

The challenge of parameter uncertainty:

Finding parameter distributions from hydrological field data for conceptual rainfall-runoff models.

Nicolette van der Tak

**The challenge of parameter uncertainty:
Finding parameter distributions from hydrological field data for
conceptual rainfall-runoff models.**

By
Nicolette van der Tak

To obtain the degree of Master of Science
at the Delft University of Technology
to be defended publicly on Friday December 7th, 2018 at 14:00 p.m. (GMT+1)

Student number:: 4496507
Project duration: March 19, 2018 – December 7, 2018

Thesis committee:
Dr. MSc. habil. M. Hrachowitz
Dr. T.A. Bogaard
Prof. dr. ir. T.J. Heimovaara
Ir. S.E. van den Driest-van der Kruijs

Delft University of Technology
Delft University of Technology
Delft University of Technology
Antea Group



Abstract

Conceptual hydrological models attempt to describe the non-linear behavior between climate drivers (precipitation and evaporation) and system output (e.g. stream or river flow). These models are a simplified and abstract representation of a catchment and should represent system characteristics such as topography and geology. The models are a tool to simulate the movement of water in the hydrological cycle. Hydrologists use the models for research, water resource assessment, planning and management of water.

System representation in models is a challenge due to underlying uncertainties. Two main categories can be distinguished: data uncertainty and model uncertainty. Data uncertainty originates from difficulties in the measurements and from the interpretation of the data. Model uncertainty consists of two kinds of uncertainties: the model structure (e.g. is the model simplification appropriate?) and uncertainty in the model parameters. Effective parameters are used in conceptual models. The model parameters represent characteristics or properties of a catchment, such as the maximum storage capacity of the soil. These parameters are not directly based on observable quantities. They are an integration of spatial heterogeneous parts of the system above the scale of available observations. To estimate these parameters, calibration of the models is needed.

The focus of this thesis is the parameter uncertainty in models. Many studies have been conducted to increase the physical meaning of the parameters. The use of constraints (regarding relations between parameter, fluxes and states) for the parameter sets prevents the model from overfitting in the calibration period. Finding adequate parameter values based on field observations, which represents the heterogeneity of a catchment on the spatial resolution and scale of the model, is considered much more challenging since we cannot directly derive the parameters from observable quantities.

The objective of this thesis is to analyze the extent to which it is possible to make an estimation of parameter values or parameter distributions based on field observations (the precipitation, evaporation and discharge) and a given hydrological model structure. The goal is to avoid the use of uninformed prior parameter distributions during calibration by using available field data to generate informed prior distributions. The hypothesis is that a hydrograph can be divided into sub-periods and that each period has different “dominant” processes with associated parameters. Dividing the hydrograph into sub-periods to find individual parameters of the model could further increase model realism. In theory, uncertainty in the selected parameter sets should be reduced if the parameter information is directly extracted from the observed data.

In this thesis, six different expert-knowledge inverse modelling methods are developed to find four parameter distributions. Each method uses sub-periods in the data and is coupled to the parameter of the model component representing that specific type of sub-period. For the first parameter, the ground water drainage parameter (C_s) of the slow reservoir, the fit and S-D methods are developed. These methods focus on long dry periods where the discharge of the slow reservoir is dominant. In the second method, percolation rates are calculated in periods with low flow. The parameter distribution restraining the maximum percolation rate (P_{max}) is determined using calculated percolation rates. For the third parameter, which regulates peak discharges (D), clear individual peak discharge moments need to be filtered out in a discharge series. The method to find the D distribution uses the most suitable sub-periods, which are during spring and fall season. For the last parameter, distribution of the maximum catchment storage ($S_{U,max}$), a water balance method is developed over sub-periods of the data set. In the bound method, winter periods are used for the lower bound of the distribution while the upper bound originates from summer periods. The inter-peak method is also based on a water balance but focuses on two subsequent peak flow moments.

To test the expert-knowledge inverse modelling methods, the study was conducted in a synthetic environment, which made it easier to validate the parameter distributions obtained with the expert-knowledge inverse modelling methods. In this synthetic experiment, discharge data was produced by a model driven by real rainfall data and potential evaporation data. All forms of uncertainty were excluded.

The effect of data uncertainty in the methods was investigated separately by conducting a sensitivity analysis. The same synthetic data was used; however the synthetic data was corrupted to

simulate data uncertainty. The validation of the methods was achieved based on the original parameters used to produce the discharge data.

Last, an application of the methods upon real measured data was completed. The performance of the methods to find parameter distributions can no longer be assessed since in the real world the “correct” parameter values are not known. However, a comparison of the resulting informed prior parameter distributions of the methods with uninformed prior parameter distributions could be made with a Monte-Carlo sampling strategy calibration.

In the synthetic experiment, all parameter distributions of the investigated model were correctly determined using the expert-knowledge inverse modelling methods. The sensitivity analysis revealed that the method to determine the P_{max} parameter distribution was sensitive to data uncertainty. The determined P_{max} parameter distributions did not include the original parameter of the corrupted synthetic data. Since many of the expert-knowledge inverse modelling methods use P_{max} as input, other parameter distributions will be affected by the errors in the P_{max} distribution. However, this issue does not lead to parameter distributions that do not include the original parameters for the affected parameters.

In real-world application, insight is gained into the performance of the developed methods to find parameter distributions. A comparison was made in a Monte-Carlo sampling strategy between the calibration results obtained using an uninformed prior parameter distributions and the informed parameter distributions of the expert-knowledge inverse modelling methods. An uncertainty interval was constructed with the Generalized Likelihood Uncertainty Estimation (GLUE) method. The total area of the constructed uncertainty interval using the calibration results of the informed prior parameter distributions was less than half than the uncertainty interval constructed using the uninformed prior parameter distributions. The posterior parameter distributions of the informed parameter distributions was two to five times smaller than for the uninformed parameter distribution. The model performance of both calibrations did not deviate significantly, indicating sufficient performance of the expert-knowledge inverse modelling methods to find parameters. If the performance of the model using the informed parameter distribution decreased significantly compared to uninformed parameter calibration, could this decrease be an indication that the model structure is unsuitable for the catchment and essential hydrological processes are left out.

Further research should provide insight into the performance of the expert-knowledge inverse modelling methods in a synthetic experiment with model uncertainty. The assumption could be confirmed in this test that the expert-knowledge inverse modelling methods would be an extra indication for how suitable a model structure is for a catchment. In addition, further research should be conducted to see how the expert-knowledge inverse modelling methods behave in a larger catchment with more heterogenic characteristics.

Table of contents

Abstract	3
List of figures.....	7
List of tables.....	9
List of abbreviations.....	10
Chapter 1 Introduction	12
1.1 Hydrological modelling.....	12
1.2 Model types.....	12
1.3 Challenges in conceptual modelling.....	12
1.3.1 Data uncertainty	13
1.3.2 Model uncertainties.....	13
1.4 Scope.....	14
1.5 Research objective	14
1.6 Research questions.....	15
1.7 Thesis outline.....	16
Chapter 2 Data and Models	18
2.1 Study area	18
2.1.1 Kervidy-Naizin catchment, France	18
2.1.2 Precipitation stations, the Netherlands.....	19
2.2 Model description	19
2.3 Synthetic data.....	20
Chapter 3 Methods	22
3.1 Methods to determine parameter distribution	22
3.1.1 Groundwater drainage parameter (C_S).....	22
3.1.2 Partitioning coefficient fast and slow reservoir (D).....	26
3.1.3 Maximum unsaturated zone storage ($S_{U,max}$).....	27
3.1.4 Maximum percolation rate (P_{max}).....	31
3.2 Synthetic experiment	33
3.3 Sensitivity analysis methods for data uncertainty	35
3.3.1 Precipitation error	35
3.3.2 Data corruption	36
3.4 Real-word application.....	36
3.4.1 Comparison traditional and new calibration approach	37
Chapter 4 Results	39
4.1 Results of parameter distribution methods.....	39
4.2 Results synthetic experiments	42
4.2.1 Results for model 4.....	42
4.2.2 Results for models 7 and 9	43

4.3 Results of sensitivity analysis	46
4.3.1 Results for model 9 including the forcing error	46
4.3.2 Results of model 9 including the data error	47
4.4 Results of real-world application	48
4.4.1 Results comparison.....	49
Chapter 5 Discussion.....	52
5.1 Methods and synthetic experiment.....	52
5.1.1 Methods.....	52
5.1.2 Comparison model run of models 4, 7 and 9	52
5.2 Sensitivity analysis	53
5.2.1 Precipitation error.....	53
5.2.2 Data corruption.....	53
5.3 Real world.....	54
Chapter 6 Conclusions.....	55
6.1 Further research.....	56
Literature list	57
Appendix 1 Model descriptions	60
Appendix 2 Evaporation distribution function	66
Appendix 3 Parameter values and discharge data synthetic experiment	68
Appendix 4 Method results.....	76
C_s fit method results.....	76
C_s S-D method	78
D method	79
$S_{U,max}$ bounds method.....	80
$S_{U,max}$ inter-peak method	81
P_{max} method.....	82
Appendix 5 Results synthetic experiments.....	83
Model 4.....	83
Model 7.....	85
Model 9.....	87
Appendix 6 Results data uncertainty	89
Precipitation error.....	89
Data corruption.....	91

List of figures

Figure 1. A visual representation of the objective of this research. The objective is to estimate the parameter values or parameter distribution based on field observations and a given hydrological model structure.....	15
Figure 2. Representation of research outline: theory portion where methods are developed to find parameter distributions, modelling portion to test the developed methods, which is divided into three sub-parts: (1) synthetic experiment, (2) sensitivity analysis and (3) real-world application. Conclusions are drawn in the last section.....	17
Figure 3. Kervidy-Naizin catchment in France, which is part of the ORE-ArtgHys network (Aubert, et al., 2013).....	18
Figure 4. Measuring stations Twenthe and Hupsel, part of the KNMI network (www.meteobase.nl).....	19
Figure 5. A schematic representation of the hydrological model structures used in this study. The fluxes and reservoir names are shown in black. The related parameters are shown in red.....	21
Figure 6. Observed time series of hourly precipitation (light blue), hourly potential evaporation (orange) and discharge (blue dotted line). Data originates from model 1 parameter set 3 in the years 2005-2011. Example of the selected recession (black) was used in the fit method. The recession has a length of 10 h, no precipitation was measured two days prior nor during the recession, no potential evaporation was measured during the recession and the end of the recession was just before the start of a precipitation event.....	23
Figure 7. Observed time series of hourly precipitation (light blue), hourly potential evaporation (orange) and discharge (blue dotted line). Overview of all selected recessions (black) were used in the fit method. Data originates from model 1 parameter set 3 in the years 2005-2011.....	23
Figure 8. A selected recession from the data of model 4 parameter set 1 (blue dotted line) with an original C_s value of 200 h and a fit (orange line) with an estimated C_s value in the fit method of 199.5 h.....	24
Figure 9. A selected recession from the data of model 9 parameter set 1 (left); the corresponding storage-discharge relation with estimated C_s value of 636 h (slope S-D relation) (right).....	25
Figure 10. Original storage-discharge relation, time integral over the recession (blue line) and the new recalculated (orange line) storage-discharge relation, time integral over the difference between the recession (outflow of the slow reservoir) and the P_{tot} (inflow to the slow reservoir). Data originates from model 9 parameter set 1.....	25
Figure 11. Selected overflow moment observed discharge (orange, red and brown dots); expected flow during the peak flow calculated with the C_s (green dot) and expected flow from the saturated zone calculated with the C_s (purple dot). The difference between the red and purple dots represents the volume of fast lateral runoff processes, while the difference between the purple and green dots represents the extra discharge from the slow reservoir due to additional water from the preferential recharge.....	26
Figure 12. Result of the bound method using different event lengths for data originating from model 4; the estimated minimum value of $S_{U,max}$ = 160 mm, and the estimated maximum = 287.5 mm.....	28
Figure 13. Selected individual event for the time series originating from model 4 set 2; $S_{U,max}$ = 235 mm; length of the event is 150 days.....	28
Figure 14. Selected individual event for the time series originating from model 4 set 2; $S_{U,max}$ = 235 mm; length of the event is 150 days.....	28
Figure 15. Previous precipitation of the selected event already partly filled the system, resulting in a low minimum and maximum $S_{U,max}$	29
Figure 16. Part of the hydrograph starting and ending with an overflow moment with a storage change of the unsaturated zone in between due to evaporation (orange line) and precipitation (light blue line). This part is used in the inter peak method.....	29
Figure 17. Observed discharge and filtered discharge of model 9 data set 2. In the filtered discharge all peak flows are excluded for the P_{max} method.....	31

Figure 18. The blue line represents the observed discharge. The orange and red dotted lines represent expected discharges if no percolation would be present. The difference between the red and blue dots represents the extra discharge from the slow reservoir due to the additional percolated water in the slow reservoir. 32

Figure 19. The green dots represent all the time steps used in the P_{max} analysis..... 32

Figure 20. The green dots represent the top 20% highest calculated percolation rates..... 33

Figure 21. The interdependencies of the different expert-knowledge inverse modelling methods to find the parameter distributions 33

Figure 22. Schematic representation of the procedure to find the parameter distributions of models 7 and 9. 34

Figure 23. Cumulative sum of the measured precipitation (mm) in Hupsel compared to the cumulative sum of the measured precipitation (mm) in Twenthe (data of 2007) 35

