, 1995). For unsaturated hydraulic conductivity, the parameter value can vary up to 59% (Russo, Russo, & Laufer, 1997). The values for measured porosity is one of the least variable soil parameters (Warrick, 2001).

The measured porosity and conductivity can differ from effective average values, resulting in different model outcomes. Sensitivity analysis showed that the change in conductivity does not result in major changes in groundwater elevation, so the impact of spatial variability is limited. The effective values for porosity are less variable, but the impact on groundwater elevation is bigger. For future research, it is recommended to measure the values for porosity and conductivity on various locations within the study area. By measuring the parameters at various locations, the parameter uncertainty can be reduced.

5.2 Porosity values

The sensitivity analysis showed that porosity is the most sensitive parameter in the Storm Water Management Model. The input values for porosity in the model are determined by an estimation of the soil type and a corresponding range of minimum and maximum porosity values. Despite the fact that spatial variability of porosity is one of the least variable soil parameters, the estimate of the porosity can vary significantly from exact porosity at the study area. By varying the porosity values during model calibration it is expected that used porosity values are representative for the actual situation. However, to reduce parameter uncertainty in future research the exact value for porosity is recommended to be determined by taking samples with a Kopecky ring. With this method, a 100 cm^3 sample of the soil is taken and dried in an oven for 24 hours and 105 °C. After this, the sample is weighed and the porosity can be determined by comparing the volume of the sample with the dry mass of the sample (Dingman, 2015).

5.3 Multi-parameter calibration

The model in this research is calibrated with the trial and error method for mainly groundwater elevations. It is expected that the highest possible accuracy is not yet achieved because it is difficult to obtain an exact optimal solution because of the limited enumeration (Wu, Liu, Cai, Li, & Jiang, 2017). When calibrating the model on multiple parameters, such as groundwater elevation and ponding depth, model output accuracy can be increased. An effective method for multiple model calibration is proposed by Wu et al. (2017). In this method, a probability distribution for each parameter is specified. According to this probability distribution, several ranges for the parameters is selected. These parameter ranges are then compared with calibrated values to determine the optimal range. This multi-parameter optimal range selection method is superior to the single-parameter an application of this method is able to increase NSE values with a minimum of 0.01 (Wu et al., 2017). Combining the proposed multi-parameter optimal ranging method of Wu et al. (2017) and an automatic calibration software, such as SWMMR, can increase model efficiency and reduce calibration time.

5.4 Uncertainty climate durability

The climate durability of the wadies at the Klifrakplantsoen and Karel Doormanlaan is determined by integrating the expected climatological changes in a design precipitation event, resulting in a single scenario analysis. In the calibration and validation events, the modelled values for ponding depths had a maximum deviation of 36% compared to the observed values. When the modelled ponding depths for climate durability are increased with 36%, the wadies still meet the set requirements. However, the design precipitation is selected based on the highest intensity, or the highest percentile, during the precipitation. The effect of a precipitation event with a lower percentile, resulting in a lower but longer peak precipitation, is not researched. More extensive scenario analysis can ensure that assumptions about future developments are more transparent (Refsgaard, van der Sluijs, Højberg, & Vanrolleghem, 2007). For future research, it is recommended to assess climate durability with multiple scenario analysis. This can be done by simulating multiple percentile precipitation patterns for the expected climate scenario. This gives more transparency to what extent the infiltration facility is able to function within the set requirements or regulations.

5.5 Field data for wadi design

Currently, consultants design wadies by determining the infiltration rate by the conductivity of the soil. Sensitivity analysis in this research showed that conductivity has little effect on groundwater elevations compared to porosity values. Also, measurements of porosity have a smaller spatial variability then conductivity measurements (Russo, Russo, & Laufer, 1997). An improvement for designing wadies can be using measured porosity values instead of conductivity values. However, no sensitivity analysis have been done on ponding depths. Future research is recommended to determine whether wadies can be better designed by using porosity values over conductivity values.

6. Conclusion

This research was conducted with the aim of answering the following research questions:

“What hydrological infiltration model is fit-for-purpose to assess the infiltration capacity of infiltration facilities (wadies) within the design process?”

“What is the accuracy of this hydrological infiltration model in relation to groundwater elevation and ponding depths?”

and

“What is the future infiltration capacity of infiltration facilities?”

To answer the research questions, a literature review and two case studies were done. By reviewing existing scientific literature of different models and arrange these models with the benchmark criteria it can be concluded that the Storm Water Management Model is fit-for-purpose to assess current and future infiltration capacity of infiltration facilities.

The accuracy of the Storm Water Management Model is determined by comparing reference data of groundwater elevations and ponding depths with simulated values during different precipitation events. From these comparisons, it can be concluded that the accuracy of the model can acquire acceptable to good output values in relation to groundwater elevation. The accuracy of the model in relation to groundwater elevation for the Klifrakplantsoen is on average 0.77 and for the Karel Doormanlaan the average accuracy is 0.73. The accuracy of the model in relation to ponding depths is divided. The model is able to determine the maximum ponding depths in precipitation events with high accuracy up to 0.99. However, the accuracy of predictions on ponding depths over time is less satisfactory. The time in which a wadi is empty can therefore not be accurately determined.

By scenario modelling a precipitation event with climatological changes and simulating this event at the Klifrakplantsoen and Karel Doormanlaan, it can be concluded that the infiltration capacity of both study areas is enough to infiltrate a precipitation event of 109 millimetres in 48 hours without reaching the maximum ponding depths. Therefore, both the wadi at the Klifrakplantsoen and the Karel Doormanlaan are climate durable for future precipitation events.

It is recommended to do additional research to improve the model for more accurate estimations on the ponding depth over time. Due to the spatial variability, it is recommended to collect data on conductivity and porosity on multiple locations at the study areas. It is also recommended to improve the model accuracy by using autocalibration software in combination with multi-parameter calibration methods. To reduce parameter uncertainty for the porosity value, it is also recommended to take samples with a Kopecky ring and measuring dry mass of the sample.

References

Actueel Hoogtebestand Nederland. (2019, May 7). *Over AHN*. Opgehaald van Actueel Hoogtebestand Nederland: https://www.ahn.nl/over-ahn

Anderson, E. (2002). *Calibration of Conceptual Models for Use in River Forecasting.* Silver Spring, MD: Office of Hydrologic Development, US National Weather Service.

Berger, A. (1988, November). Milankovitch theory and climate. *Reviews of geophysics*, 624-657.

Bergström, S. (1992). *The HBV model - its structure and applications.* Norrköping: SMHI.

Beven, K. (1990). A discussion of distributed hydrological modelling. In K. Beven, *Distributed Hydrological Modelling* (pp. 255-278). Dordrecht: Springer.

Beven, K. (2018). On hypothesis testing in hydrology: Why falsification of models. *Wiley Interdisciplinary Reviews: Water, 5*(3).

Boogaard, F., & Wentink, R. (2007). *Dichtslibben van infiltratievoorzieningen.* Ede: RIONED.

Boyle, D., Gupta, H., & Sorooshian, S. (2000, December). Toward improved calibration of hydrologic models: Combining the strengths of manual and automatic methods. *Water Resources Research, 36*(12), 3663-3674.

Bruni, G., Reinoso, R., van de Giesen, N., Clemens, F., & ten Veldhuis, J. (2015). On the sensitivity of urban hydrodynamic modelling to rainfall spatial and temporal resolution. *Hydrology and Earth System Sciences*, 691-709.

Core Writing Team, Pachauri, R., & Meyers, L. (2014). *Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change.* IPCC. Geneva, Switzerland: IPCC.

Crowley, T. (2000). Causes of climate change over the past 1000 years. *Science*, 270-277.

Cunderlik, J., & Simonovic, S. (2004). *Calibration, Verification and Sensitivity Analysis of the HEC-HMS Hydrological Model.* The University of Ontario Department of Civil and Environmental Engineering.

Das, T., Bárdossy, A., Zehe, E., & He, Y. (2008). Comparison of conceptual model performance using different representations of spatial variability. *Journal of Hydrology*, 106-118.

Deltares. (2018). *Sobek - Hydrodynamics, Rainfall Runoff and Real Time Control - User Manual.* Delft: Deltares.

Devi, G., Ganasri, B., & Dwarakish, G. (2015). A Review on Hydrological Models. *Aquatic Procedia, 4*, 1001-1007.

Dingman, S. L. (2015). *Physical Hydrology.* Long Grove: Waveland Press.

Gemeente Utrecht. (2011). *Verbreed Gemeentelijk Rioleringsplan Utrecht 2011-2014.*

Gemeente Utrecht. (2013, March 14). Grondwatercontourenkaart Gemeente Utrecht.

Gemeente Utrecht. (2015). *Plan gemeentelijke watertaken 2016-2019.* Utrecht: Gemeente Utrecht Stadswerken.

Gómez-Hernández, J., & Gorelick, S. (1989). Effective Groundwater Model Parameter Values: Influence of Spatial Variability of Hydraulic Conductivity, Leakance, and Recharge. *Water Resources Research*, 405-419.

Grayson, R., Moore, I., & McMahon, T. (1992). Physically Based Hydrologic Modeling - 2. Is the Concept Realistic? *Water Resources Research*, 2659-2666.

Green, C., & van Griensven, A. (2007). Autocalibration in hydrologic modeling: Using SWAT2005 in small-scale watersheds. *Environmental Modelling & Software*, 422-434.

Hamby, D. (1994). A review of techniques for parameter sensitivity analysis of environmental models. *Environmental Monitoring and Assessment*, 135-154.

Harbaugh, A. (2005). *MODFLOW-2005, The U.S. Geological Survey Modular Ground-Water Model — the Ground-Water Flow Process.* U.S. Geological Survey Techniques and Methods 6-A16.

Hoes, O., & Schuurmans, W. (2006). Flood standards or risk analyzes for polder management in the Netherlands. *Irrigation and Drainage*, S113-S119.

Hoogheemraadschap de Stichtse Rijnlanden. (2015). *Handboek Watertoetsproces - Deel 2: Water in ruimtelijke plannen.* Hoogheemraadschap de Stichtse Rijnlanden.

Hunt, W., Kannan, N., Jeong, J., & Gassman, P. (2019). Stormwater best management practices: Review of current practices and potential incorporation in SWAT. *International Agricultural Engineering Journal*, 73-89.

KNMI. (2008, Januari 1). *neerslag - 1-uurneerslagaccumulaties van klimatologische met regenmeterdata gecorrigeerde radardataset voor Nederland (1 km)*. Opgeroepen op April 3, 2019, van KNMI DataCentrum: https://data.knmi.nl/datasets/rad_nl25_rac_mfbs_01h/2.0?q=neerslag

KNMI. (2015). *KNMI '14 klimaatscenario's voor Nederland.* De Bilt: KNMI.

KNMI. (2019, April 3). *Over WOW-NL*. Opgehaald van WOW-KNMI: https://wow.knmi.nl/over-wow-nl

Lindström, G., Johansson, B., Persson, M., Gardelin, M., & Bergström, S. (1998). Development and test of the distributed HBV-96 hydrological model. *Journal of Hydrology*, 272-288.

McCuen, R. (1973). The Role of Sensitivity Analysis in Hydrologic Modeling. *Journal of Hydrology*, 37-53.

McCutcheon, M., & Wride, D. (2013). Shades of Green: Using SWMM LID Controls to Simulate Green Infrastructure. *Journal of Water Management Modeling*, 289-301.

Morbidelli, R., Corradini, C., Saltalippi, C., Flammini, A., Dari, J., & Govindaraju, R. (2018). Rainfall infiltration modelling: A review. *Water*.

Moser, S. (2010). *Communicating climate change: history, challenges, process and future directions.* Santa Cruz: John Wiley & Sons.

Nash, J., & Sutcliffe, J. (1970). River flow forecasting through conceptual models part I - A discussion of principles. *Journal of Hydrology*, 282-290.

Neitsch, S., Arnold, J., Kiniry, J., & Williams, J. (2011). *Soil Water Assessment Tool - Theoretical Documentation.* Texas: Texas Water Resources Institute.

Niswonger, R., Prudic, D., & Regan, S. (2006). *Documentation of the Unsaturated-Zone Flow (UZF1) Package for Modeling Unsaturated Flow Between the Land Surface and the Water Table with MODFLOW-2005.* Reston: Techniques and Methods 6-A19.

Refsgaard, J., van der Sluijs, J., Højberg, A., & Vanrolleghem, P. (2007). Uncertainty in the environmental modelling process - A framework and guidance. *Environmental Modelling & Software*, 1543-1556.

Ritter, A., & Muñoz-Carpena, R. (2013). Performance evaluation of hydrological models: Statistical significance for reducing subjectivity in goodness-of-fit assessments. *Journal of Hydrology*, 33-45.

Ritzema, H. (2006). *Drainage Principles and Applications.* Wageningen: International Institute for Land Reclamation and Improvement.

Rossman, L., & Huber, W. (2016). *Storm Water Management Model Reference Manual - Volume I - Hydrology (Revised).* Cincinatti: U.S. Environmental Protection Agency.

Russo, D., Russo, I., & Laufer, A. (1997). On the spatial variability of parameters of the unsaturated hydraulic conductivity. *Water Resources Research*, 947-956.

Saloranta, T., Kamari, J., Rekolainen, S., & Malve, O. (2003). Benchmark criteria: A tool for selecting the appropiate models in the field of water management. *Environmental Management, 32*(3), 322-333.

Seyfried, M., & Wilcox, B. (1995). Scale and the nature of spatial variability: Field examples having implications for hydrologic modeling. *Water Resources Research*, 173-184.

Smits, I., Wijngaarden, J., Versteeg, R., & Kok, M. (2004). *Statistiek van extreme neerslag.* Utrecht: STOWA.

Stichting Toegepast Onderzoek Water. (2012). *Handboek meten van grondwaterstanden in peilbuizen.* Amersfoort: STOWA.

STOWA. (2003). *Vooronderzoek natuurvriendelijke wadi's.* Utrecht: Stichting Toegepast Onderzoek Waterbeheer (STOWA).

STOWA. (2004a). *Nieuwe neerslagstatistiek voor waterbeheerders.* Utrecht: STOWA.

STOWA. (2004b). *Statistiek van extreme neerslag in Nederland - Uitwerking neerslagpatronen.* Utrecht: STOWA.

STOWA. (2015). *Actualisatie meteogegevens voor waterbeheer 2015.* Amersfoort: STOWA.

StructX. (2019, July 16). *Typical Porosity Values for Various Soil Types*. Opgehaald van StructX: https://www.structx.com/Soil_Properties_006.html

Swedish Meteorological and Hydrological Institute. (2019, September 9). *HBV*. Opgehaald van Swedish Meteorological and Hydrological Institute: https://www.smhi.se/en/research/research-departments/hydrology/hbv-1.90007

Tchakoua, P., Wamkeue, R., Slaoui-Hasnaoui, F., Tameghe, T., & Ekemb, G. (2013). New Trends and Future Challenges for Wind Turbines Condition Monitoring. *International Conference on Control, Automation & Information Sciences*, 238-245.

U.S. Army Corps of Engineers. (2000). *Hydrologic Modeling System HEC-HMS - Technical Reference Manual.* U.S. Army Corps of Engineers.

U.S. Army Corps of Engineers. (2008). *Hydrologic Modeling System HEC-HMS - Applications Guide.* U.S. Army Corps of Engineers.

Uhlenbrook, S., Seibert, J., Leibundgut, C., & Rodhe, A. (1999). Prediction uncertainty of conceptual rainfall-runoff models caused by problems to identify model parameters and structure. *Hydrological Sciences*, 779-798.

United Nations. (2018, May 16). *68% of the world population projected to live in urban areas by 2050, says UN*. Opgehaald van United Nations - Department of Economic and Social affairs: https://www.un.org/development/desa/en/news/population/2018-revision-of-world-urbanization-prospects.html

United States Department of Agriculture. (2017). *Soil Survey Manual.* Washington: USDA.

Vaes, G., Bouteligier, R., & Berlamont, J. (2004). Het ontwerp van infiltratievoorzieningen. *Waterniveau*, 1-13.

van Beers, W. (1983). *The Auger hole method.* Wageningen: International Institute for Land Reclamation and Improvement.

van den Hurk, B., Siegmund, P., Klein Tank, A., Attema, J., Bakker, A., Beersma, J., . . . van Zadelhoff, G.-J. (2014). *KNMI '14: Climate Change scenarios for the 21st Century – A Netherlands perspective.* De Bilt: KNMI.

van Griensven, A., Meixner, T., Grunwald, S., Bishop, T., Diluzio, M., & Srinivasan, R. (2005). A global sensitivity analysis tool for the parameters of multi-variable catchment models. *Journal of Hydrology*, 10-23.

Vasseur, B., Devillers, R., & Jeansoulin, R. (2003). Ontological approach of the fitness of use of geospatial datasets. *Proceedings of the 6th AGILE*, (pp. 1-8). Lyon.

Vergroesen, T., Verschelling, E., & Becker, B. (2014). Modelling of sustainable urban drainage measures. *Ingeniería Innova*, 1-16.

Vlaams Instituut voor de Zee. (2019a, Juli 29). *Model Calibration*. Opgehaald van VLIZ: http://www.vliz.be/wiki/Model_calibration

Vlaams Instituut voor de Zee. (2019b, August 12). *Model Validation*. Opgehaald van VLIZ: http://www.vliz.be/wiki/Model_validation

Wagener, T., Boyle, D., Lees, M., Wheater, H., Gupta, H., & Sorooshian, S. (2001). A framework for development and application of hydrological models. *Hydrology and Earth System Sciences*, 13-26.

Warrick, A. (2001). Classical measures of variability. In A. Warrick, *Soil Physics Companion* (pp. 344-347). CRC Press.

Wijngaard, J., Kok, M., Smits, A., & Talsma, M. (2005). Nieuwe statisitiek voor extreme neerslag. *H2O*.

Wu, Q., Liu, S., Cai, Y., Li, X., & Jiang, Y. (2017). Improvement of hydrological model calibration by selecting multiple parameter ranges. *Hydrology and Earth System Sciences*, 393-407.

Zhang, H., Wang, Y., Wang, Y., Li, D., & Wang, X. (2013). The effect of watershed scale on HEC-HMS calibrated parameters: a case study in the Clear Creek watershed in Iowa, US. *Hydrology and Earth System Sciences*, 2735-2745.

Annex I: Model requirements

Requirements related to output of the model:

- Model is able to generate ponding depths as an output.
- From the output of the model, it is possible to determine the time to empty the infiltration facility and infiltration capacity.

Requirements related to the model's span and resolution in time and space:

- The model is able to generate output in timesteps of ten minutes.
- Model is able to generate output for infiltration facilities which are 10 m2 or greater.

Requirements related to the model's input parameters:

- Model's relevant input parameters are limited to:
    - Surface area of infiltration facility.
    - Maximum depth of infiltration facility/depth to where ponding can occur.
    - Surface area of catchment area.
    - Hydraulic conductivity.
    - Porosity
- Other parameters are able to determine by using values from literature.

Requirements related to model's user-friendliness:

- non-experienced users are able to use the model within a few days.
- Parameters can be adjusted with a few acts inside the model.
- Output data can be visualized by graphs.

Requirements related to models flexibility for adaptation:

- Input data, such as rainfall data, is easy to import and adapt.
- The time to set-up the model for various project locations is limited and can be done within a few hours.

Annex II: Benchmark criteria – boundary conditions answers

Model applicability and relevance for the management task		
How well does the model's output relate to the management task?	Good (at least 2 items)	- The model's output can be directly related to the "core" of the management task. - The model's output (relevant to the management task) consists of variables that are commonly applied and easy to measure or quantify (e.g., the quality elements and pollutants listed in the Annex V of the WFD). - The model allows the simulation of a variety of relevant management operations.
	Adequate	- The model's output can be related to the management task via clear, well-known, and well-established links.
	Inadequate (at least 1 item)	- The model's output is peripheral in relation to the management task. - The links between the model's output and management task are not clear or adequately scientifically established.

How well does the model's span and resolution in time and space compare with the requirements of the management task?	Good (all items)	- The model can be run with any desired spatial and temporal resolution. - The model can be run over the desired spatial and temporal span (e.g. it allows simulations to be run over many years).
	Adequate	- There are restrictions on the model's spatial or temporal resolution or span, but the model is still expected to produce useful and meaningful results for the management task.
	Inadequate	- The model's spatial or temporal resolution or span cannot be chosen to be appropriate for the management task.

How well has the model been tested?	Good (at least 3 items)	- There are at least 10 documented previous model applications. - At least five model applications are published in peer-reviewed journals. - The model has been evaluated against independent data sets. - The model has been evaluated in various conditions or geographical regions. - Some previous model use and evaluation is closely related to the management task in question.
	Adeqate (at least 2 items)	- There are at least three reported model applications. - The model has been evaluated in different conditions or geographical regions. - The model is specific to the site of the management task.
	Inadequate (at least 1 item)	- The model is site-specific to other type of site than that of the management task and it has not been evaluated in different conditions or geographical regions. - There are less than three documented previous model applications.