Figure 24. Observed (blue line) and most optimal simulated (orange line) discharges in the calibration and validation periods. The simulated discharge is shown with a 95 percent uncertainty interval (shaded grey area)..... 50

Figure 25. The prior parameter distributions and the posterior parameter distributions of all behavioral models. For both the uninformed parameter distribution and the informed parameter distribution. (boxplots: the dots represents the outliers in de data set, the lower and upper whisker 2.5/97.5th percentiles and, the horizontal orange line the median)..... 51

List of tables

Table 1. Water balance equations of the models used in this study; the x indicates presence in the model structure.....	20
Table 2. Constitutive formulas in the model structures; the x indicates presence in the model structure	20
Table 3. Annual volumes for precipitation and potential evaporation Twenthe and Hupsel stations .	35
Table 4. Prior parameter bounds for MC simulation of model 9	37
Table 5. Objective functions used to assess the model performance	38
Table 6. Examples of the results of the synthetic experiment using the C_s fit method.....	39
Table 7. Examples of the results of the synthetic experiment using the S-D method	40
Table 8. Examples of the results of the synthetic experiment using the D method.....	40
Table 9. Examples of the $S_{U,max}$ results of synthetic experiment using the inter peak method.....	41
Table 10. Examples of the results of synthetic experiment using the P_{max} method.....	41
Table 11. Summary of the results of synthetic experiment model 4 for C_s and D	42
Table 12. Summary of the results of synthetic experiment for model 4 for $S_{U,max}$	42
Table 13. Summary of the results of synthetic experiment for model 7 for C_s and P_{max}	43
Table 14. Summary of the $S_{U,max}$ results of synthetic experiment for model 7	44
Table 15. Summary of the results of synthetic experiment for model 9 for C_s , P_{max} and D	44
Table 16. Summary of the $S_{U,max}$ results of synthetic experiment for model 9	45
Table 17. Summary of the results for $S_{U,max}$ with precipitation error.....	46
Table 18. Summary of the results for C_s , P_{max} and D parameter distributions with precipitation error.	46
Table 19. Summary of the results for $S_{U,max}$ with data corruption.....	47
Table 20. Summary of the results for C_s , P_{max} and D parameter distributions with data corruption.....	47
Table 21. Results of real-world application for C_s , P_{max} and D	48
Table 22. Results of real-world application for $S_{U,max}$	48
Table 23. Performance results of the uninformed prior parameter distribution and the informed prior parameter distributions over multiple periods. The mean performance represents the mean of all simulated discharges in the uncertainty bound, the most balanced (optimal) performance in the uncertainty bound is also given in this table.....	49
Table 24. Sunrise and sunset time of Naizin (loosely based on https://www.aroundtheworld360.com/sunrise_sunset/france/naizin/)	66

List of abbreviations

	Unit	
C_s	mm/h	Groundwater drainage parameter
D	-	Partitioning coefficient fast and slow reservoir
E	mm/h	Evaporation
ED	-	Euclidean distance
E_p	mm/h	Potential evaporation
GLUE		Generalized Likelihood Uncertainty Estimation
logNS	-	Nash-Sutcliffe model efficiency of the logarithm of the flow
MRC		Master recession curve
NS	-	Nash-Sutcliffe model Efficiency
Perc	mm/h	Percolation
P_{max}	mm/h	Maximum percolation rate
P_{tot}	mm/h	Inflow to the slow reservoir
Q_f	mm/h	Fast lateral runoff discharge
Q_s	mm/h	Slow discharge
Q_t	mm/h	Total discharge
Q_{uf}	mm/h	Overflow unsaturated zone
R_s	mm/h	Preferential recharge
S-D relation		Storage-Discharge relation
S_s	mm	Slow reservoir
std		Standard deviation
$S_{U,max}$	mm	Maximum unsaturated zone storage
S_U	mm	Unsaturated zone reservoir
T	mm/h	Transpiration

Chapter 1 Introduction

1.1 Hydrological modelling

Hydrological models attempt to describe the non-linear behavior between climate drivers (e.g. rainfall, potential evaporation) and system output (e.g. stream or river flow, groundwater table fluctuations). The observed climatic drivers are included in mathematical equations describing the fluxes and different system states (water storage) to reproduce the observed hydrological response. The models are a tool to simulate the movement of water in the hydrological cycle. The hydrological models are a simplified and abstract representation of our hypothesis on the processes occurring in nature (Savenije, 2009).

Hydrological models are used for many reasons, for example, as powerful tool in research as explained by Savenije (2009). In models, we try to encapsulate our knowledge. The models are the translation of our perception of the complex processes acting in nature. After formulating the model and the accompanied mathematical or numerical expressions of the processes, we need to carefully test these models to avoid misrepresentation of real-world processes. If the model is able to reproduce the physical processes, these well-tested models can be used to test the impact of interventions in the hydrological system. For example, a model might test what happens when the climate changes or if the land use in a catchment changes. Different analyses can be carried out to assess the availability of water or extrapolate in time to use the models for future predictions. The well-tested models give insight into the overall behavior of the system at a level unobservable in reality (Savenije, 2009).

Assuming the models are well-tested and capable of reproducing the physical processes, these models are used as a tool to develop policy for water resource management (Hrachowitz & Clark, 2017). Changes in land use or climate regime can be tested to quantify the effect of these factors on flood risk or water availability. Furthermore, effects of water withdraw or storage creations could be analyzed. Last, models are used for early warning of floods or droughts. This information is used in the daily operation of our water control systems (Brauer, 2014; Savenije, 2009).

1.2 Model types

Hydrological models can have different degrees of complexity. The least complex models are called black box models. These are based on empirical relations, for example, the unit hydrograph (Clark, 1945). In contrast to these simple models, the most complex models are often referred to as distributed physically-based models, with high complexity and spatial resolution. These models are based on differential equations, such as the Darcy-Richards' equation for groundwater flow and the St. Venant equations for open water. MIKE-SHE (Refsgaard & Storm, 1995) is an example of these models.

In between these two types of models are conceptual parametric rainfall-runoff models. These models consist of reservoirs representing storages and processes interacting with the storages (fluxes) (Savenije, 2009; Brauer, 2014; Hrachowitz & Clark, 2017). Fluxes always occur per unit of time, and examples are precipitation, evaporation, percolation, subsurface flow and discharge. They can be represented by observed data or mathematically described. Examples of conceptual models are HBV (Bergström & Forsman, 1973), FLEX (Fenicia, et al., 2006) and WALRUS (Brauer, et al., 2014). This study uses conceptual model structures.

1.3 Challenges in conceptual modelling

The most prominent challenge in conceptual modelling is to identify a quantitative and functional relationship between climate drivers and output, like river discharge at the catchment scale. The models are a mathematical replica of the real hydrological system, which should reflect catchment characteristics like geology and topography (Gharari, 2016). This challenge originates from different underlying sources of uncertainties. Two main categories can be distinguished, data uncertainty and model uncertainty, and result in unreliable predictions even though the models perform well during the calibration period (Renard, et al., 2010; Hrachowitz & Clark, 2017). This could indicate an insufficient

representation of underlying processes (Seibert, 1997; Hrachowitz, et al., 2014). The next sections will elaborate on the uncertainties in modelling.

1.3.1 Data uncertainty

The driver in the modelling and calibration process is measured data. Therefore data uncertainty has a large influence on the performance of the models (Kavetski, et al., 2006; Ajami, et al., 2007). Both input data (like rainfall and evaporation) and measured output, like discharge, contain uncertainties. The data uncertainty originates from errors and uncertainty in the measurement technique. Interpretation errors also have a large influence on the total data uncertainty (Renard, et al., 2010).

Interpretation errors are, for example, the extrapolation of the measured data over the catchment. Rainfall input into the models is typically measured by rain gauges. This measurement represents a small area of several square centimeters, while the hydrological model acts on a scale exceeding various square kilometers (Gharari, 2016). Rainfall is known to be a heterogeneous process (highly variable in both space and time), but in modelling, it is assumed to be uniform over (part of) the catchment, forming an important source of uncertainty (Butts, et al., 2004; Kavetski, et al., 2006; Gharari, 2016). In addition to the rainfall data, is evaporation another climate driver. Evaporation is also known to be a heterogeneous process and is also assumed to be uniform over (part of) the catchment, forming an extra source of uncertainty.

The potential evaporation is often used as model input and forms another source of data uncertainty. Potential evaporation cannot be measured but is estimated using empirical formulas like the Penman equation (1948). The convergence to potential evaporation is achieved based on measurements of air temperature, wind velocity, net solar radiation and relative humidity. Since the potential evaporation is indirectly estimated, a larger uncertainty is associated with this forcing data.

The estimation of discharge data includes multiple sources of uncertainty. Different techniques, such as rating curves, have been developed to determine discharge. For rating curves, measurements of the water level and flow velocity are needed. In particular, during high and extremely low flows, these measurements include errors due to underflow or bypassing of the gauging stations (Savenije, 2009). In addition to the measuring errors, errors in the rating curve itself are a large uncertainty source. Hydraulic conditions change over time or flow rate, which influences the rating curve. These changes in rating curves are often not considered, which makes the discharge estimates uncertain. Furthermore, more modern techniques for measuring discharge, like an Acoustic Doppler Current Profiler boat, have uncertainties in the measurements (González-Castro & Muste, 2007).

1.3.2 Model uncertainties

All models are our perception of the world, and models are simplifications of the hydrological system, regardless of the degree of spatial or physical complexity in a particular model (Beven, 2011; Gupta, et al., 2012). The model's uncertainty is a consequence of the simplification and assumptions made by developing the model with a mathematical hypothesis (Clark, et al., 2008; Renard, et al., 2010; Gupta, et al., 2012). Identifying and selecting the most appropriate model structure is a significant challenge for the hydrological community (Clark, et al., 2008). Savenije (2009) argued that hydrologist should always remember a degree of "art" is included when developing a model structure. Many studies have focused into finding more appropriate model structures (Clark, et al., 2008; McMillan, et al., 2011; Willems, et al., 2014).

Another challenge in modelling the hydrological system is the presence of numerous non-linear processes. Stream discharge often depends non-linearly on rainfall, which implies that unique precipitation-discharge relationships do not exist (Savenije, 2009). There are several reasons for this non-linearity, such as the hysteretic processes. An example of a hysteric process is the water flow in the unsaturated zone that acts differently under wet and dry conditions. Another cause of non-linearity is threshold behavior. An example is the interception process (immediate evaporation from water stored on the canopy and soil surface). This is the first process after the start of a rainfall event and will start

other processes only when it reaches maximum capacity. The maximum capacity is different for each catchment and season (for example, in summertime, interception storage is large due to the leaves on the trees). Last, non-linear equations, such as the St. Venant equations, contribute to the non-linearity of a system. Lack of knowledge about the non-linearity in a system forms a source of uncertainty in the model structure (Savenije, 2009).

The issue of scale adds to the uncertainty in hydrological models. Hydrological processes act on many different scales. We can observe processes in scales from millimeters to maybe 100 meters. The question arises whether hydrological processes acting on the catchment scale are similar to our observable scale (Beven, 2006; Savenije, 2009). How processes act on a larger scale and how we model the processes are connected to the fact that models are our perception of the complete hydrological cycle. This makes it likely that our assumptions are incorrect.

Parameter uncertainty also plays a role in model uncertainty. In conceptual models, effective parameters are used. These parameters are not directly based on observable quantities. The small-scale observations, for example measurements of the soil hydraulic conductivity, often fail to represent the larger features, like macropores, of a system. Parameters are frequently an integration of spatial heterogeneous parts of the system above the scale of available observations. To estimate these parameters, calibration of the models is needed (Gharari, et al., 2014; Willems, 2014; Hrachowitz & Clark, 2017). Due to the difficulties with identifying the effective parameter values in a heterogenic and complex system, the degree of realism in the conceptual models is limited (Gharari, et al., 2014). Many studies have been conducted to increase the physical realism of the parameters, for example, by incorporating different data sources (Freer, et al., 2004) or extracting more information from data, like hydrological signatures (Euser, et al., 2013)

During the calibration process, another problem arises: the issue of equifinality, wherein multiple sets of parameters can give equally accurate reproduced hydrological responses (Beven, 1996). During calibration the goodness-of-fit is based on the comparison of measured stream discharge and modeled discharge. However, a strong interaction between parameters is present, and by evaluating the model, the efficiency of the complete set is tested. The evaluation of the individual parameters is difficult (Willems, 2014). Equifinality can indicate that the model has an insufficient representation of the underlying processes (Gharari, et al., 2014) or that models are simple too complex (Savenije, 2009; Willems, 2014). With increased complexity, over-fitting to the observed data is possible during model calibration.

The use of constraints (regarding relations between parameter, fluxes and states) to the parameter sets prevents the model from overfitting in the calibration period. This results in a decrease of model uncertainty. Thus, the models have a higher predictive power and a higher skill to reproduce the overall system response (Gharari, et al., 2014; Hrachowitz, et al., 2014). However, these studies focused on relations between parameters and fluxes. Finding adequate parameter values from field observations, which represent the heterogeneity of a catchment on the spatial resolution scale of the model, is considered much harder (Gharari, et al., 2014; Hrachowitz & Clark, 2017).

1.4 Scope

As mentioned above, the model uncertainties are twofold: (1) structural and (2) parameter. For this thesis, the parameter uncertainty is investigated. This study will look into the possibilities to find adequate parameter values or distributions from field observations to further limit the equifinality.