How complicated is the model in relation to the management task?	Good	- The model has an optimally simple structure, i.e., it includes mostly only those processes and parameters that are known to be relevant for the management task.
	Adequate (one of the items)	- The model has a somewhat too complicated structure, i.e., most of the model's processes and parameters are relevant but the model seemingly includes also some irrelevant processes and parameters. - Alternatively, the model is somewhat too simple, i.e., its relevance to the management task could be enhanced somewhat (but not radically) by introducing some additional processes.
	Inadequate (one of the items)	- The model is too complex, and most of the model's features could clearly be omitted or simplified (or a more simple model could be chosen) without loss in model relevance for the management task. - Alternatively, model is too simple, and many key processes relevant to the management task are not included.

How is the balance between the model's input data requirements and data availability?	Good	- The required model input data are available from monitoring and field observations, either from the management site or from other applicable site close to it.
	Adequate	- Most of the required model input data are available from monitoring and field observations, either from the management site or from other applicable site close to it. However, some surrogate input data (e.g., results from other models or data from other remote sites) must be used.
	Inadequate	- A majority of the required model input data are not available from monitoring and field observations from the management site (or from other applicable site close to it).

How is the identifiability of the model parameters?	Good (at least 1 item)	- All relevant model parameter values are well documented in scientific literature or can be estimated directly based on available data. - Available data (corresponding to model output variables) will allow the establishment of all relevant model parameter values via model calibration.
	Adequate	- There seems to be enough data or documentation available to allow an adequate estimate of most of the relevant model parameter values (either directly or via model calibration).
	Inadequate	- There are clearly not enough calibration data or other parameter documentation available to allow for an adequate establishment of many of the relevant model parameter values.

How easily are the model results understood and interpreted?	Good	- Nonspecialist users are generally capable of understanding and interpreting the model output results.
	Adequate	- Assistance from research staff or modeling specialist is necessary to clarify and interpret the model's output results.
	Inadequate (at least 1 item)	- Expert skills, long experience, and deep insight (e.g., those of a model developer) are needed to understand and interpret the model results. - Much "tacit" (i.e., difficult-to-express) knowledge or intuition is involved in the interpretation of the model results.

How is the peer acceptance for the model and the model's consistency with scientific theory?	Good (at least 2 items)	- The model has gained wide and international acceptance among the scientific community. - The model is widely used in many countries. - The whole model is based on well-established scientific theory.
	Adequate (at least 1 item)	- Model is used and has gained peer-acceptance mostly locally/nationally. - most of the model components are based on well-established science.
	Inadequate (at least 1 item)	- The model is based on speculative or immature scientific theory and/or assumptions. - The model is used only by few persons.

Model uncertainty and sensitivity		
How well is the model suited for sensitivity and uncertainty analyses and how well have these analyses been performed and documented?	Good (at leat 3 items)	- Thorough analysis of model sensitivity has been performed and reported. - Model sensitivity analysis is published in peer-reviewed journal(s). - A variety of sensitivity/uncertainty analysis techniques or software can easily and with reasonable effort be applied to the model. - The model software contains tools for sensitivity/ uncertainty analysis. - Uncertainty ranges, associated with the model parameter values, can be adequately established.
	Adequate (at least 1 item)	- Screening of the most sensitive model parameters has been done and published in technical report(s). - Sensitivity/uncertainty analysis techniques or software can be applied to the model, but this will be a rather laborious task.
	Inadequate	- No model sensitivity analysis has been performed because the model is generally not suitable for adequate analysis of sensitivity/uncertainty.

Model transparancy, Ease of understanding, Ease of use		
How is the model's version control?	Good (all items)	- Different model versions are numbered and description of version development exists. - It is easy to check the version of the executable model. - User manual and other model documentation matches with the particular model version.
	Adequate (one of the items)	- Model versions are numbered. - User manual and other model documentation is known to be sufficiently consistent with the particular model version. - Alternatively, only one version exists.
	Inadequate	- No consistent numbering between different model versions exists.

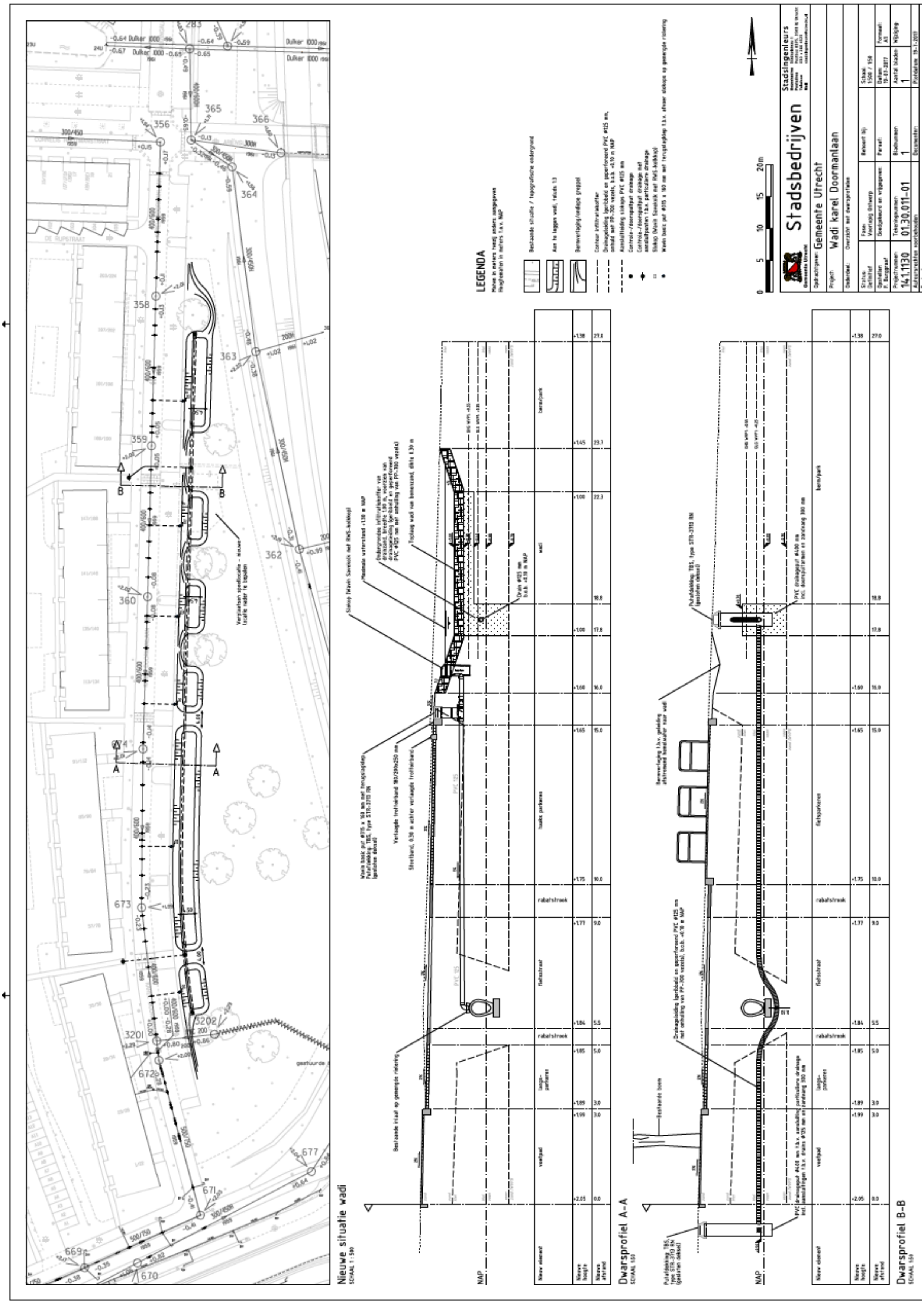
How are the model's user manual and tutorial?	Good (at least 2 items)	- Instructions for use are comprehensive and detailed, yet operative and clear. - The scope of the model, its application domain, input file structures, and parameter estimation methods are explained. - There are application examples, or a well-structured tutorial section.
	Adequate	- User manual is less comprehensive, but includes clear operating instructions.
	Inadequate	- Adequate user manual is not available.

How is the model's technical documentation?	Good (at least 2 items)	- Model documentation gives comprehensive and detailed description of the processes, algorithms, and numerical methods. - The science behind the model is reviewed in the documentation. - Documentation is published in peer-reviewed scientific journal(s).
	Adequate	- Technical document of model processes and equations is available.
	Inadequate	- No adequate technical document of the model and its structure is available.

How are the model's interactiveness, user-friendliness and suitability for end-user participation?	Good (at least 3 items)	- The model is well structured, transparent, and has informative user interface with easy visualisation of the model output. - Input data format is user-friendly and model parameters are easily modified (or the model is connected to parameter databases). - Active user support is available, either from model developers or from a user-group. - Nonspecialist users are generally capable of running the model. - The model can contribute to the process of negotiation among relevant stakeholders.
	Adequate (at least 1 item)	- The model is less transparent and the facilitation of a model specialist is required to guide the model use. - The model has a well-functioning user interface offering the user some insight and control on model parameters and functioning.
	Inadequate (at least 1 item)	- The model is an "opaque box," and allows the user no interaction with the model and its parameters. - Only a specialist (e.g. a model developer) can use the model.