1.5 Research objective

Based on the problem statement and the scope, the objective of this research is to identify, determine and quantify the extent to which it is possible to estimate the parameter values or parameter distribution based on field observations and a given hydrological model structure. In a traditional calibration an uninformed parameter distribution was determined for each individual parameter. The goal of the research is to avoid the use of uninformed prior parameter distributions during calibration by using

available information from the field observations to generate informed prior distributions. The model functions as the hypothesis on the processes occurring in the catchment (Figure 1).

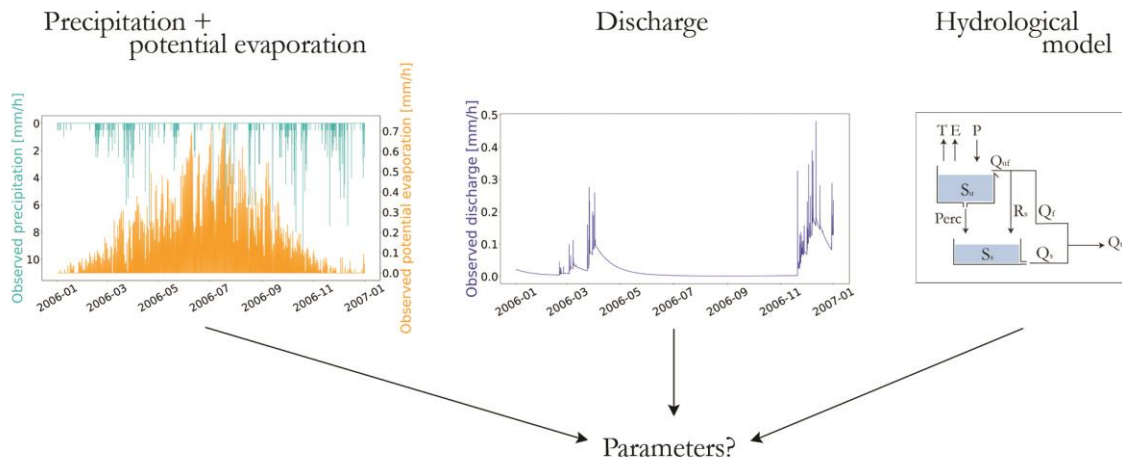


Figure 1. A visual representation of the objective of this research. The objective is to estimate the parameter values or parameter distribution based on field observations and a given hydrological model structure.

1.6 Research questions

Based on the research objective formulated above, the following main research question was developed to be answered in this study:

To what degree is it possible to avoid the use of uninformed prior parameter distribution in the calibration of a conceptual hydrological model by using available information from the field observations to generate informed prior distributions?

Sub-questions were formulated to provide additional information to answer the main research question. These sub-questions indicate the important issues to be addressed.

1. Can we select sub-periods from a hydrograph typical of different hydrological process, which can be coupled to model components, and how can this help to determine model parameters?

The hydrograph is an integrated representation of the system and is built from multiple interacting and spatially variable hydrological processes. The measured hydrograph was separated into periods that cover different response modes, like dry periods, draining periods and wet periods. The parameter identification was based on sub-periods or different event types.

For this question the key moments, or timing, of the different processes were analyzed to split the hydrograph into periods that cover different response modes. The parameters connected to these sub-periods were estimated.

The hypothesis is that each period has different “dominant” processes with associated parameters. Using the sub-periods of the hydrograph to find individual parameters of the model could further increase realism. In theory, the uncertainty and equifinality in the parameter sets should be reduced if the parameter information is directly extracted from the observed data. Validation of the model parameters against observations remains problematic; therefore, a synthetic experiment was conducted to validate the methods to obtain parameter distributions.

2. To what extent are the methods to obtain the parameter distributions sensitive to data errors?

In the synthetic experiment, all data and model uncertainty was normally excluded. However, for this question, the effect of data uncertainties on the parameter determination methods was investigated by corrupting the synthetic data set.

3. To what extent can the methods be used for real rainfall-runoff data?

For this last question, measured data was used to find a parameter set. In this case, real data and model uncertainty were present. The performance of the methods to determine parameter distributions could no longer be assessed by looking at parameters to produce the data. However, a comparison of the methods could be made with a Monte-Carlo sampling strategy with uninformed parameter distributions.

1.7 Thesis outline

A research outline was developed to answer the research questions and is presented in Figure 2. The research outline consists of a theory section, modelling section and conclusions. Methods developed in the theory section were applied in the modelling section, which is divided into three sub-parts according to the sub-questions.

An introduction of study area is given in Chapter 2. In addition, the investigated models are presented to assess the methods used to define parameter distributions; the extensive model descriptions can be found in Appendix 1. Furthermore, an explanation of the data sets used in this study is provided. The distribution function for daily evaporation to hourly resolution is given in Appendix 2.

The third chapter has two sections. First the theory portion of the thesis outline. This consists of a method explanation for each parameter to determine the parameter distribution under consideration. Second is the description of the method used to obtain the desired results for the application of the methods. This section consists of three different components, according to the research outline: the synthetic experiment, the sensitivity analysis and the real-world application.

The most important summarized results of the methods to determine parameter distributions and the different research sections are presented in the fourth chapter. The complete results are given in Appendix 4, Appendix 5 and Appendix 6. The methods and results are discussed together with their limitations or weaknesses in Chapter 5. In the last chapter, the answers to the research questions are presented.

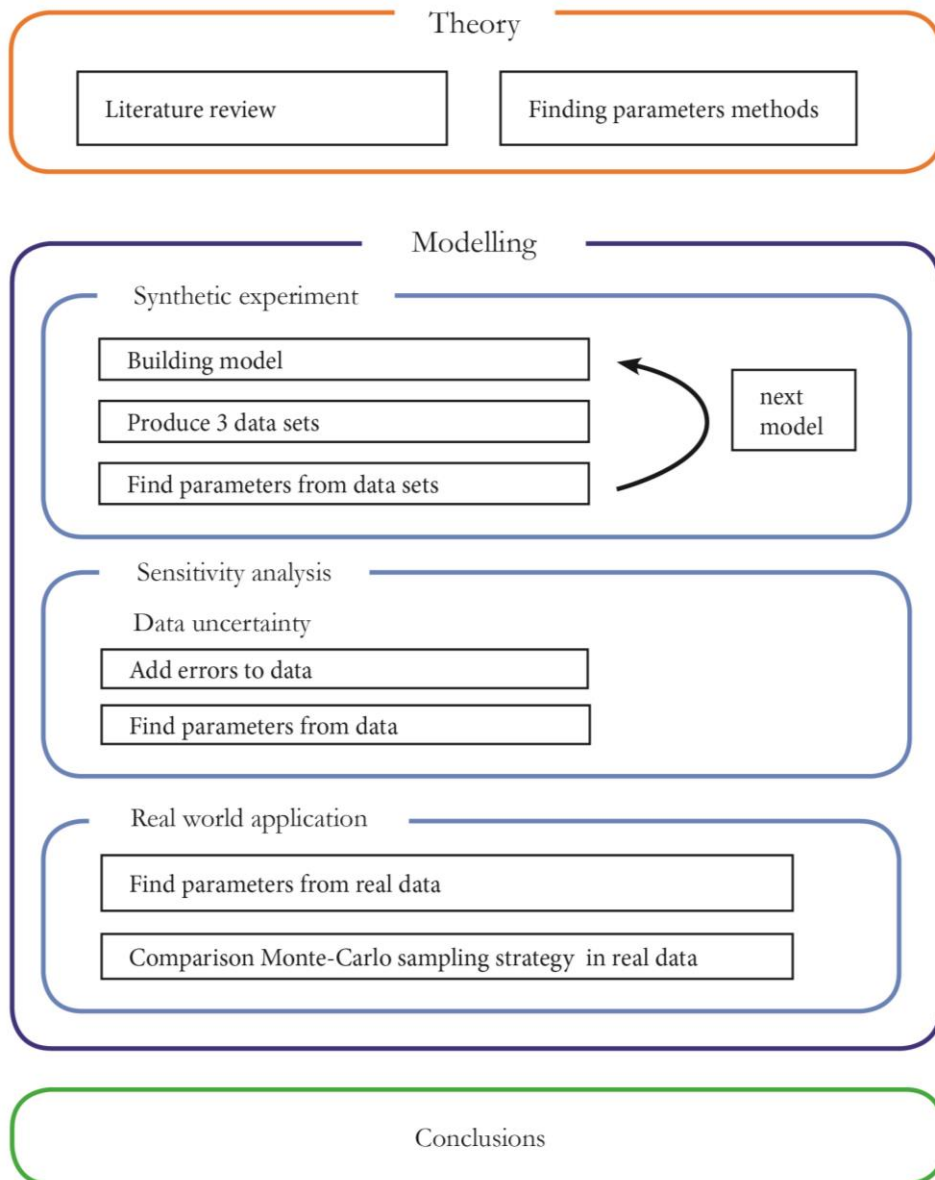


Figure 2. Representation of research outline: theory portion where methods are developed to find parameter distributions, modelling portion to test the developed methods, which is divided into three sub-parts: (1) synthetic experiment, (2) sensitivity analysis and (3) real-world application. Conclusions are drawn in the last section.

Chapter 2 Data and Models

2.1 Study area

2.1.1 Kervidy-Naizin catchment, France

In the main portion of the research uses data originated from the Kervidy-Naizin catchment in France (Figure 3). The catchment is located in the center of French Brittany, approximately 100 km west of Rennes. The surface area of the catchment is 4.9 km².

The catchment observatory is part of the French network of Drainage Basins (Réseau des Bassins Versants, RBV). The data originates from the years 2005 through 2011 and can be downloaded from the ORE-AgrHyS website (https://www6.inra.fr/ore_agrhys).

The catchment area has a humid climate with an average annual precipitation of 900 mm. The monthly average temperature varies between 5.4°C in January to 17.4°C in August. The potential evapotranspiration has an annual average of 700 mm and is considered stable over the years.

The measurements were conducted at a weather station at Kervidy. This station is located approximately 1 km from the catchment outlet and records hourly rainfall and all required meteorological data needed to calculate the potential Penman evaporation: air and soil temperatures, air humidity, incoming net radiation, wind direction and speed.

With the Penman equation, it is possible to calculate the daily potential evaporation. As the available precipitation data has an hourly resolution, the potential evaporation is distributed to hourly potential evaporation. A short description of the method to distribute daily potential evaporation data to hourly data, suggested by Fleming (1970), is presented in Appendix 2.

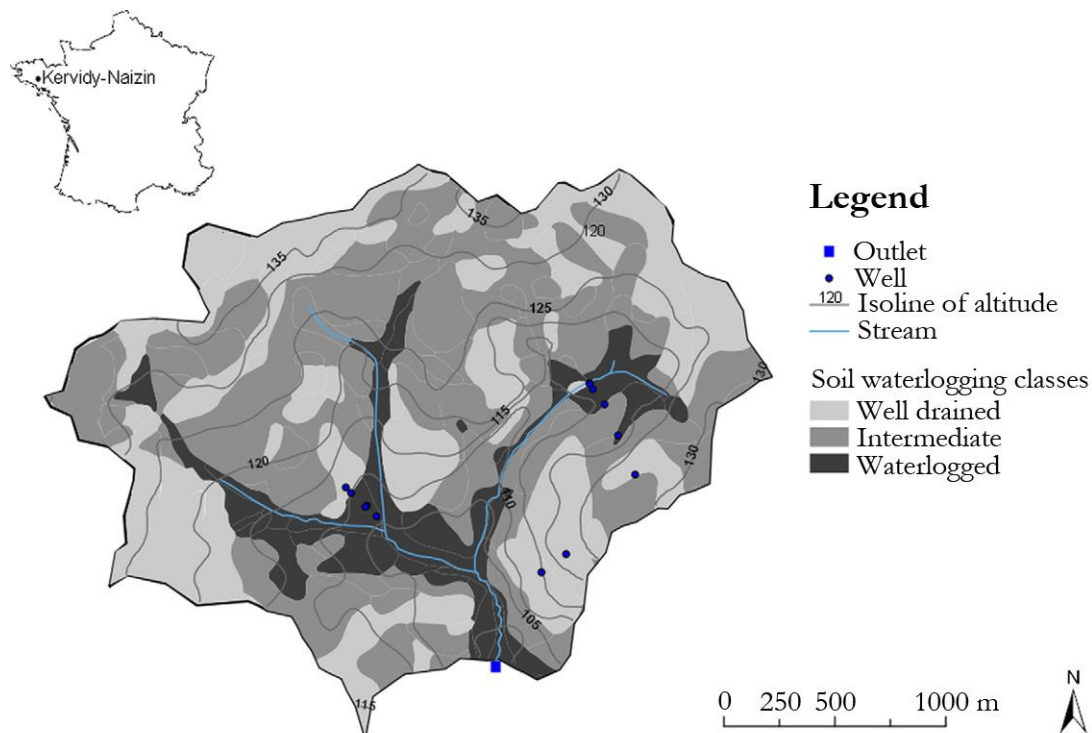


Figure 3. Kervidy-Naizin catchment in France, which is part of the ORE-ArtgHys network (Aubert, et al., 2013).

2.1.2 Precipitation stations, the Netherlands

Precipitation errors were added in the sensitivity analysis by using two different time series. The stations used are both located in the Netherlands and owned by the KNMI, the royal Dutch meteorological institute. The first station is the Hupsel station (52° 04' N.B. 06° 39' O.L.), and the second station is the Twenthe station (52° 16' N.B. 06° 53' O.L.). The distance between the stations is approximately 40 km (Figure 4).



Figure 4. Measuring stations Twenthe and Hupsel, part of the KNMI network (www.meteobase.nl)

2.2 Model description

In this study is worked with lumped conceptual model structures. The model configuration consists of structures with an increasing complexity to identify the extent to which it is possible to extract parameter values or distributions of individual processes using the hydrograph. By increasing the conceptualized processes, the heterogeneity of a catchment was better captured in each step. Each model consists of several components that represent one or more different hydrological processes. The models are based on the model structures described in Fenicia, et al. (2006), Kavetski and Fenicia, (2011) and Euser et al. (2013).

The model schematizations of each individual model are depicted in Figure 5. There are two main lines in the model structures: preferential flow paths (left) and percolation (right). The last model (model 9) is a combination of models 4 and 7. All models take the observed precipitation (P) and potential evaporation (E_p) as input. The output of the models is the simulated total discharge (Q_t) and the simulated actual evaporation and transpiration (E, T). The models consist of, at most, two reservoirs: the unsaturated zone reservoir and the slow runoff reservoir. The reservoir storage is referred as S_x in which the index x equals U (for unsaturated zone reservoir) or S (slow runoff reservoir).

The water balance equations for each reservoir in the models are presented in Table 1. The fluxes entering and leaving the reservoirs are given in these water balance equations. The constitutive formulas of the models are provided in Table 2 and represent the mathematical approximation of the fluxes.

Due to time constraints during the research, the choice is made to exclude models 5 and 8. These models include also a fast reservoir. As a consequence of excluding models 5 and 8, this reservoir is not included in model 9.

A complete description of each model is given in Appendix 1.

<i>Water balance equations</i>	<i>M 1</i>	<i>M 2</i>	<i>M 3</i>	<i>M 4</i>	<i>M 6</i>	<i>M 7</i>	<i>M 9</i>
$\frac{dS_u}{dt} = P(t) - E(t) - Q_s(t)$	x						
$\frac{dS_u}{dt} = P(t) - E(t) - Q_s(t) - Q_f(t)$		x					
$\frac{dS_u}{dt} = P(t) - E(t) - Q_{uf}(t)$			x	x			
$\frac{dS_u}{dt} = P(t) - E(t) - Perc(t)$					x		
$\frac{dS_u}{dt} = P(t) - E(t) - Perc(t) - Q_f(t)$						x	
$\frac{dS_u}{dt} = P(t) - E(t) - Perc(t) - Q_{uf}(t)$							x
$\frac{dS_s}{dt} = R_s(t) - Q_s(t)$				x			
$\frac{dS_s}{dt} = Perc(t) - Q_s(t)$					x	x	
$\frac{dS_s}{dt} = R_s(t) + Perc(t) - Q_s(t)$							x

Table 1. Water balance equations of the models used in this study; the x indicates presence in the model structure

<i>Constitutive formulas</i>	<i>M 1</i>	<i>M 2</i>	<i>M 3</i>	<i>M 4</i>	<i>M 6</i>	<i>M 7</i>	<i>M 9</i>
$E+T = \min(S_u, E_p)$	x				x		
$E+T = \frac{S_u}{S_{U,max}} E_p$		x	x	x		x	x
$Perc = P_{max}$					x		
$Perc = P_{max} \frac{S_u}{S_{U,max}}$						x	x
$Q_f = \max(0, S_u - S_{U,max})$		x				x	
$Q_{uf} = \max(0, S_u - S_{U,max})$			x	x			x
$R_s = D * Q_{uf}$				x			x
$Q_f = (1 - D) * Q_{uf}$				x			x
$Q_s = \frac{S_s}{C_s}$	x	x	x	x	x	x	x

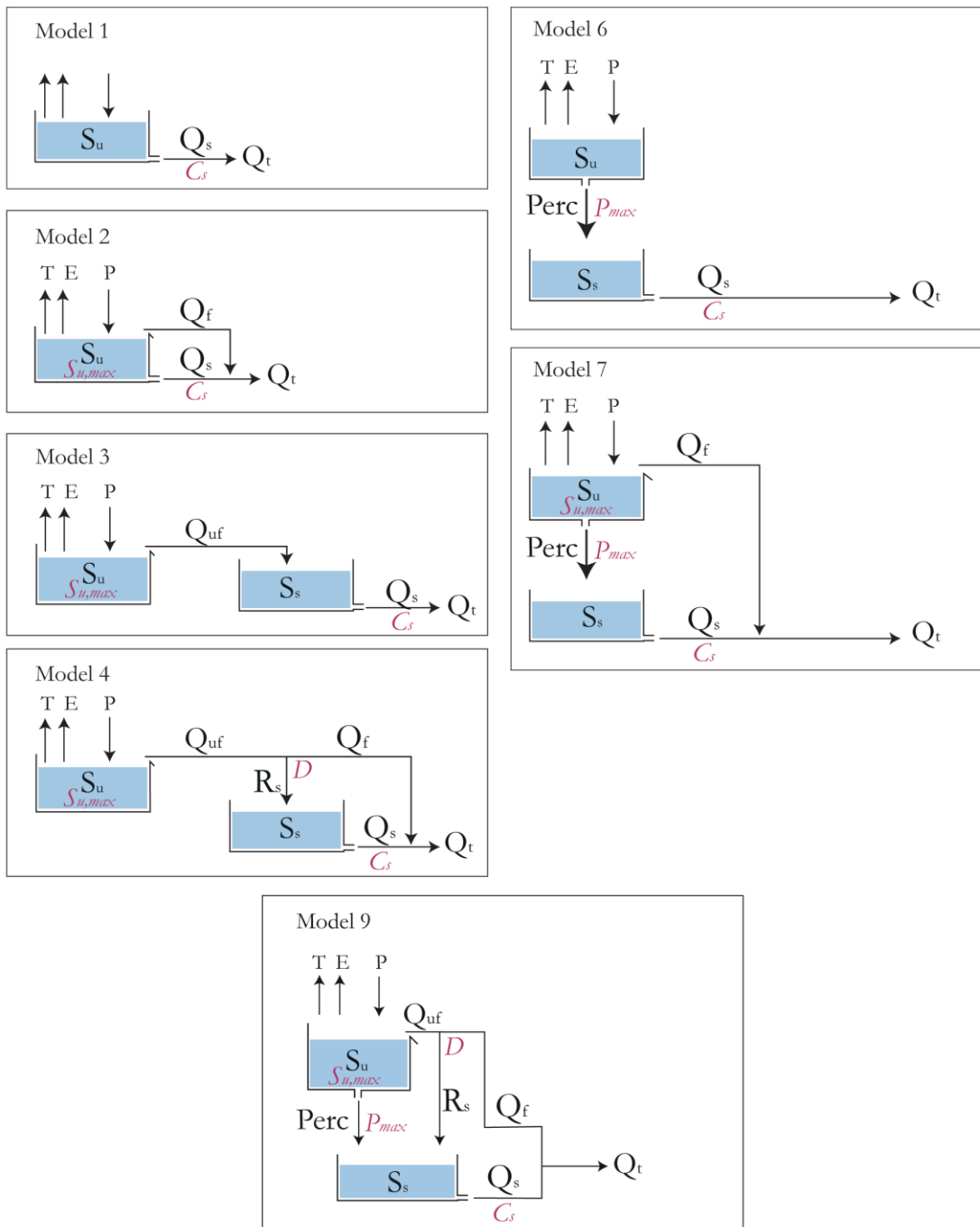
Table 2. Constitutive formulas in the model structures; the x indicates presence in the model structure

2.3 Synthetic data

Synthetic data sets are used for both the synthetic experiment and sensitivity analysis. The discharge data of these sets originates from model output which uses real precipitation and potential evaporation data as input, resulting in discharged data with no errors. The French forcing data is used for these experiments. Only the precipitation error analysis of paragraph 3.3.1 makes use of the Dutch forcing data.

To produce the discharge data parameter sets are chosen randomly. Each model produces three sets of data using three unique sets of parameters. The only exception are the parameter sets for model 9, which originate from a Monte-Carlo sampling strategy were set 1 and 2 have an equal performance regarding the observed discharge in France.

The parameter sets for the different models and produced discharge data are given in Appendix 3.



Fluxes	Parameters
Q_t = total discharge	C_s = Groundwater drainage parameter
Q_s = slow discharge	P_{max} = Maximum percolation rate
Q_f = fast discharge	$S_{u,max}$ = Maximum unsaturated zone storage
Q_{uf} = overflow unsaturated zone	D = Partitioning coefficient fast and slow reservoir
R_s = preferential recharge	
$Perc$ = percolation	
T = transpiration	
E = evaporation	
Storages	
S_u = unsaturated zone reservoir	
S_s = slow runoff reservoir	

Figure 5. A schematic representation of the hydrological model structures used in this study. The fluxes and reservoir names are shown in black. The related parameters are shown in red.

Chapter 3 Methods

The methods used in this research are explained in this chapter. The theory to determine parameter distributions for the different parameters in the models is explained in Section 3.1. A description of the approach to obtain insight into the behavior of the methods explained in the theory section is given in the second half of this chapter. The methods are applied in three tests, according to the research outline: the synthetic experiment (Section 3.2), sensitivity analysis (Section 3.3) and the real-world application (Section 3.4).

3.1 Methods to determine parameter distribution

3.1.1 Groundwater drainage parameter (C_s)

The groundwater drainage parameter (C_s) is a characteristic of the linear approximated reservoir that links the storage to the discharge. The physical meaning of this parameter could be explained as the mean response time of water in the saturated zone.

3.1.1.1 Fit method

To determine the C_s value of a reservoir, multiple methods have been developed, for example, the master recession curve analysis (Lamb & Beven, 1997). In this analysis, various individual recession curves of the hydrograph are used to construct the master recession curve (MRC). To select the recession curves, some conditions are set. For a minimum period (multiple hours or days), all other fluxes, like precipitation and evaporation, should be zero (for example, during the night). During this period, the storage change of the reservoir is driven by the discharge (1).

$$\frac{dS_s}{dt} = -Q_s(t) \quad (1)$$

The master curve recession analysis has its drawbacks. With short data series, not enough recession periods can be selected to construct one synthetic curve, resulting in a large error in the estimation of the C_s value. This research suggests another approach in which individual recessions are used to determine the C_s value. This approach provides insight into the distribution of the individual C_s values.

A recession must meet the following requirements:

1. The recession has a minimum length (automatically set to 10 h).¹
2. During a recession period, no rainfall is measured and, for model 1, also no potential evaporation is measured during the recession.
3. The sum of the discharge is larger than 0.1e-10 to prevent errors.²
4. To exclude the effects of other fluxes with delays in the catchments, for multiple days (automatically set to a minimum of two days) prior to the selected recession, no precipitation is measured.³
5. The end of the recession is just before the start of a rainfall event.

A part of a discharge series with a selected recession is illustrated in Figure 6. A complete discharge series with all selected recessions is depicted in Figure 7.

¹ This is the maximum length possible during night time; shorter recession lengths have a large impact on the final C_s value.

² With lower values, the method provides unreliable results, even though the model can produce lower discharges. In addition, measurement techniques cannot be measured with this precision.

³ Shorter dry periods appeared to have a significant impact on the final C_s value.

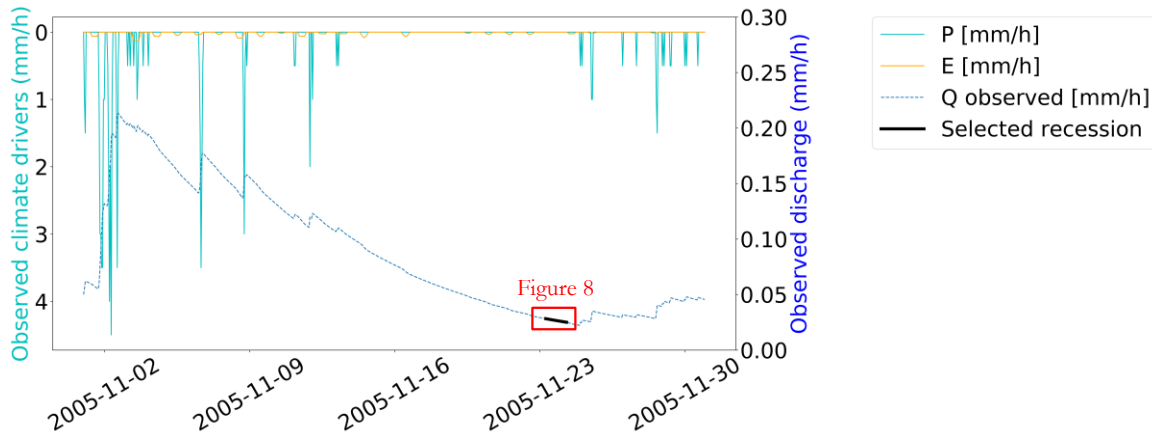


Figure 6. Observed time series of hourly precipitation (light blue), hourly potential evaporation (orange) and discharge (blue dotted line). Data originates from model 1 parameter set 3 in the years 2005-2011. Example of the selected recession (black) was used in the fit method. The recession has a length of 10 h, no precipitation was measured two days prior nor during the recession, no potential evaporation was measured during the recession and the end of the recession was just before the start of a precipitation event.