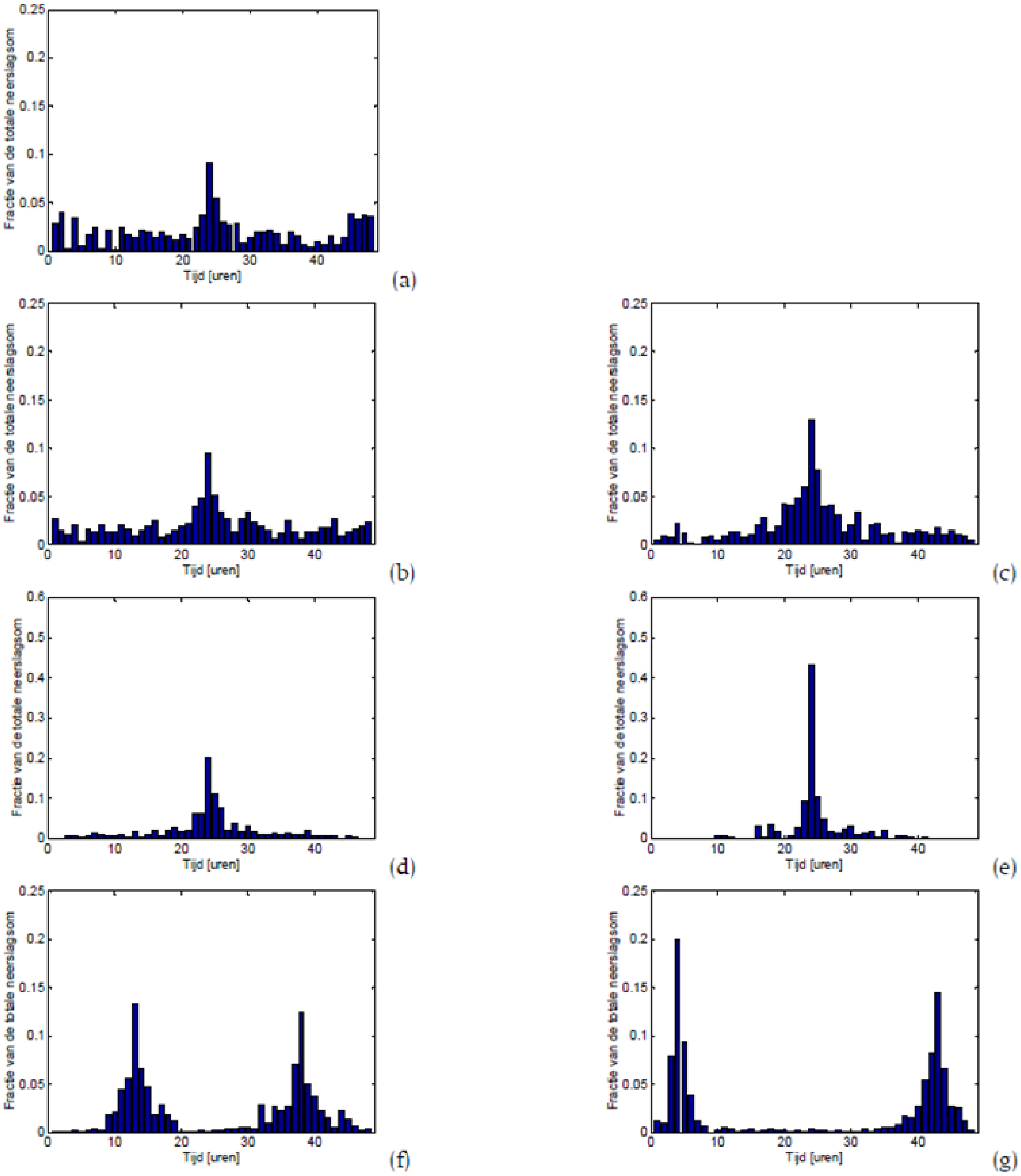
How is the model's flexibility for adaptation and improvements?	Good (at least 2 items)	- The model's source code is available to the model user and is well structured and documented. - The model is flexible, i.e., different processes can easily be added to (or removed from) the model in the form of, e.g., add-in modules. - The model is easily adaptable for inclusion in integrated model systems.
	Adequate (one of the items)	- The model's source code is available to the model user. - Alternatively, the model's source code is not generally available, but model developers may give support for adaptation and improvements.
	Inadequate	- The model's source code is not available and no active model development exists.

Annex III: Karel Doormanlaan wadi design



Annex IV: Percentiles of characteristic precipitation patterns

KARAKTERISTIEKE NEERSLAGPATRONEN VOOR NEERSLAGGEBEURTENISSEN MET EEN DUUR VAN 48 UUR. EEN UNIFORM PATROON ZONDER PIEK (A), 1-PIEKS PATRONEN GEBASEERD OP HET 12,5% PERCENTIEL (B), HET 37,5% PERCENTIEL (C), HET 62,5% PERCENTIEL (D) EN HET 87,5% PERCENTIEL (E) EN 2-PIEKS PATRONEN MET EEN KORTE TUSSENDUUR (F) EN EEN LANGE TUSSENDUUR (G)



Annex V: Benchmark criteria - complete answers

Model applicability and relevance for the management task		HBV-model	HEC-HMS model	ModFlow-2005	Sobek Urban	Storm Water Management Model	SWAT	Geonius Excel model
How well does the model's output relate to the management task?	Good			The model's output can be directly related to the management task and allows variety of relevant management operations.	The infiltration capacity and ponding depth of a swale is possible to model according to Vergroesen, Verschelling and Becker in 2014.	With the help of LID (Low Impact Development) tools, the model can be related to the management task. (Rossman, 2010) Output consists of variables which are easy to measure/quantify. (McCutcheon, 2013)		
	Adequate		Main output of the model is discharge, but values such as infiltration rate are possible to relate. (US Army Corps, 2008)				It is possible to use the model for the management taks, but it needs multiple changes to obtain the relevant model output. Most model output relate to water quality instead of water quantity. (Hunt, 2009)	
	Inadequate	Model output is mainly the amount of discharge in river basins. (Bergström, 1992) It is difficult to relate the discharge to the infiltration capacity and ponding depth in the wadi.						Model output is limited to time in which wadi is empty. Ponding depths and groundwater elevation levels are not presented. Also the formula's are not scientifically established.
How well does the model's span and resolution in time and space compare with the requirements of the management task?	Good			The model can be run with any spatial and temporal resolution and span. Small spatial grids are ideal for the management task.	The spatial resolution used by Vergroesen, Verschelling and Becker in 2014 was 1800m2 which is relevant for the management task. The temporal resolution can be minutes or hours.	The models spatial and temporal resolution meets the requirements of the management task. The model is capable to run over the desired spatial and temporal span.		
	Adequate							The model's spatial resolution is in line with the management task. The temporal resolution is limited to only two different rainfall events.
	Inadequate	Good fits were obtained for the larger, mesoscaled basins. The results of the smallest, microscaled basin were less satisfactory. (Stefan Uhlenbrook et al.) The HBV model is originally a conceptually lumped model. Though a great deal of effort has been invested in accounting for spatial variability of precipitation and temperature in the "distributed" HBV model, it still falls short in its ability to account for full spatial variability due to its conceptual and lumped nature (for example no representation of lateral flow component in the designed "distributed" HBV model). (Das, 2008) The model runs in timesteps of 1 day or 12 hours, which is too big for the management task of this research.	However, configurations with 1 (30km2) and 157 (0,86km2) subbasins do not yield as good of results as the others, and we may expect a threshold level of sub-basin delineations beyond which model parameters have little possibility to enhance model performance. (Zhang, 2013) This is much bigger than the spatial resolution for the management task. Temporal scale can be chosen for this model (minutes or hours)				Temporal resolution of the model is daily and spatial resolution is too big for management task. Research of Green (2008) shows that small watersheds of 4 to 8 ha give not representative results. For appropriate management task, model needs to function below 10 ha. The model is not designed for simulating detailed, single-event flood routing.	
How well has the model been tested?	Good	We know that the model has been applied in more than these 30 countries (Bergstrom, 1992). However no model application is known for the management task in this study.	The model has been applied various times in different watersheds/regions with independent datasets. The results of these applications can be found in peer-reviewed journals. However no model application is known for the management task in this study.	The model contains more than 10 previous model applications. Most of them are published in peer-reviewed journals and some of those model applications is closely related to the management task.		There are over 10 model applications, and over five are published in peer-reviewed journals. The model has been evaluated in various conditions and geographic regions and some of them include model use which is closely related to the management task.		
	Adequate				Model is applied in multiple situations in different conditions/regions. Scientific literature of these applications is limited.		The model has multiple reported applications which are published in peer-reviewed journals. These applications have been evaluated in different conditions or geographical regions. However, few model applications are closely related to the management task.	
	Inadequate							The model is used various times to consult on the design of infiltration facilities. However, the model has never been tested to what extent model outcome is accurate.
How complicated is the model in relation to the management task?	Good				The model seems to have an optimally structure in terms of processes and parameters relevant for the management task.	The model has almost an optimal structure. Most relevant processes and parameters for the management task are included.		
	Adequate		The model includes most of the processes and parameters which are known to be relevant. However, the model is intended to calculate the discharge of an basin. (US Army Corps, 2000)	The model is somewhat too complicated. Almost all parameters or processes are included, also the more irrelevant processes and parameters.			The model is somehwat too complicated for the management task. There are too many irrelevant processes in the model.	
	Inadequate	The model is relatively simple. Basic equations are used, but give representative outcomes in terms of discharge. However, the output of the model is missing relevance for the management task. (Lindström, 1997)						Model is too simple, infiltration is only determined by conductivity and needed infiltration surface. Groundwater level is not included in the model.
How is the balance between the model's input data requirements and data availability?	Good	The input data can be obtained from monitoring and/or field observations.		Required model input data is available from monitoring and field observations. Less important parameters can be estimated based on scientific literature.	Required input data is available from monitoring and field observations.	Required model input data is available from monitoring and field observations.		Required input data is available from field observations.
	Adequate		Most of the model data is available by monitoring and field observations. Other input data have to obtained by calculations and derivations (but not from another model result).				Required data is available from monitoring and field observations. However, due to input format of the model data, most of the data has to be transformed.	
	Inadequate							
How is the identifiability of the model parameters?	Good	Relevant model parameter values are documented in different scienitic literature. Also parameter uncertainty is well documented.	The technical reference manual describes all model parameters, what values these parameters needs to have or how these can easblished by calibration.	Relevant model parameters are documented and a parameter estimation tool is incorporated in the model. (Harbough, 2005)	Relevant model parameter values can be estimated directly based on available data.	Relevant model parameter values are documented in the Reference Manual. (EPA, 2016)	All model variables/parameters are documented in theoretical documentation of the model. (SWAT, 2009)	
	Adequate							
	Inadequate							Because the model has never been tested, there is not known what relevant model parameter values are.
How easily are the model results understood and interpreted?	Good	Nonspecialists in the field of water management are capable of understanding and interpreting the model output results.	Nonspecialists in the field of water management are capable of understanding and interpreting the model output results.	Nonspecialists in the field of water management are capable of understanding and interpreting the model output results.	Nonspecialists in the field of water management are capable of understanding and interpreting the model output results.	Nonspecialists in the field of water management are capable of understanding and interpreting the model output results.	Nonspecialists in the field of water management are capable of understanding and interpreting the model output results.	Model output results are generally easy to understand.
	Adequate							
	Inadequate							