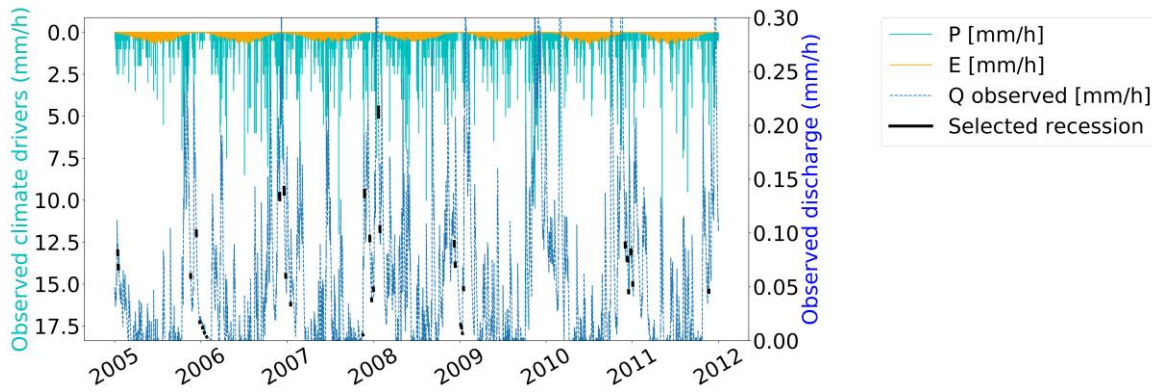


Figure 7. Observed time series of hourly precipitation (light blue), hourly potential evaporation (orange) and discharge (blue dotted line). Overview of all selected recessions (black) were used in the fit method. Data originates from model 1 parameter set 3 in the years 2005-2011.

The C_r is determined in two steps. First, a fit is made through each recession using an analytical equation (2). To assess the performance of the fit, the Nash-Sutcliffe (NS) objective is calculated⁴. All recessions with an $NS < 0.8$ are discarded. Recessions with a lower performance are considered influenced by other fluxes or measurement errors.

$$Q = Q_o e^{\frac{t}{C_r}} \quad (2)$$

An example of a recession and a fit is provided in Figure 8. The original C_r value is 200 h and the C_r of the fit is 199.5 h. The small error of 0.5 h can be explained. The recession is produced in the models with the empirical relation (4), while the fit is made using an analytical solution (equation (2)).

⁴ The Nash-Sutcliffe objective function is given in Table 5.

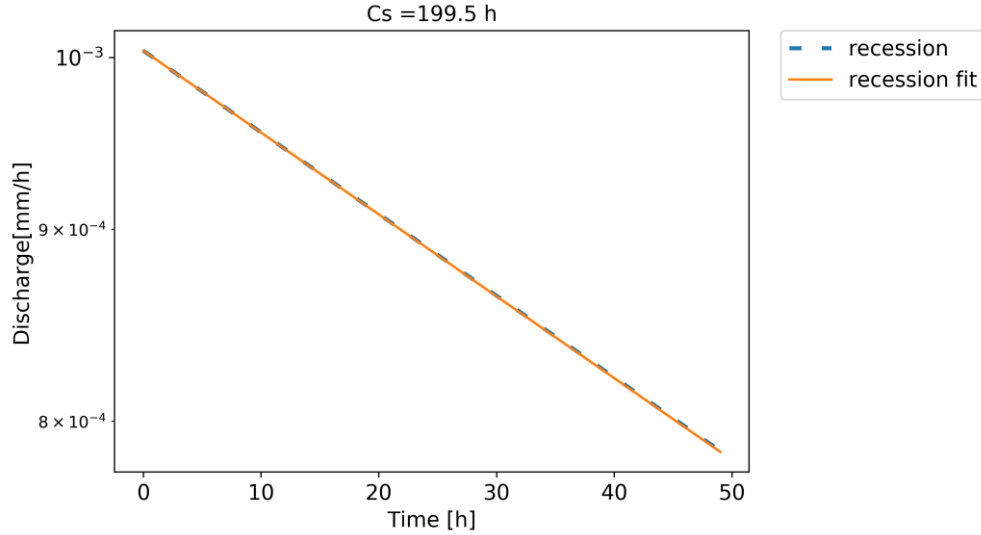


Figure 8. A selected recession from the data of model 4 parameter set 1 (blue dotted line) with an original C_s value of 200 h and a fit (orange line) with an estimated C_s value in the fit method of 199.5 h.

The next step in the method is to determine the exact C_s value. The relations below are used to determine the empirical relation of the recession. Here $S_s(t=0)$ originated from equation 3 with $Q_i = Q_{observed\ at\ t=0}$. A solve package⁵ in the Python programming language is used to solve the problem. The start value for solver is the C_s value originating from step 1 (the analytical fit).

The C_s distribution of the catchment is given by all calculated C_s values with multiple no rain periods. Multiple no rain periods are used to limit the errors in the distribution.

$$S_s(Q) \quad S_{s,i} = Q_i C_s \quad (3)$$

$$Q(S) \quad Q_i = \frac{S_{s,i}}{C_s} \quad (4)$$

$$S(S,Q) \quad S_{s,i+1} = S_{s,i} - Q_i dt \quad (5)$$

$$S_s(S) \quad S_{s,i+1} = S_{s,i} \left(1 - \frac{dt}{C_s}\right) \quad (6)$$

3.1.1.2 S-D method

Recharge of the slow reservoir has a significant influence on the recession curve, according to the paper by Fenicia et al. (2006). To minimize the influence of the recharge, the Storage-Discharge (S-D) method was developed based on this paper. Models 6, 7 and 9 indicate recharge over an extended period due to percolation. In an iterative process with an initial guess of the C_s value and model runs, the “correct” C_s can be reached for recessions with recharge. The following steps were developed based on Fenicia et al. (2006).

Step 1 : Initial Storage-Discharge relation (S-D relation)

The fit-method functions as an initial estimation of the C_s under the assumption that the slow reservoir does not receive recharge. It is possible to construct a S-D relation of that recession by integrating the recession over time (Figure 9). For the starting point of the S-D relation, the initial estimation of the C_s is needed. This starting point is determined using equation 3. The slope of the S-D relation is equal to the C_s .

⁵ The solve package used is `scipy.optimize.curve_fit`

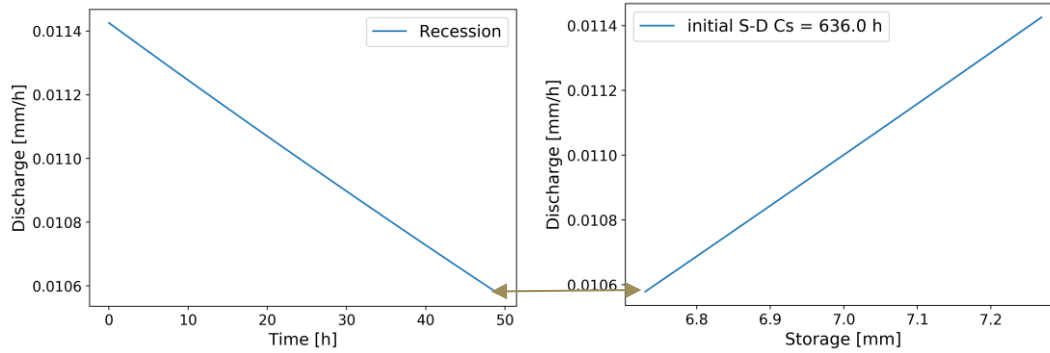


Figure 9. A selected recession from the data of model 9 parameter set 1 (left); the corresponding storage-discharge relation with estimated C_s value of 636 h (slope S-D relation) (right)

Step 2: Calculate recharge (P_{tot})

For the initial S-D relation is assumed that during recession periods, no recharge enters the slow reservoir. Recharge could, for example, be percolation (P_{peri}) or preferential recharge (R_p). To assess the impact of the total recharge (P_{tot}), the recharge first needs to be calculated for the corresponding recession segments. This P_{tot} is calculated by running the model using the estimate of the C_s value and the other parameters determined with the methods using the estimate of the C_s . In the examples of this section, the other parameters are assumed to be known.

Step 3: Recalculate the S-D relation

With the calculated recharge, it is possible to determine a new S-D relation for each recession (Figure 10). In contrast to the calculation of the S-D relation in step 1, where only the time integral is taken over the discharge, now the time integral is taken over the difference between the discharge (outflow of the slow reservoir) and the P_{tot} (inflow to the slow reservoir).

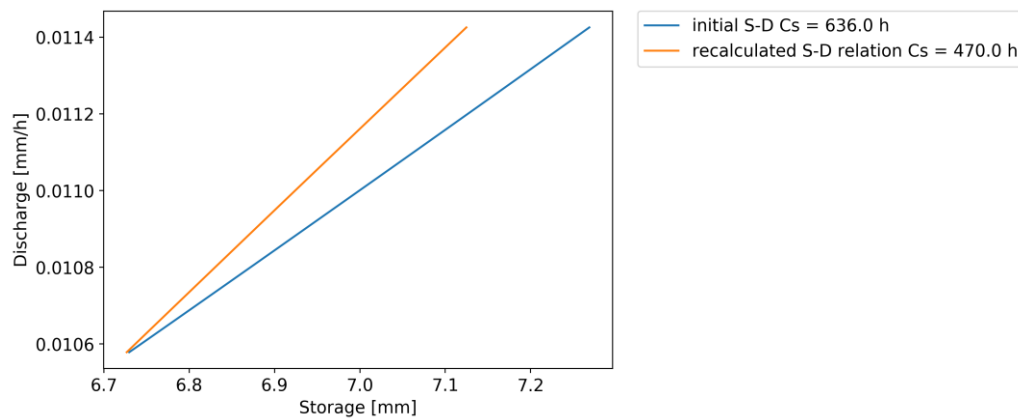


Figure 10. Original storage-discharge relation, time integral over the recession (blue line) and the new recalculated (orange line) storage-discharge relation, time integral over the difference between the recession (outflow of the slow reservoir) and the P_{tot} (inflow to the slow reservoir). Data originates from model 9 parameter set 1.

As Figure 10 illustrates, the new S-D relation is much steeper. The flattening of the old S-D relation is caused by the constant increase in storage due to the recharge of the slow reservoir. In contrast to what the individual recessions show, the slow reservoir needs to empty faster than initially thought to compensate for the recharge.

Small errors are made in this method since the initial storage calculated with equation 3 assumed that the S-D relation is without recharge. The method does not take into account the P_{tot} at the end of a recession. To minimize this error, only recessions with a tail end lower than a pre-set maximum discharge value were used in the analysis. The pre-set maximum discharge was chosen by visual inspection of the data. The pre-set maximum discharge is equal to the maximum discharge during summer.

3.1.2 Partitioning coefficient fast and slow reservoir (D)

Excess water from the unsaturated zone, the proportion of water that cannot be stored, is routed toward the saturated zone (S_s) and the fast lateral runoff processes (Q_f). The partitioning coefficient (D) regulates the distribution between the preferential groundwater recharge (R_s) toward the S_s and Q_f (e.g., preferential flow or saturated overland flow).

To determine D , the total volume of the overflow (O) needs to be calculated, as well as one of the two flows (R_s or Q_f). First, the overflow moments need to be found in the time series. An overflow moment is defined as the following:

1. In the next time step, the discharge is increased by 1.5 $Q[t_i]$. Where $Q[t_i]$ is the discharge of the current time step.
2. After an overflow moment, the discharge is decreased significantly, by at least 0.5 $Q[t_i]$.⁶
3. During wet periods overflow is happening often, sometimes even each time step. The calculation of D will be impacted if it is tried to calculate during this period with multiple peak flows. An extra criteria is introduced: the discharge of both point 1 and 2 needs to be below a predefined maximum discharge ($max Q$). This value is the discharge which is exceeded for the most peak discharges and is defined by visual inspection.

With the information from the observed discharge it is possible to determine the total overflow volume, the Q_f and the R_s . The first step is to calculate the expected flow during the peak flow using the C_s (Q_{expect} (equation 7)) represented by the green dot in Figure 11. With the discharge measured after the peak flow event, it is possible to calculate the expected flow from the saturated zone the purple dot (Q_{Ss} (equation 8)). From here, it is possible to calculate the volume of the overflow (equations 9, 11 and 12). The last step is to calculate D using (13).