How is the peer acceptance for the model and the model's consistency with scientific theory?	Good	The model is widely used and based on well-established scientific theory.	The model has been used in many countries and different circumstances. The outcomes of these model calculations are published in several scientific reports and international acceptance among the scientific community is present.	The model is used in many countries and is established based on scientific theory. The iMOD GUI is mostly used within Dutch scientists.			The model has been used in many countries and has gained international acceptance among the scientific community.	
	Adequate				Model is used and has gained peer-acceptance mostly in the Netherlands.	Model is mostly used in the United States and Korea. Model components are based on well-established science.		
	Inadequate							The model is based on assumptions and simplified equations. Also, the model is only used within Geonius.
Model uncertainty and sensitivity								
How well is the model suited for sensitivity and uncertainty analyses and how well have these analyses been performed and documented?	Good	Model sensitivity analysis are published in peer-reviewed journal(s). Uncertainty ranges with model parameters are published as well.	The models sensitivity has been performed and reported. This analysis is published in a report (Cunderlik & Simonovic, 2004).			Various sensitivity analysis of the model have been performed for different model applications. These model sensitivity analysis are published in peer reviewed journals.	Model sensitivity has been performed and published in peer-reviewed journal (van Griensven, 2006). Uncertainty ranges with the model parameters can be adequately established.	
	Adequate			There are sensitivity analysis techniques for the software which can be applied but due to the wide variety of model application it takes some time to execute.				
	Inadequate				A model sensitivity analyses has been performed on the temporal and spatial scale, but not for the model parameters itself. (Bruni, 2015)			No model sensitivity analysis has been performed.
Model transparancy, Ease of understanding, Ease of use								
How is the model's version control?	Good		Different versions of the model are numbered and a description of changes in the version are given. Documentation matches with the model version.	Model versions are numbered and a description of the development exists. It is easy to check which version you execute, all previous versions are still downloadable. For each core version there is a particular user manual and other documentation.	Model versions are numbered. Only the most recent version is downloadable. With every new version, release notes describe the changes compared to previous model version. User manual and other documentation corresponds with current model version.	Model versions are numbered and description of version development exists. User manual corresponds with most recent model version. The most recent model version is the only executable model.	The model versions are numbered and has his own model documentation. It is easy to check which version is the most recent and executable one. (SWAT website)	
	Adequate	Model versions are numbered.						Model versions are numbered.
	Inadequate							
How are the model's user manual and tutorial?	Good		User manual is detailed, operative and clear. The scope of the model, application domain and input file structures are explained. (US Army Corps, 2008). However, the scope of the model does not correspond with the management task.	For the iMOD graphic user interface which is tested, there is a comprehensive and detailed user manual. The scope and application domain is clear. Every GUI uses other input file structures, but these are mentioned clearly. For iMOD there are several tutorials available.	There is a detailed and comprehensive user manual in which the scope of the model, input files and parameters estimation methods are explained.	The user manual is comprehensive and detailed. Scope of the model, input file structures and parameter estimation methods are described. Other documentation gives examples about the application of the model.		
	Adequate						User manual is not comprehensive, it only describes the different input and output files. Video tutorials are available but onlys show one model application.	
	Inadequate	Adequate user manual is not available.						User manual is not available.
How is the model's technical documentation?	Good	Model documentation is detailed and published in peer-reviewed scientific journals.	Model documentation is detailed and is reviewed in the documentation (US Army Corps, 2000).	Model documentation gives a comprehensive and detailed description. Science behind the model is reviewed in the documentation. (Harbough, 2005)		Model documentation about processes, algorithms and numerical methods is detailed. In the documentation is referred to scientific literature to explain the science behind processes or formulas.		
	Adequate				Model processes and equations are described in an technical reference document. No review of science behind the model and peer review available.		Technical documentation is available, but does not correspond with the most recent model version. (SWAT, 2009)	
	Inadequate							No technical document of the model and its structure is available
How are the model's interactiveness, user-friendliness and suitability for end-user participation?	Good				The model is well structured with a user friendly interface. Model input and parameters are easily to change. User support is available from Deltares.	The model is well structured and has informative user interface with clear visualisation of the model output. Input data is described in the user manual and model parameters can be modified easily. Support is available from developers by a contact form on the website.	The model is run in ArcGis, which has a well structured and informative user interface. Support is available from both user groups as developers. Input data is described and parameters are easily to change.	The model is structured, formulas are easy to trace and the output is easily visualized. Input data format is user friendly and easy to modify. Nonspecialist are generally capable to run the model.
	Adequate		The model has a well-functioning user interface in which some insight and control on model parameters can be found. To get along with the model, more experience is needed or a model specialist is required to guide the model use.	The model is less transparent and model specialists are needed to guide model use. For end-users it is somewhat too complicated compared to the management task.				
	Inadequate	Based on the information available it is not able to understand the model without relevant knowledge.						
How is the model's flexibility for adaptation and improvements?	Good			Model's source code is available. Own executables can be implemented in the iMOD GUI. How to do this is described on the iMOD forum pages. Support on how to do this is available as well.				The model is flexible, different formulas and processes can be added or changed. Excel is easily adaptable for inclusion in integrated model systems.
	Adequate	Model source code is available, but specialists/developers are needed for adaptation and implementation			Model's source code is not available but support by developers may include adaptation and improvements.	Model's source code is available. Source code is divided in computational engine and graphical user interface. (not sure to what extent model source code is flexible, if yes than good)	Model's source code is available. No add-in modules or support is available for the model to create add-in modules.	
	Inadequate		Model's source code is not available and no add-ons exist. The model itself is being developed by the US Army Corps of Engineers itself.					
		7x Good 2x Adequate 5x Inadequate	8x Good 4x Adequate 2x Inadequate	11x Good 3x Adequate 0x Inadequate	9x Good 4x Adequate 1x Inadequate	12x Good 2x Adequate 0x Inadequate	6x Good 7x Adequate 1x Inadequate	4x Good 2x Adequate 8x Inadequate
		16 pt	20 pt	25 pt	22 pt	26 pt	19 pt	10 pt

Annex VI: Measurement of saturated conductivity Klifrakplantsoen

Projectomschrijving:	Verzadigde doorlatendheidsmeting	Opdrachtnr.	Thesis Universiteit Utrecht
Locatie:	Klifrakplantsoen	Traject (m-mv)	-
Boornummer:	DB01	Meting	DM01

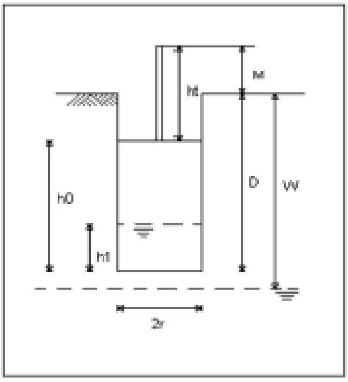
Formule om de doorlatendheid volgens Porchet te bepalen :

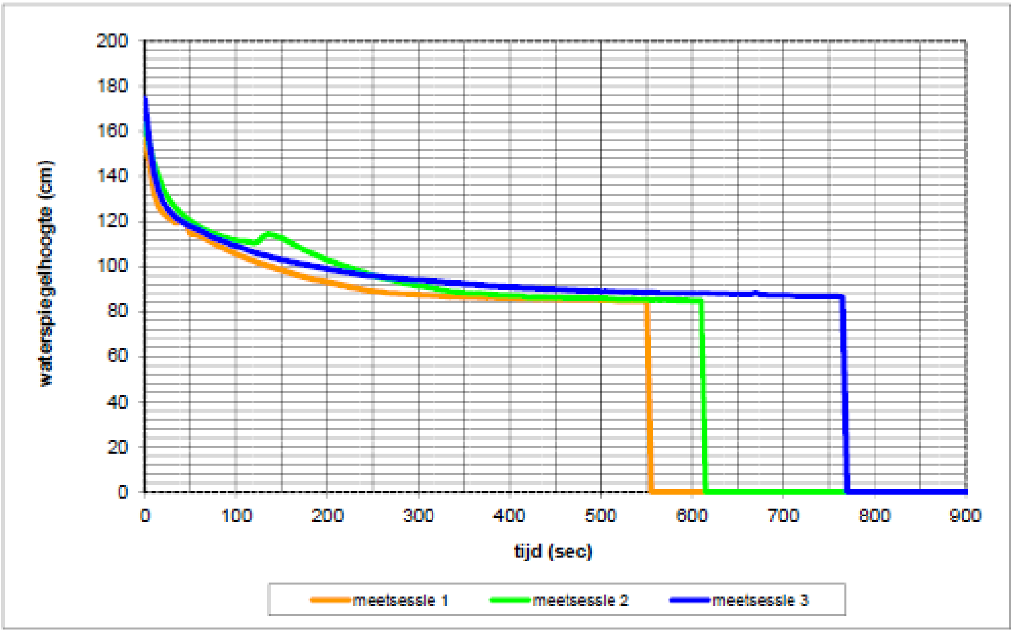
kf = 1,15 * r * (log(h0+r/2)-log(h1+r/2))/dt [cm/s]

Hierbij is :
h0 = waterhoogte in boorgat op tijdstip t = t0
h1 = waterhoogte in boorgat op tijdstip t = t1
r = boorgatradius
dt = verlopen tijd van t = t0 tot t = t1

Onderzoekswaarden

Diepte boorgat	D :	180	cm
Standaardhoogte	M :	0	cm
Radiusboorgat	R :	1,2	cm
Grondwater	W :	103	cm





Meetsessie 1

t0 =	0	sec
h0 =	156,33	cm
t1 =	550	sec
h1 =	84,52	cm
kf =	6,67E-06	m/s
kf =	0,58	m/dag
rc =	-1,31E-03	m/s

Meetsessie 2

t0 =	0	sec
h0 =	164,78	cm
t1 =	610	sec
h1 =	84,69	cm
kf =	6,51E-06	m/s
kf =	0,56	m/dag
rc =	-1,31E-03	m/s

Meetsessie 3

t0 =	0	sec
h0 =	174,41	cm
t1 =	760	sec
h1 =	86,68	cm
kf =	5,49E-06	m/s
kf =	0,47	m/dag
rc =	-1,15E-03	m/s

Geonius Geotechniek BV
Postbus 1097
6160 BB Geleen

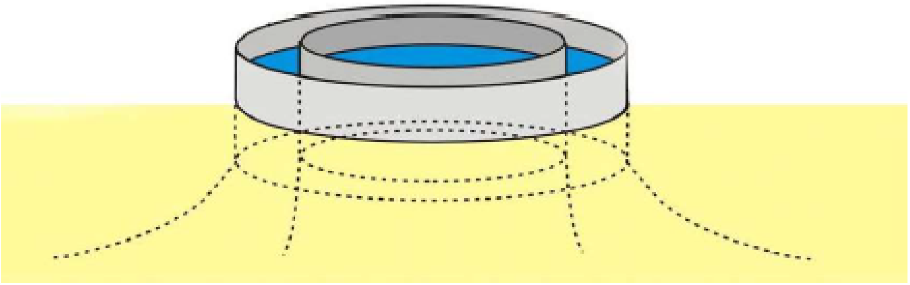
GEONIUS
CIVIEL GEOTECHNIEK MILIEU

Tel. 088-130 06 00
Fax. 088-130 06 69
info@geonius.nl

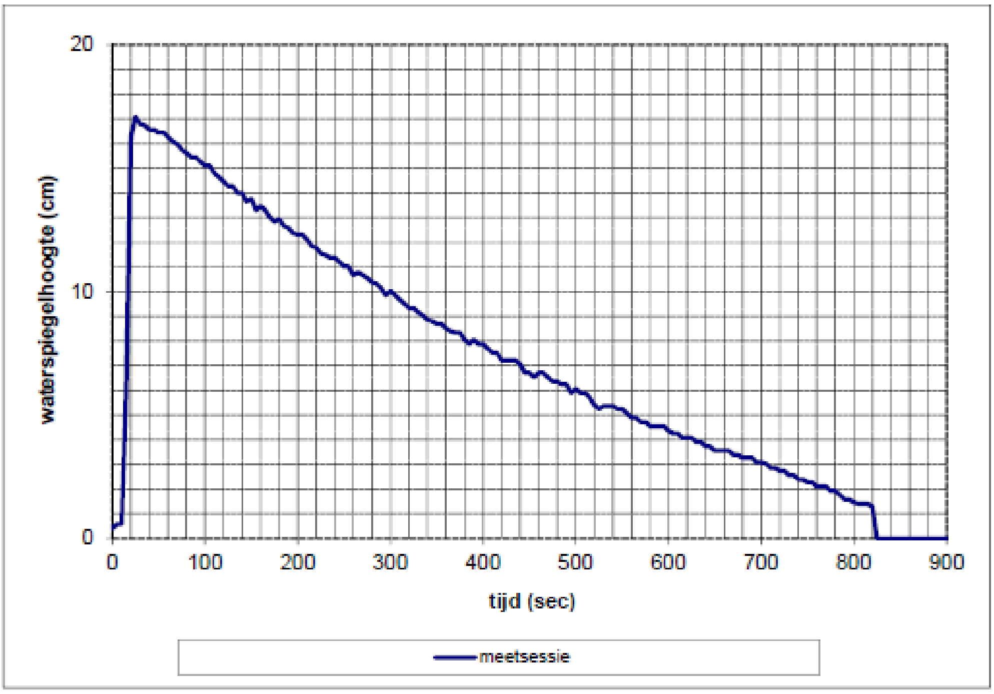
Annex VII: Measurement of unsaturated conductivity Klifrakplantsoen

Projektomschrijving: Onverzadigde doorlatendheid
Lokatie: Klifrakplantsoen
Meting: DM01
Opdrachtnr. Thesis Universiteit Utrecht

dubbele ring infiltrometer



Diameter binnenring: 32 cm
Diameter buitenring: 57 cm



Meetsessie

t0 =	25	sec
h0 =	17,09166667	cm
t1 =	820	sec
h1 =	1,283333333	cm
kf =	1,99E-04	m/s
kf =	17,18038	m/dag

Geonius Geotechniek BV
Postbus 1097
6160 BB Geleen

Tel. 08-130 06 00
info@geonius.nl
www.geonius.nl

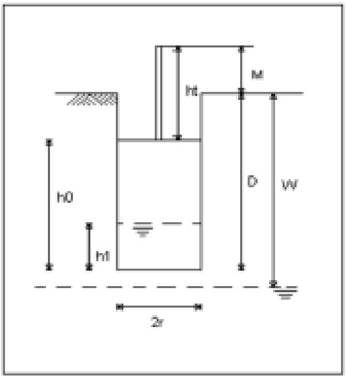
Annex VIII: Measurement of saturated conductivity Karel Doormanlaan

Projectomschrijving:	Verzadigde doorlatendheidsmeting	Opdrachtnr.	Thesis Universiteit Utrecht
Locatie:	Karel Doormanlaan	Traject (m-mv)	-
Boornummer:	DB01	Meting	DM01

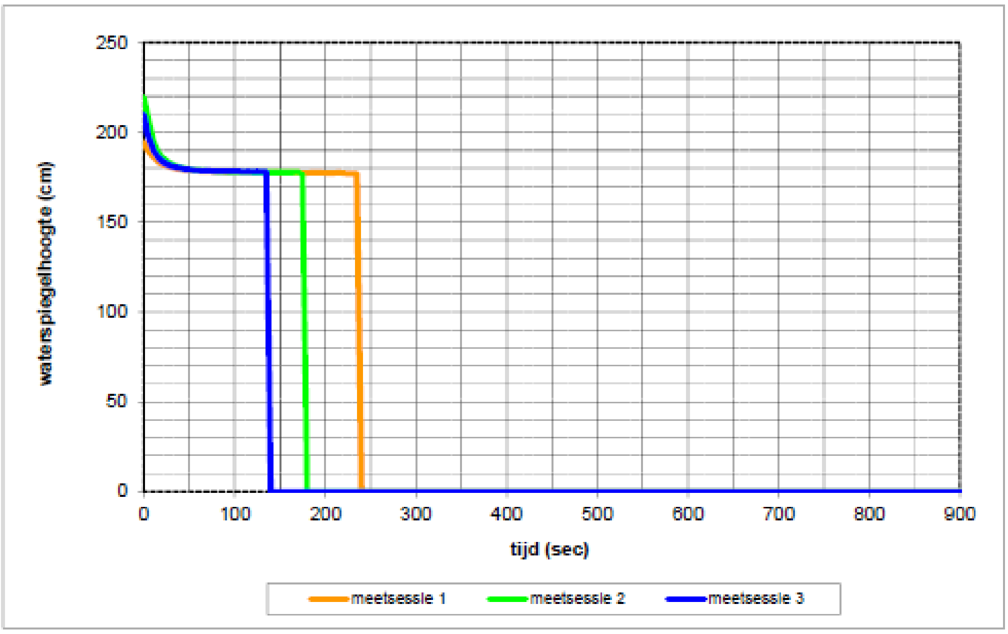
Formule om de doorlatendheid volgens Porchet te bepalen :

kf = 1,15 * r * (log(h0+r/2)-log(h1+r/2))/dt [cm/s]

Hierbij is :
h0 = waterhoogte in boorgat op tijdstip t = t0
h1 = waterhoogte in boorgat op tijdstip t = t1
r = boorgatradius
dt = verlopen tijd van t = t0 tot t = t1



Onderzoekswaarden			
Diepte boorgat	D :	240	cm
Standaardhoogte	M :	0	cm
Radiusboorgat	R :	1,2	cm
Grondwater	W :	63	cm

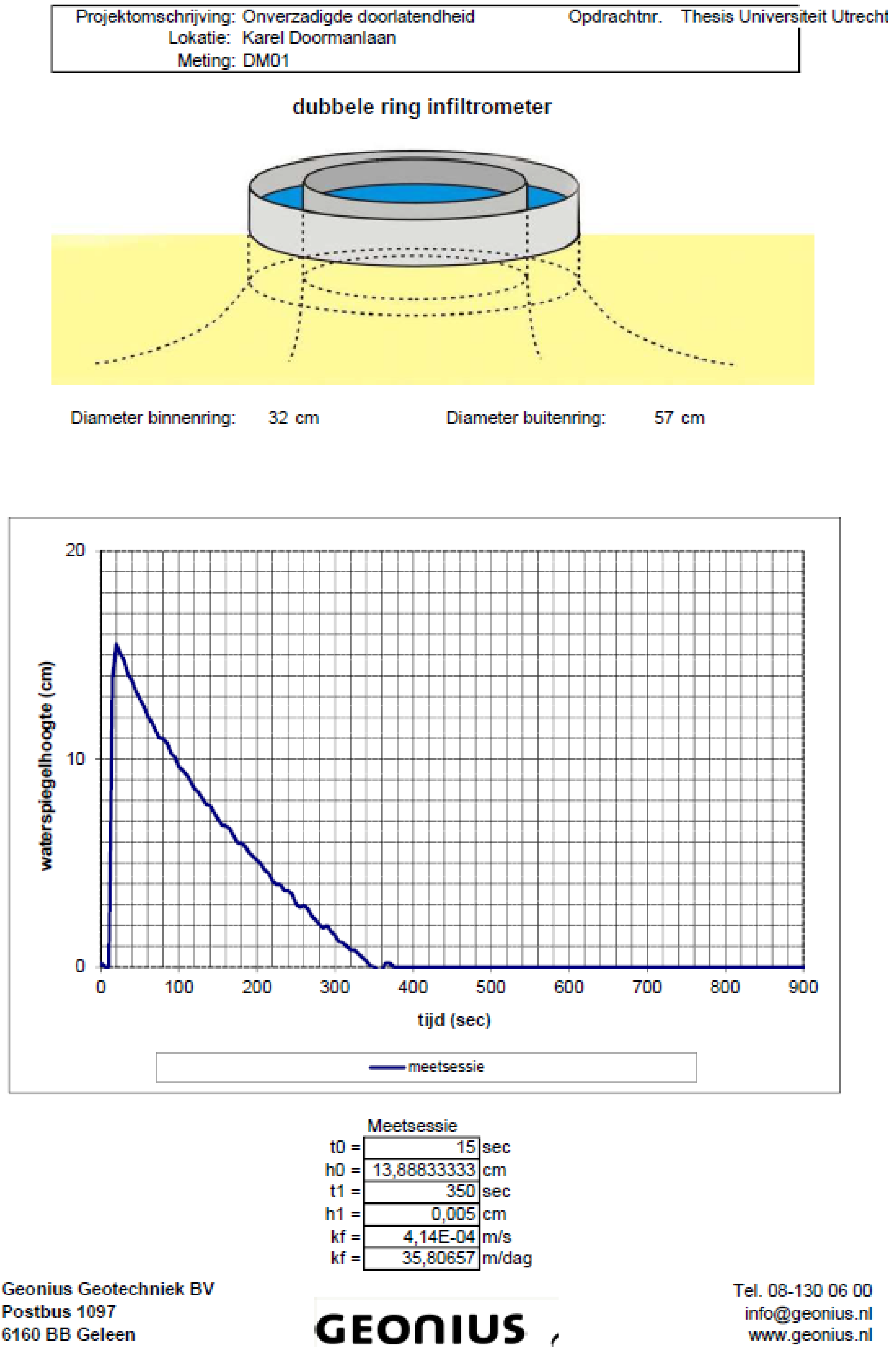


	Meetsessie 1			Meetsessie 2			Meetsessie 3	
t0 =	0	sec	t0 =	0	sec	t0 =	0	sec
h0 =	194,16	cm	h0 =	219,42	cm	h0 =	209,79	cm
t1 =	235	sec	t1 =	175	sec	t1 =	135	sec
h1 =	177,07	cm	h1 =	176,37	cm	h1 =	178,06	cm
kf =	2,34E-06	m/s	kf =	7,46E-06	m/s	kf =	7,26E-06	m/s
kf =	0,20	m/dag	kf =	0,64	m/dag	kf =	0,63	m/dag
rc =	-7,27E-04	m/s	rc =	-2,46E-03	m/s	rc =	-2,35E-03	m/s