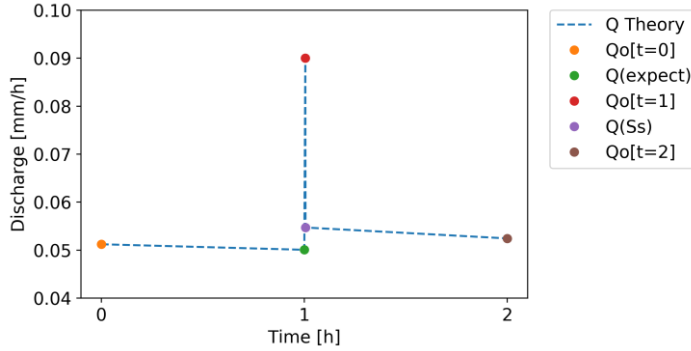


Figure 11. Selected overflow moment observed discharge (orange, red and brown dots); expected flow during the peak flow calculated with the C_s (green dot) and expected flow from the saturated zone calculated with the C_s (purple dot). The difference between the red and purple dots represents the volume of fast lateral runoff processes, while the difference between the purple and green dots represents the extra discharge from the slow reservoir due to additional water from the preferential recharge.

$$Q \text{ (expect)} \quad Q_{expect} = S_{s,t+1}C_s = (S_{s,t+0} - Q_{t+0})C_s \quad (\text{uses (3) and (5)}) \quad (7)$$

$$Q \text{ (Ss)} \quad Q_{Ss} = \frac{S_{s,t+1}dt}{C_s} = \frac{Q_o[t+2]C_s}{(1-\frac{dt}{C_s})C_s} \quad (\text{uses (3) and (6)}) \quad (8)$$

$$Q_f \quad Q_f = Q_o[t+1] - Q(S_s) \quad (9)$$

$$Q_{Rs} \quad Q_{Rs} = Q_o[t+1] - Q_f - Q_{expect} \quad (10)$$

$$R_s \quad R_s = Q_{Rs} * C_s \quad (11)$$

$$O \text{ (D)} \quad O = Q_f + R_s = (1 - D)O + D * O \quad (12)$$

$$D(Q_f, O) \quad D = 1 - \frac{Q_f}{O} \quad (13)$$

⁶ The factors of 1.5 and 0.5 were set by expert judgment; variations in these values did not give significantly different results.

3.1.3 Maximum unsaturated zone storage ($S_{U,max}$)

The maximum unsaturated zone storage ($S_{U,max}$) is a characteristic of the unsaturated zone reservoir. The $S_{U,max}$ reflects the maximum soil moisture capacity in the root zone of the catchment. When the maximum is exceeded, the fast lateral flow is activated. This parameter also indirectly controls the evaporation and percolation rates.

The method to determine the $S_{U,max}$ consists of three steps. A data preparation step is followed by the determination of the bounds for the $S_{U,max}$. The last step calculates the $S_{U,max}$ distribution with the inter-peak method.

Step 1: Data preparation

The models assume that during rainfall events, no evaporation flux is present. However, the daily potential evaporation is continuously distributed over the sun hours. The data set is prepared to take this error into account. In the data series, for the time steps in which precipitation is measured, the potential evaporation is set to zero.

Step 2: Defining the bounds for $S_{U,max}$

The water balance characteristics of the catchment are used to define the boundary values of the $S_{U,max}$. The method used in this thesis is based on the paper of McMillan et al. (2011). This method examines the response to the catchment's rainfall events by dividing the time series into individual events. The volume of rainfall causing a rapid increase in discharge (e.g. the threshold behavior of a catchment) could be coupled with the maximum storage of the unsaturated zone.

A storm event is identified using the rainfall time series; the values of rainfall intensities were taken from the paper by McMillan et al. (2011) and are defined by expert judgment:

1. The main portion of the event has a predefined length: standard set to 20 days.⁷ This portion should begin with a rainfall intensity greater than 0.5 mm/day for a given hour. The mean rainfall over this part of the event should be larger than 5 mm/day, or events longer than 60 days should exceed a total volume of 300 mm.
2. The rainfall tail is considered to discover the development of the discharge after the main volume of rain has fallen. The end of a storm event is defined when the rainfall intensity in the following 36 h after the main event length is less than 0.5 mm/day for a given hour.

The water balance is calculated over a particular part of the rainfall event. Here, the discharge time series is leading. With an observed significant increase in discharge after the start of the rainfall event, the threshold of the catchment is reached. A significant increase in discharge is defined as an increase in discharge larger than 105% of the current discharge⁸. The red dots in Figures 13, 14, and 15 indicate a significant increase in discharge.

The water balance is taken over the time series from the start of the event until the time step in which the discharge is increased:

1. The minimum threshold is defined as the sum of the rainfall minus the losses. The losses are not constant: for models 1 through 4, only the evaporation is defined as a loss; for models 6 through 9, percolation is also considered a loss.
2. The maximum threshold is defined as the sum of the rainfall minus 30% of the maximum percolation in the water balance period. The 30% percolation is subtracted to prevent an over estimation of the maximum. The maximum is only set if the rainfall exceeds the potential evaporation and the minimum threshold of that event is above zero.

⁷ Depending on the catchment properties, this period can be extended or shortened. In catchments with expected large storage capacities, longer periods are needed. Testing for multiple durations is recommended.

⁸ The value of 105% is set to prevent interruptions in the water balance period due to small variations in discharge often caused by percolation. For different data sets, this value could vary.

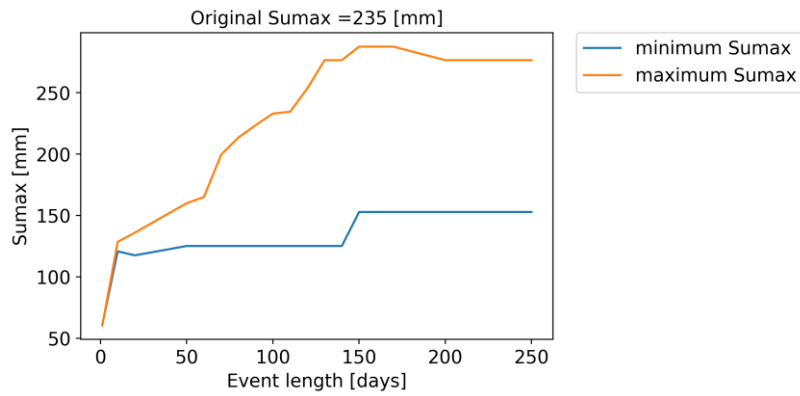


Figure 12. Result of the bound method using different event lengths for data originating from model 4; the estimated minimum value of $S_{U,max}$ = 160 mm, and the estimated maximum = 287.5 mm

For different event lengths, the largest threshold can be calculated for both the minimum and the maximum thresholds (Figure 12). The storm (time series used for the water balance) used to determine the minimum threshold (Figure 13) may be different from the storm used to calculate the maximum threshold (Figure 14).

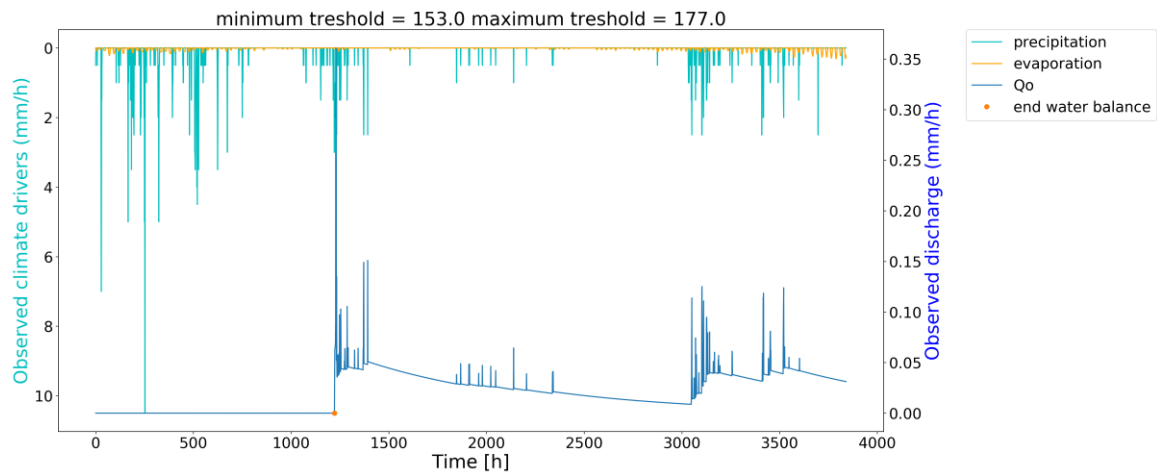


Figure 13. Selected individual event for the time series originating from model 4 set 2; $S_{U,max}$ = 235 mm; length of the event is 150 days

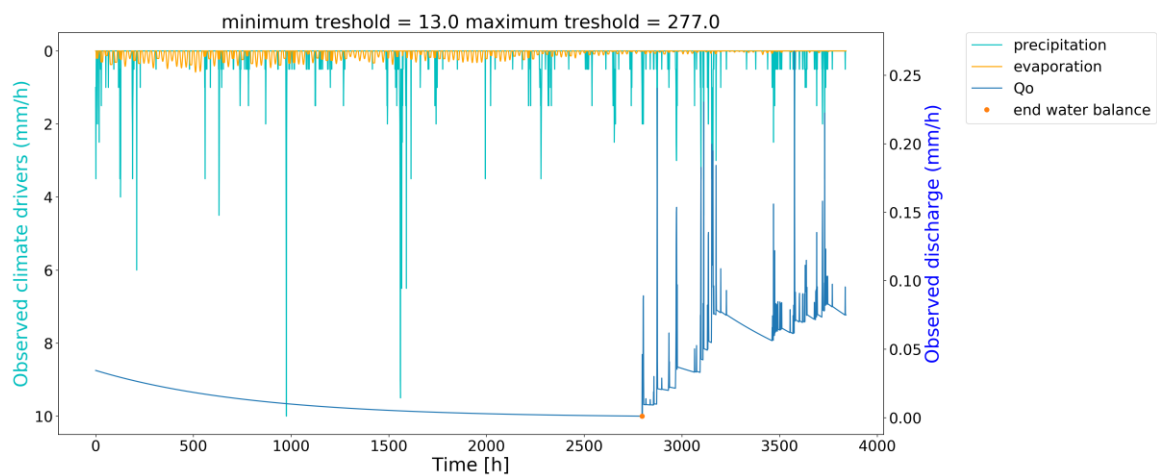


Figure 14. Selected individual event for the time series originating from model 4 set 2; $S_{U,max}$ = 235 mm; length of the event is 150 days

It is not necessary to check whether a long dry period was present before the start of an event. If there was rainfall prior to the event, the threshold is quickly reached, resulting in an early increase of discharge (Figure 15).

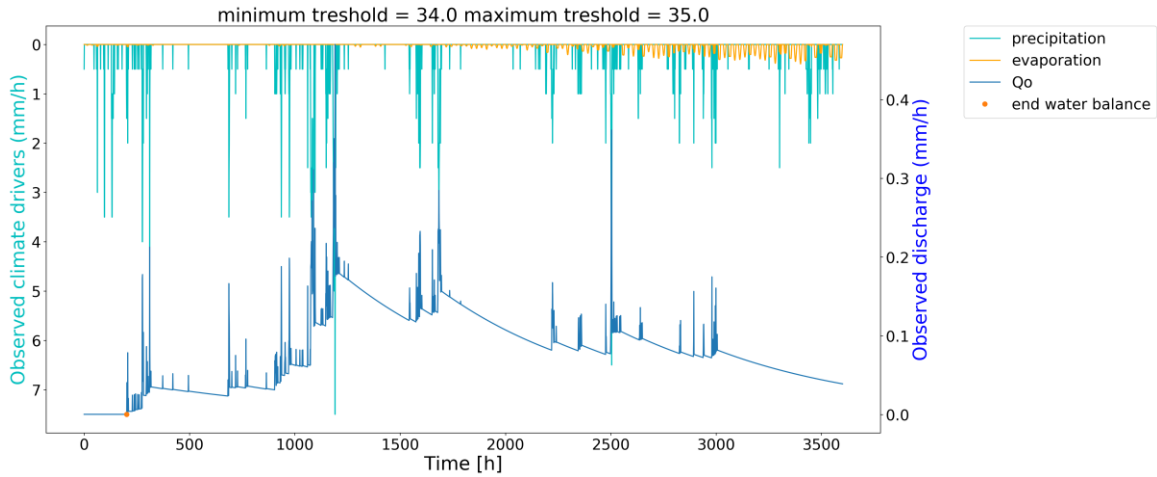


Figure 15. Previous precipitation of the selected event already partly filled the system, resulting in a low minimum and maximum $S_{U,max}$

Step 3: Determine $S_{U,max}$ (inter peak method)

The distribution of the $S_{U,max}$ value can be made using a system of equations. Here, the discharge, potential evaporation and precipitation data is used. The ground water drainage parameter (C_s) and maximum percolation parameter (P_{max}) are also needed. For the more complex models, the overflow data calculated for the D is required as well.

The basic assumption is that at overflow moments, the unsaturated zone (Su) is filled to the maximum. Then, in each time step, water is evaporated and percolated from the Su leading to a drop of the water level in the unsaturated zone. At precipitation moments, the Su is filled again. At the next overflow moment, the storage is filled to the maximum storage. Figure 16 presents an example hydrograph with climate drivers where at $t = 0$ overflow of the unsaturated reservoir occurs. At $t = 150$, overflow occurs again.

The parts of the hydrograph used to determine $S_{U,max}$ should have a minimum length of 4 time steps and a maximum length of 1000 time steps. Furthermore, in this period, a storage change should happen, so the sum of E_p should be larger than zero.

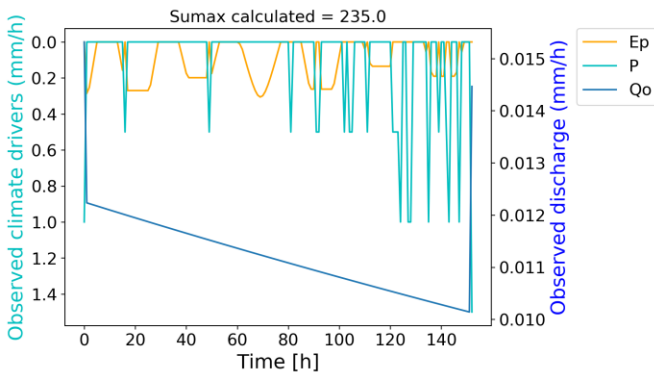


Figure 16. Part of the hydrograph starting and ending with an overflow moment with a storage change of the unsaturated zone in between due to evaporation (orange line) and precipitation (light blue line). This part is used in the inter peak method.