Geonius Geotechniek BV
Postbus 1097
6160 BB Geleen

Tel. 088-130 06 00
Fax. 088-130 06 69
info@geonius.nl

Annex IX: Measurement of unsaturated conductivity Karel Doormanlaan



Projektomschrijving: Onverzadigde doorlatendheid Opdrachtnr. Thesis Universiteit Utrecht
Lokatie: Karel Doormanlaan
Meting: DM01

dubbele ring infiltrometer

Diameter binnenring: 32 cm Diameter buitenring: 57 cm

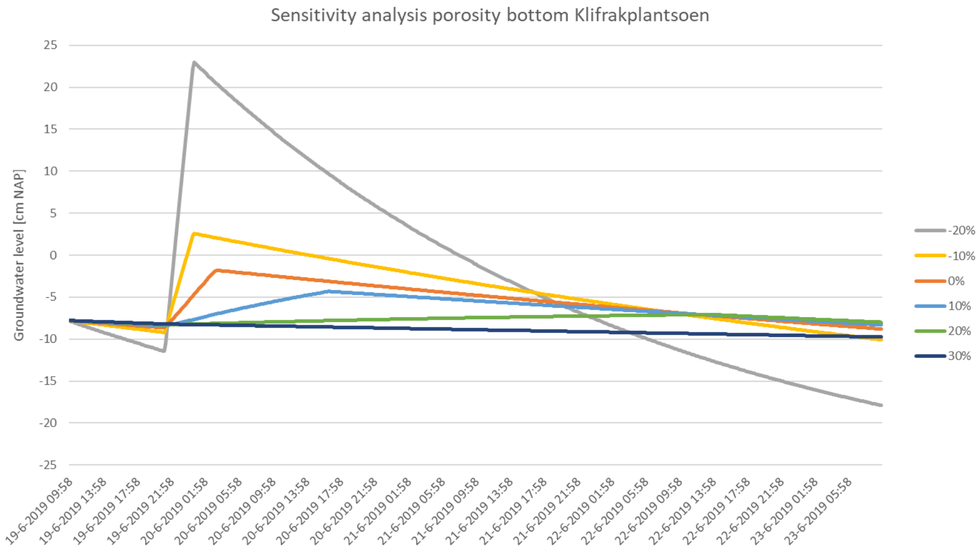
Meetsessie
t0 = 15 sec
h0 = 13,88833333 cm
t1 = 350 sec
h1 = 0,005 cm
kf = 4,14E-04 m/s
kf = 35,80657 m/dag

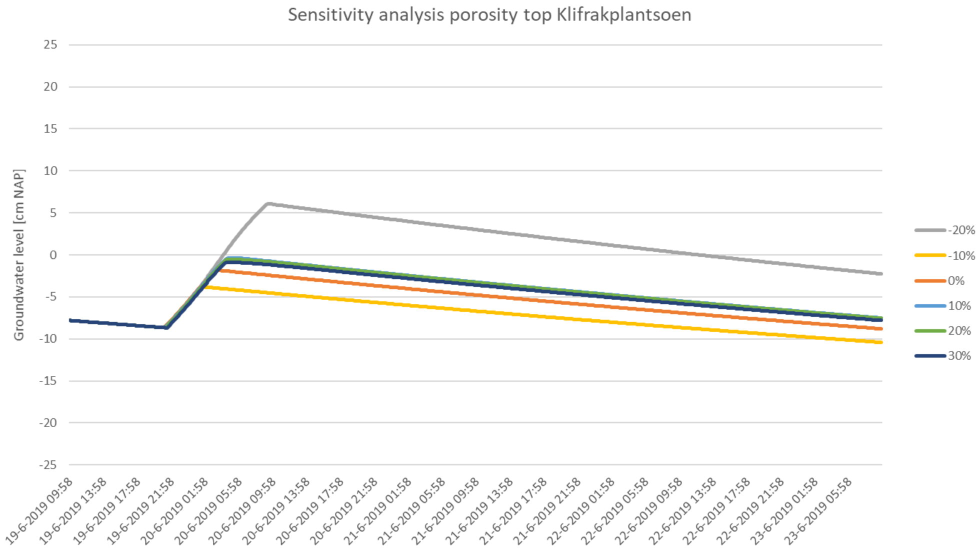
Geonius Geotechniek BV
Postbus 1097
6160 BB Geleen

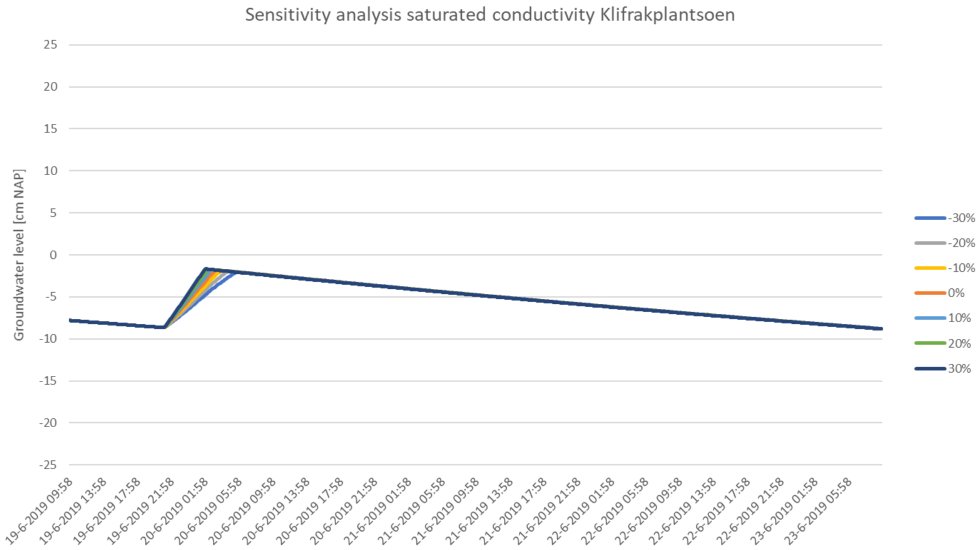
GEONIUS

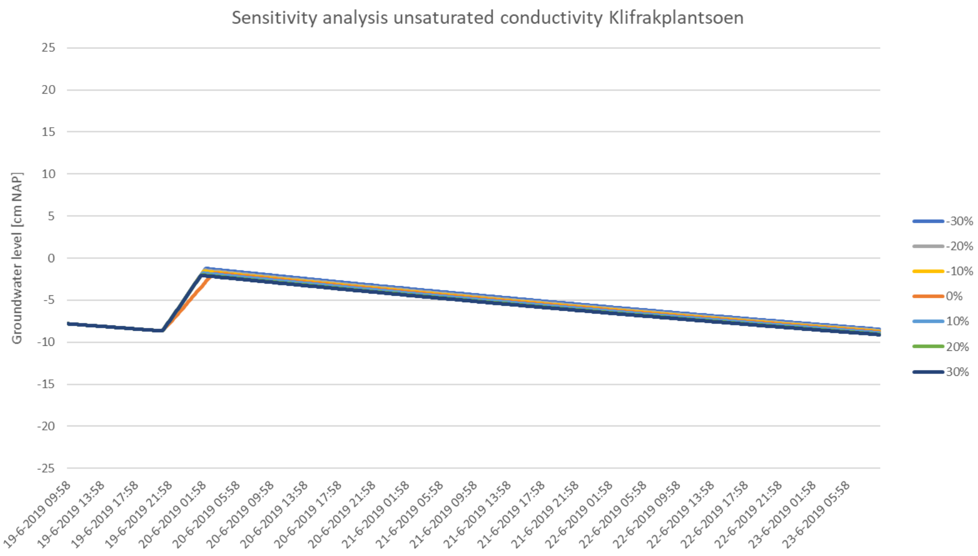
Tel. 08-130 06 00
info@geonius.nl
www.geonius.nl

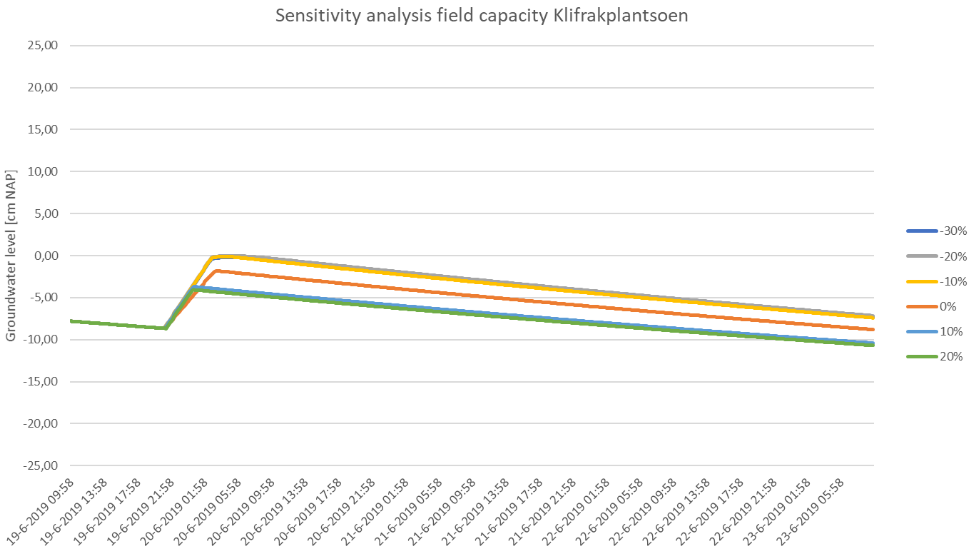
Annex X: Groundwater level sensitivity Klifrakplantsoen