In the following example, the relations are provided. At $t = 0$, overflow occurs, and at $t = 4$, overflow reoccurs. In models 1 through 4, no percolation flux is present resulting in a P_{max} equal to zero.

$$Su [t + 0] = S_{umax} - P_{max} \quad (14)$$

$$Su [t + 1] = Su[t + 0] - E_p[t + 1] \frac{Su[t + 0]}{S_{umax}} - P_{max} \frac{Su[t + 0]}{S_{umax}} + P[t + 1] \quad (15)$$

$$Su [t + 2] = Su[t + 1] - E_p[t + 2] \frac{Su[t + 1]}{S_{umax}} - P_{max} \frac{Su[t + 1]}{S_{umax}} + P[t + 2] \quad (16)$$

$$Su [t + 3] = Su[t + 2] - E_p[t + 3] \frac{Su[t + 2]}{S_{umax}} - P_{max} \frac{Su[t + 2]}{S_{umax}} + P[t + 3] \quad (17)$$

$$Su [t + 4] = S_{umax} \quad (18)$$

Since there are no measurements of the S_u , an extra equation is needed to calculate the $S_{U,max}$. This is the water balance equation during overflow:

$$Overflow [t + 4] = P[t + 4] - (S_{umax} - Su[t + 3]) \quad (19)$$

Depending on the model, overflow can be calculated using different relations. For models 4 and 9, the overflow is calculated for the D . For models 3 and 7, the calculation of the overflow is similar to overflow calculations of D only now the Q_f is not present. The unknown value in the system of equations is $S_{U,max}$. Using a curve fit solver in Python, the $S_{U,max}$ is determined. The start value of the solver is the minimum $S_{U,max}$ from step 2.

3.1.4 Maximum percolation rate (P_{max})

The slow reservoir receives water from the unsaturated zone reservoir through percolation. The maximum percolation rate (P_{max}) regulates this process. The percolation rate is linearly related to the relative soil moisture content ($S_u/S_{U,max}$) and the P_{max} . The method to determine the P_{max} consists of two steps: a data preparation step and the determination of the P_{max} .

Data preparation

For the calculation of the percolation rate, the data should not include peak discharges. During these periods, many processes are acting, which makes it hard to determine the percolation rate. A procedure was developed to filter out peak discharges. The filter method is quite similar to the method used to determine peak flows for the partitioning coefficient (D).

If the discharge increases rapidly in the next time step, it is taken to be constant until it decreases again. A rapid increase is defined as an increase larger than 150% compared to the current discharge.⁹ The period, which is set to a constant value, is not taken into account by determining the percolation rates. Although percolation is occurring during this period, the calculation does not provide accurate results because too many fluxes are occurring.

The result of the filter is that only low flow periods are present in the data set. An example of the application of the filter is provided in Figure 17. The blue lines are the observed discharges, including the peak discharges, and the orange line is result of the filter, the discharge without peaks.

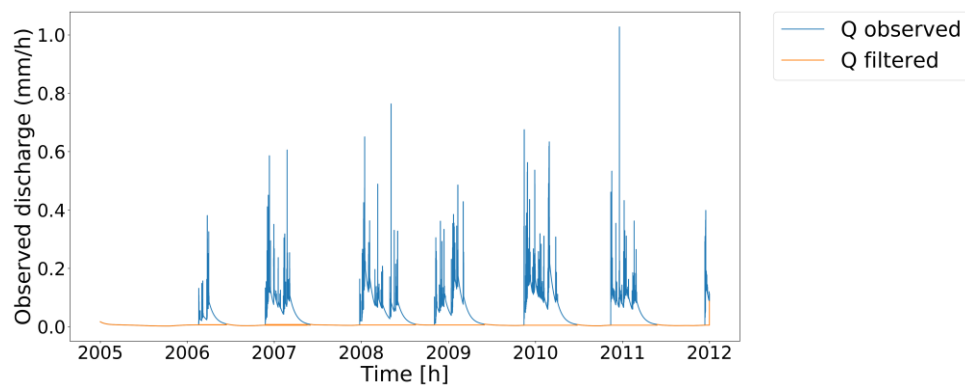


Figure 17. Observed discharge and filtered discharge of model 9 data set 2. In the filtered discharge all peak flows are excluded for the P_{max} method.

Determine P_{max}

To determine the maximum percolation rate, some data is needed: the filtered observed discharge and the ground water drainage parameter (C_d). For each time step, it is possible to calculate the percolation rate. The result of all the calculations provides a range of percolation rates in the catchment. The parameter distribution is given by the 20% highest calculated percolation rates. The most optimal maximum percolation rate (P_{max}) is set to the 75th percentile of the distribution. This will result in an underestimation of the P_{max} . The advantage, however, is that the outliers or data errors will not influence the result.

⁹ The factor of 1.5 is set by expert judgment; variations around these values did not provide significantly different results.

To calculate the percolation rate per time step, a few calculations need to be made. The steps are visualized in a graph presented in Figure 18.

1. The starting point of the method is at the last time step ($Q_o[t+1]$) in the time series. The C_s value is used to calculate the expected discharge (Q_{expect}) of the previous time step ($Q_o[t+0]$) with equations 20 and 23.
2. The expected discharge is then compared to the observed discharge in the previous time step ($Q_o[t+0]$). If percolation is present, the expected discharge is larger than 100.2 % of the observed discharge.
3. The expected discharge can be calculated as if there were no percolation (Q_{Ss}). This is calculated with equations 20, 22 and 21.
4. The difference in discharge of Q_{Ss} and $Q_o[t+1]$ represents the added discharge (Q_{perc}) due to percolation. Equation 24 is used to calculate the percolation rate.

$$S_{s,i} = Q_{o,i} C_s dt - Q_{o,i} \quad (20)$$

$$Q_i = \frac{S_{s,i}}{C_s} \quad (21)$$

$$S_{s,i+1} = S_{s,i} - Q_i dt \quad (22)$$

$$S_{s,i} = \frac{S_{s,i+1}}{1 - \frac{dt}{C_s}} \quad (23)$$

$$Perc = Q_{perc} C_s \quad (24)$$

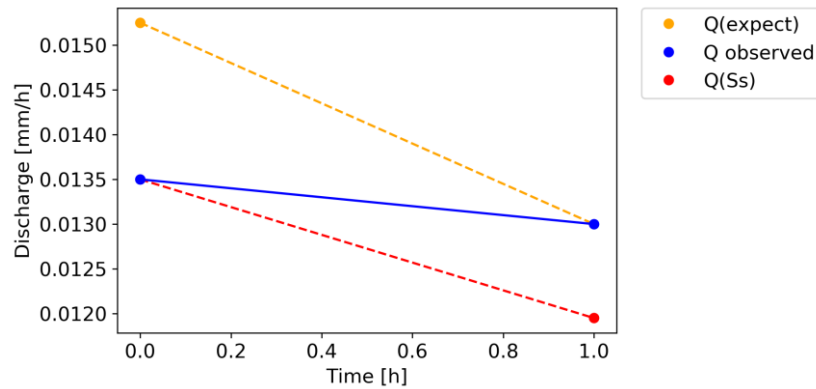


Figure 18. The blue line represents the observed discharge. The orange and red dotted lines represent expected discharges if no percolation would be present. The difference between the red and blue dots represents the extra discharge from the slow reservoir due to the additional percolated water in the slow reservoir.

The time steps that contain the analyzed percolation are displayed in Figure 19. The time steps containing the 20% highest calculated percolation rates are presented in Figure 20.

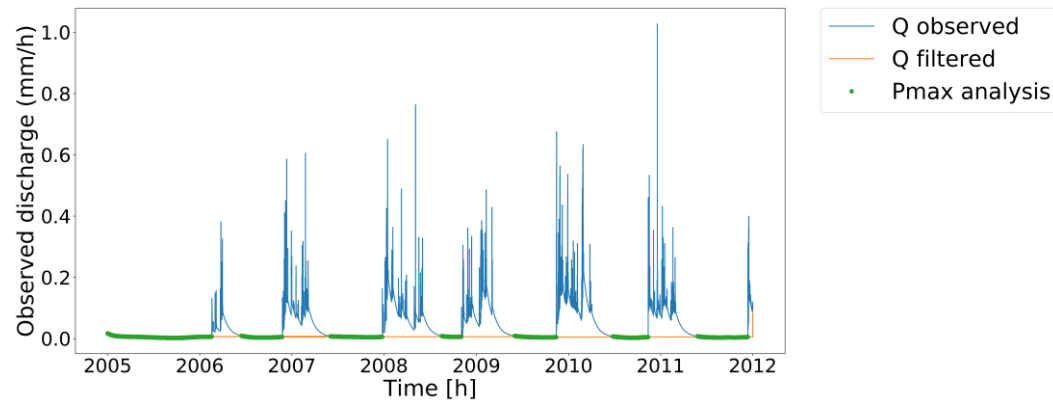


Figure 19. The green dots represent all the time steps used in the P_{max} analysis

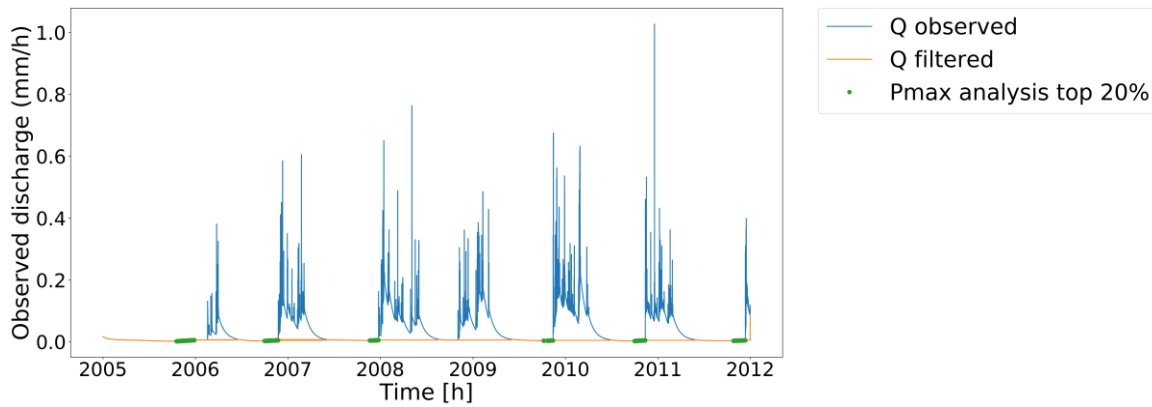


Figure 20. The green dots represent the top 20% highest calculated percolation rates

3.2 Synthetic experiment

In this synthetic experiment, discharge data was produced by a model. The model was driven by real rainfall data and potential evaporation data of the French catchment. Random parameter sets were used for the fluxes in the models. The methods to obtain parameter distributions, explained in Section 3.1, were applied to the synthetic discharge data and the forcing data.

Having a synthetic experiment allows experimentation in a controlled environment. The methods to obtain the parameters can be tested and validated since the original parameters from the data are known. This provides a chance to see whether the parameters found from the hydrograph are close to the original parameters from the model. After all, the goal of the research is to discover a reliable method to obtain parameter distributions and not to determine the exact values of the data sets. In this approach, model structure and data uncertainty were excluded. The hydrograph, forcing data and model structures were considered “perfect” and did not contain unknown processes or errors.

The effect of the method errors was investigated in this synthetic experiment. For three data sets per model, the methods were applied. The models used in this experiment were 4, 7 and 9. The value and accompanying error of the first parameter estimation formed the input for the next model parameter estimation. Thus, all the methods were tested without prior knowledge of the parameter values. This experiment indicates the extent to which the methods’ errors propagate to other parameter estimations.

The interdependencies of the different model parameter estimation methods are displayed in Figure 21. The arrow represents which method takes which parameter estimation as input to determine the current parameter. The figure clearly illustrates that many methods depend on each other.