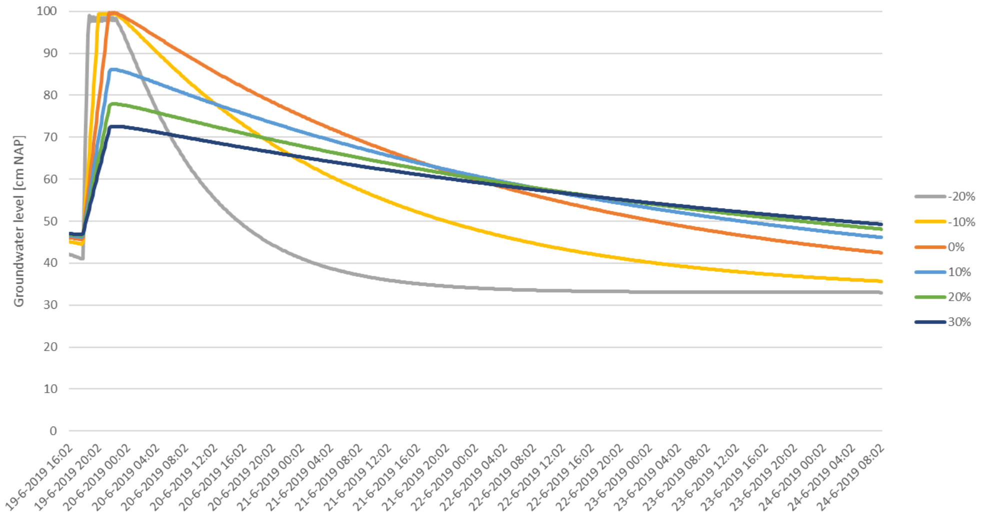




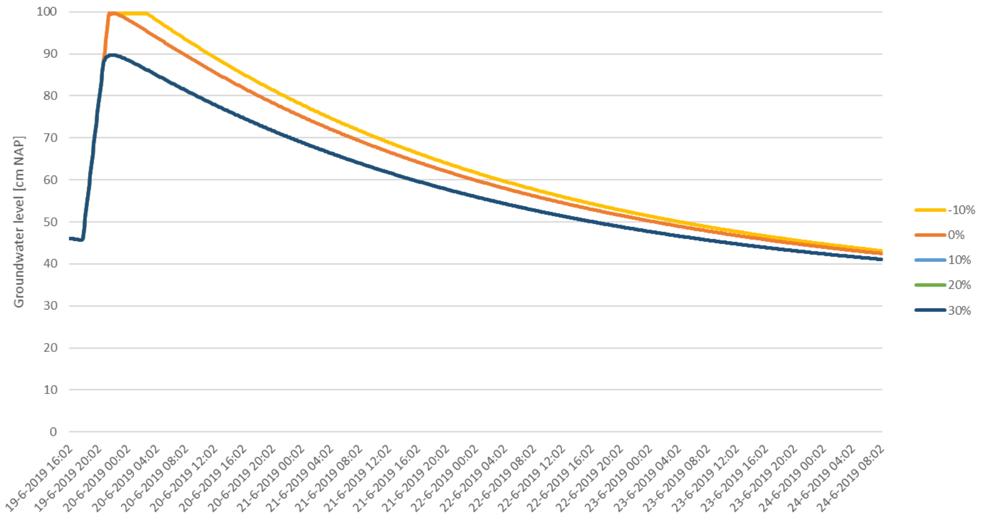

Sensitivity analysis unsaturated conductivity Klifrakplantsoen
Groundwater level [cm NAP]
25
20
15
10
5
0
-5
-10
-15
-20
-25
19-6-2019 09:58
19-6-2019 13:58
19-6-2019 17:58
19-6-2019 21:58
20-6-2019 01:58
20-6-2019 05:58
20-6-2019 09:58
20-6-2019 13:58
20-6-2019 17:58
20-6-2019 21:58
21-6-2019 01:58
21-6-2019 05:58
21-6-2019 09:58
21-6-2019 13:58
21-6-2019 17:58
21-6-2019 21:58
22-6-2019 01:58
22-6-2019 05:58
22-6-2019 09:58
22-6-2019 13:58
22-6-2019 17:58
22-6-2019 21:58
23-6-2019 01:58
23-6-2019 05:58
-30%
-20%
-10%
0%
10%
20%
30%


Sensitivity analysis field capacity Klifrakplantsoen
Groundwater level [cm NAP]
25,00
20,00
15,00
10,00
5,00
0,00
-5,00
-10,00
-15,00
-20,00
-25,00
19-6-2019 09:58
19-6-2019 13:58
19-6-2019 17:58
19-6-2019 21:58
20-6-2019 01:58
20-6-2019 05:58
20-6-2019 09:58
20-6-2019 13:58
20-6-2019 17:58
20-6-2019 21:58
21-6-2019 01:58
21-6-2019 05:58
21-6-2019 09:58
21-6-2019 13:58
21-6-2019 17:58
21-6-2019 21:58
22-6-2019 01:58
22-6-2019 05:58
22-6-2019 09:58
22-6-2019 13:58
22-6-2019 17:58
22-6-2019 21:58
23-6-2019 01:58
23-6-2019 05:58
-30%
-20%
-10%
0%
10%
20%

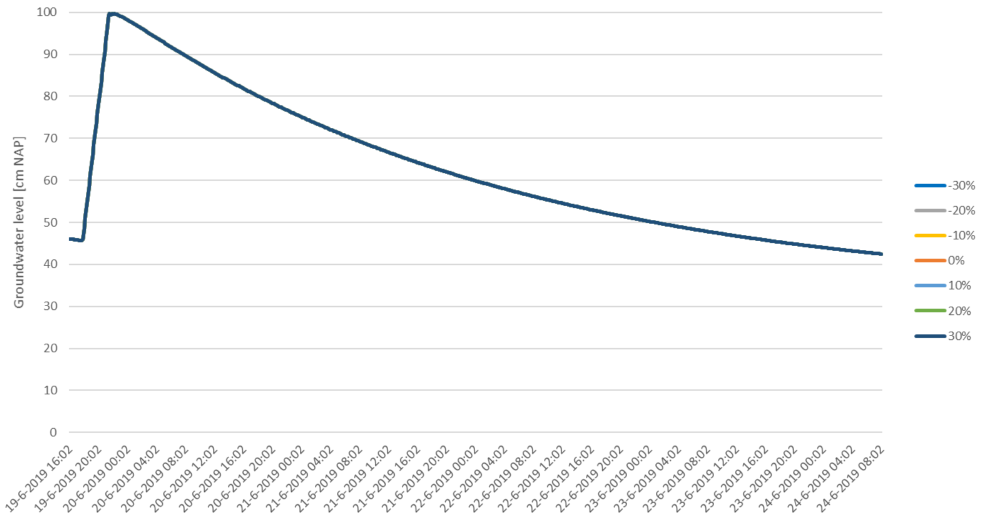
Annex XI: Groundwater level sensitivity Karel Doormanlaan



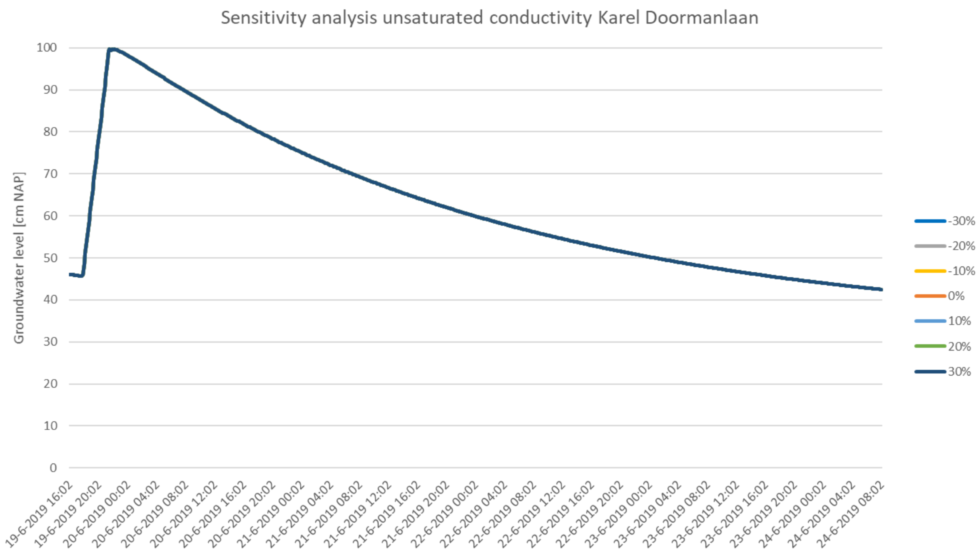
Sensitivity analysis porosity bottom Karel Doormanlaan

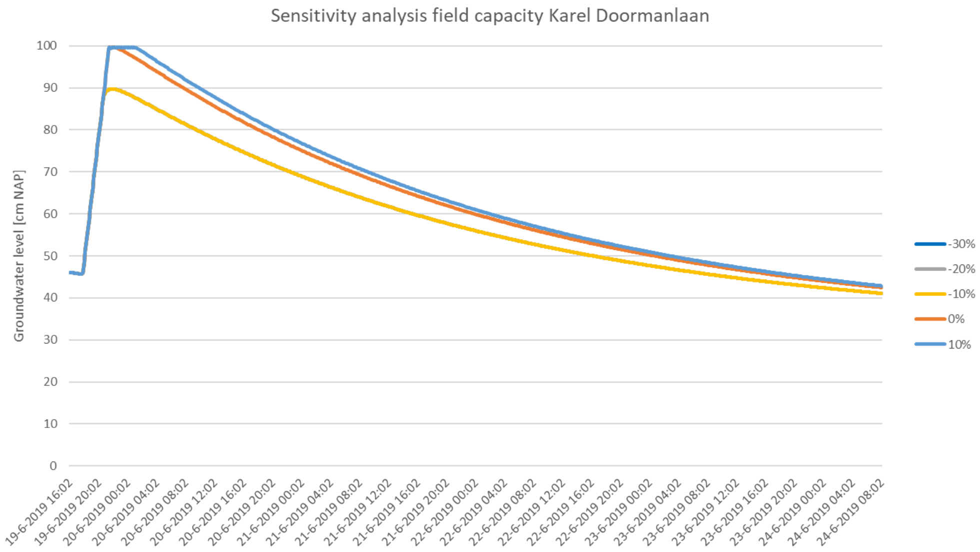


Sensitivity analysis porosity top Karel Doormanlaan



Sensitivity analysis saturated conductivity Karel Doormanlaan


Sensitivity analysis unsaturated conductivity Karel Doormanlaan
Groundwater level [cm NAP]
100
90
80
70
60
50
40
30
20
10
0
19-6-2019 16:02
19-6-2019 20:02
20-6-2019 00:02
20-6-2019 04:02
20-6-2019 08:02
20-6-2019 12:02
20-6-2019 16:02
20-6-2019 20:02
21-6-2019 00:02
21-6-2019 04:02
21-6-2019 08:02
21-6-2019 12:02
21-6-2019 16:02
21-6-2019 20:02
22-6-2019 00:02
22-6-2019 04:02
22-6-2019 08:02
22-6-2019 12:02
22-6-2019 16:02
22-6-2019 20:02
23-6-2019 00:02
23-6-2019 04:02
23-6-2019 08:02
23-6-2019 12:02
23-6-2019 16:02
23-6-2019 20:02
24-6-2019 00:02
24-6-2019 04:02
24-6-2019 08:02
-30%
-20%
-10%
0%
10%
20%
30%


Sensitivity analysis field capacity Karel Doormanlaan
Groundwater level [cm NAP]
100
90
80
70
60
50
40
30
20
10
0
19-6-2019 16:02
19-6-2019 20:02
20-6-2019 00:02
20-6-2019 04:02
20-6-2019 08:02
20-6-2019 12:02
20-6-2019 16:02
20-6-2019 20:02
21-6-2019 00:02
21-6-2019 04:02
21-6-2019 08:02
21-6-2019 12:02
21-6-2019 16:02
21-6-2019 20:02
22-6-2019 00:02
22-6-2019 04:02
22-6-2019 08:02
22-6-2019 12:02
22-6-2019 16:02
22-6-2019 20:02
23-6-2019 00:02
23-6-2019 04:02
23-6-2019 08:02
23-6-2019 12:02
23-6-2019 16:02
23-6-2019 20:02
24-6-2019 00:02
24-6-2019 04:02
24-6-2019 08:02
-30%
-20%
-10%
0%
10%