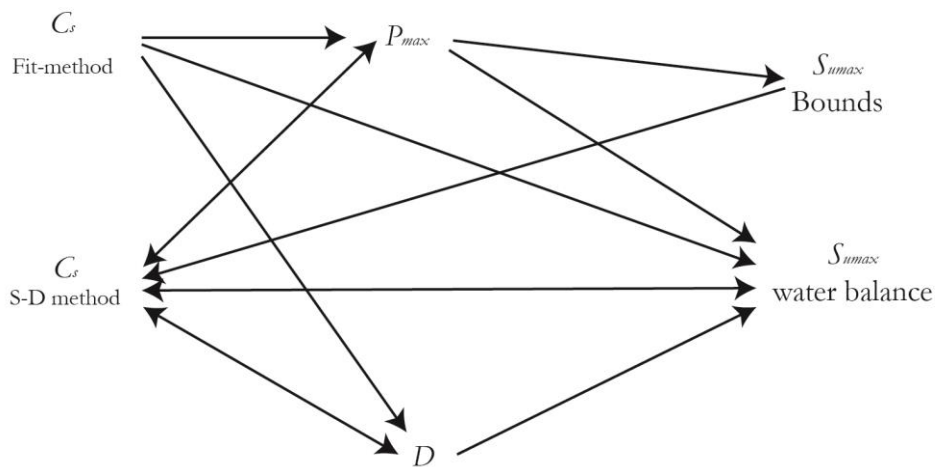


Figure 21. The interdependencies of the different expert-knowledge inverse modelling methods to find the parameter distributions

For models 7 and 9, it was assumed that the recharge to the slow reservoir plays a considerable role in the determination of the C_s value. The procedure to find the parameter distribution of the models for the different data sets consists of different stages. The whole procedure is schematically displayed in Figure 22, and the steps are outline below:

1. In the first step, an estimation of the C_s value is made with the fit method.
2. Subsequently, the estimated C_s value will be used to determine the P_{max} and for model 9 the D .
3. The estimated P_{max} is input for the bound method of $S_{U,max}$. The result of the method for $S_{U,max}$ is input for the next step.
4. The model simulations are completed using three combinations of parameter sets: one set with the minimum of the $S_{U,max}$, one set with the median of the $S_{U,max}$ and the last set with the maximum of the $S_{U,max}$. The C_s , P_{max} and D are constant for all the sets.
5. The P_{tot} of each recession is known from the model simulations, and the S-D method can be applied to update the C_s value.
6. The updated C_s value forms the input of the second step and is iterated until convergence is obtained.

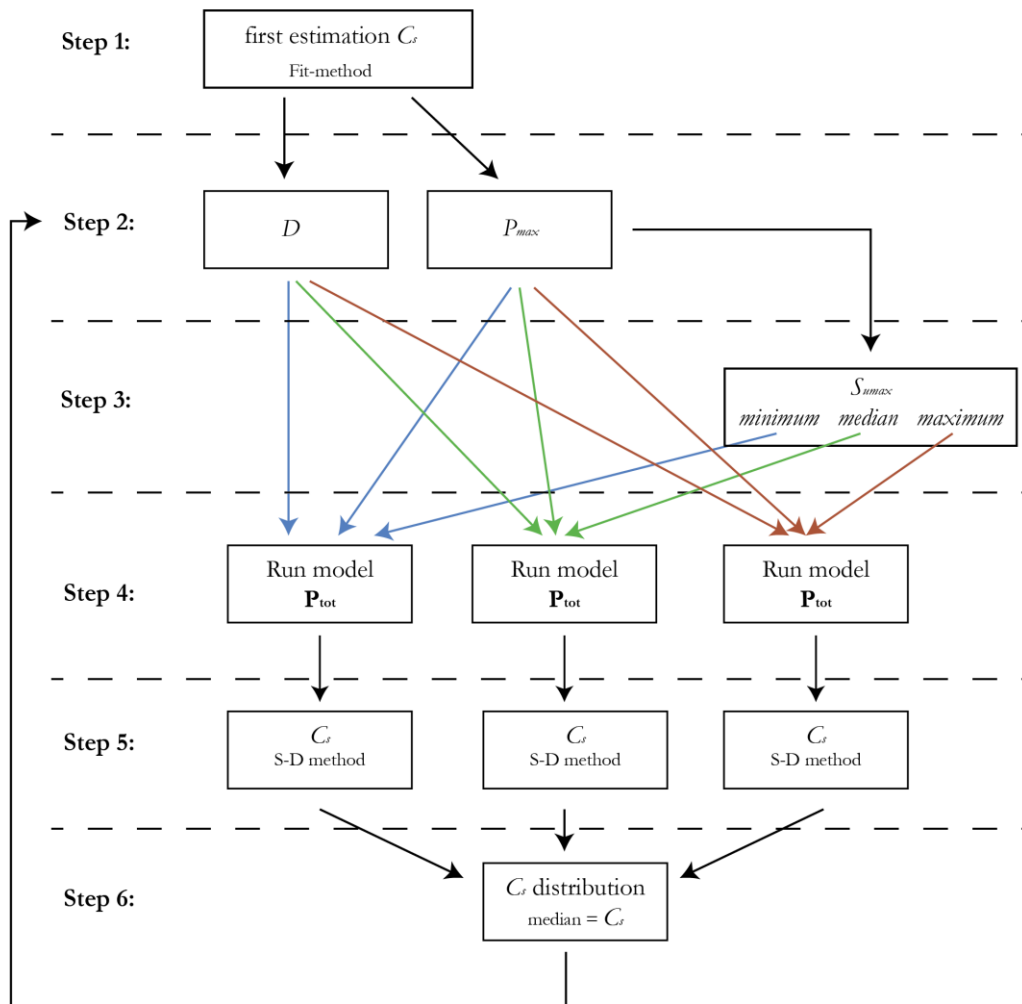


Figure 22. Schematic representation of the procedure to find the parameter distributions of models 7 and 9.

The number of iterations needed to obtain convergence was variable. No automated condition was set to define convergence. In this research, we choose to visually inspect the development of the parameters to determine convergence during the iterative process.

3.3 Sensitivity analysis methods for data uncertainty

In the previous section, different methods were presented to find parameter distributions in a synthetic experiment. The data used in the synthetic experiment was assumed to be “true forcing and discharge data” without any uncertainty. In this sensitivity analysis, the methods were tested with the introduction of data errors. As mentioned in the introduction, multiple data errors are possible. An example of data errors is the use of a point measurement of precipitation for catchment modelling while the precipitation over an area is highly variable in both time and space. Also, errors in measurement equipment contribute to the data uncertainty.

The model used for this analysis is the most complex model, number 9. The synthetic generated data sets were corrupted in both the forcing and discharge data. This analysis was executed with an absence of model structure uncertainty. The test for the sensitivity of the methods regarding forcing errors is described in Section 3.3.1. A description of the test for the sensitivity of the methods regarding both the precipitation and discharge data uncertainty is presented in Section 3.3.2.

3.3.1 Precipitation error

The first data uncertainty test was performed to examine the effect of spatial precipitation variability on the methods. Discharge was produced using forcing data from a point measurement location, and forcing data from another station was used to determine the parameter distributions. The stations used are both located in the Netherlands and owned by the KNMI, royal Dutch meteorological institute. The year volumes of the precipitation and the potential evaporation (calculated with Makking) are listed in Table 3. Small variances in the volumes were observed; however, neither Twenthe or Hupsel dominate.

Precipitation [mm/year]	2000	2001	2002	2003	2004	2005	2006	2007
Twenthe	852	836.2	780.9	618.8	854.7	796.8	723.8	916.8
Hupsel	869.5	817.7	841.8	719.8	805.5	813.5	734.4	949.2
Potential Evaporation [mm/year]								
Twenthe	528.4	553.3	547.8	638.9	562.6	584.3	592.6	562.5
Hupsel	548.5	560.9	556.6	639.4	571.5	594.9	602.1	570

Table 3. Annual volumes for precipitation and potential evaporation Twenthe and Hupsel stations

The only method that directly uses the measured precipitation is the bounds method for $S_{U,max}$. Section 3.1.3 mentions that the bounds become stable after precipitation events longer than 150 days. The long-term errors are more interesting, and Figure 23 illustrates a double mass. For 2007, the mean error is 6.6 mm, and the standard deviation in the error is 24.69 mm. A positive mean represents a larger amount of precipitation measured in Twenthe than in Hupsel. For all other years, the mean varies between -61.44 and 36.4 mm. The standard deviation varies between 9.2 and 33.1 mm.

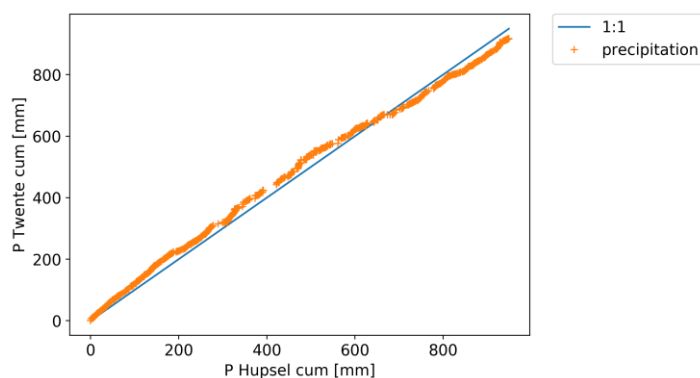


Figure 23. Cumulative sum of the measured precipitation (mm) in Hupsel compared to the cumulative sum of the measured precipitation (mm) in Twenthe (data of 2007)

3.3.2 Data corruption

The absolute error in the data is impossible to determine. However, to mimic data errors in synthetic experiments various studies use multiplying constants. The assumed “true” forcing and discharge data of the synthetic experiment are corrupted by the unique multiplying constants. These multipliers make it possible to change both the magnitude and pattern of the observed data (Butts, et al., 2004; Kavetski, et al., 2006; Ajami, et al., 2007; Renard, et al., 2010).

The error term used in this study originates from Renard (2010). The numbers, however, were changed to match the errors found in the previous section (3.3.1). Errors were added to the forcing data using a lognormal distribution. Also, the hydrograph was corrupted using a normal distribution to mimic all sorts of uncertainties, such as measurement errors.

The assumed “true” forcing data was used to generate synthetic discharge data. To test the methods, the “true” precipitation (r) was corrupted to generate an observed precipitation time series (\check{r}) equations 25 and 26. The error model used is as follows:

$$\check{r}_i = \frac{r_i}{\exp(m_i)} \quad (25)$$

$$m_i \sim N(0, 0.5^2) \quad (26)$$

Where r_i represents the i th “true” precipitation measurement, \check{r}_i is the corrupted i th measurement and m_i the random unique error. m_i originates from a normal distribution with a mean of zero and a variance of 0.5². In the corrupted time series, all precipitation values below zero were set to zero. This error led to a mean error of approximately 20 mm and a standard deviation of approximately 15 mm.

The generated discharge series (Q) with the “true” forcing data was also corrupted with an error term to produce an observed discharge series (\check{Q}) (equations 27 and 28).

$$\check{Q}_i = Q_i + e_i \quad (27)$$

$$e \sim N(0, (0.1Q_i)^2) \quad (28)$$

Q_i represents the i th assumed “true” discharge and \check{Q}_i the i th assumed observed discharge. The error term e_i originates from a normal distribution with a standard deviation of $(0.1Q_i)^2$. This error leads to a mean error of approximately -1.8e-6 mm/h and a standard deviation of approximately 1.3e-4 mm/h.

3.4 Real-word application

This last experiment was a real-world data study. This study provided insight into the performance of the methods under all sources of uncertainty. It was not possible to check whether the methods’ results had similar qualitative parameter distributions as in the synthetic experiments since the original parameter values were not known.

The methods were applied to the data of the Kervidy-Naizin catchment in France. The application of the methods could not be achieved without some adjustments to the P_{max} method. The described technique to filter out the peaks (Section 3.1.4) did not provide sufficient results. In the data set, measurements of zero discharge was present, which formed an implication for the multiplication factor. After a zero discharge measurement the filter only returns zero values and also low flow periods are filtered from the discharge data. As an alternative, all discharge measurements exceeding a threshold (set to 0.03 mm/h) were considered peak flows and were filtered from the discharge series.

For this catchment, it was assumed that during low flow, recharge enters the slow reservoir. The approach to find parameter distributions for the real-world data is the same approach as described in Section 3.2, Figure 22.

3.4.1 Comparison traditional and new calibration approach

A Monte-Carlo sampling strategy was performed to further investigate effectiveness of this study's newly developed methods to find parameter distributions. A comparison was made between an uninformed parameter distribution and the informed prior parameter distribution originating from the methods described in Section 3.1. Both parameter distributions were compared with the performance of the model. The uninformed parameter distribution and the new parameter distributions found in the methods are presented in Table 4.

After a warm-up period of one year, the calibrations were completed between January 1, 2006 and December 31, 2008. From the distributions 10⁶ parameter sets were randomly selected and assessed using a multi-objective evaluation (Gupta, et al., 1998) with three objective functions. In Table 5, the objective functions are described. The Euclidean distance (29) was used as an overall performance indicator of the model.

$$\text{Euclidean distance} = \sqrt{(1 - NS)^2 + (1 - \log NS)^2 + (1 - R)^2} \quad (29)$$

The GLUE method (Beven & Binley, 1992) was used to determine a 95% confidence interval. All parameter sets that perform below a Euclidean distance of 0.75 were accepted as behavioral models. For each time step, all behavioral parameter sets were used to calculate all possible modelled discharges. The 2.5 percentile and 97.5 percentile were calculated from the modeled discharges per time step. Then, the 2.5 percentile and 97.5 percentile were used to construct the 95% uncertainty interval.

The parameter distribution for D originating from the D -method was changed in this real-world test. The lower bound was originally set to 0.9951. Whether this is a "correct" number is debatable. The method to calculate the D distribution was limited to three time steps. The time resolution in the data set was one hour. However, the peak volume could not reach the measuring point in only one hour since the catchment size is 4.9 km². To account for this unsuitable assumption, the lower bound was changed to 95% of the original lower bound of the method.

	<i>Uninformed prior parameter distributions</i>	<i>Informed prior parameter distributions from expert-knowledge inverse modelling methods</i>
$C_s [h]$	0-3500	87-294
$P_{max} [mm/h]$	0.0009-0.16	0.0193 -0.041
$S_{U,max} [mm]$	50-500	108-270
$D [-]$	0-1	0.94-1

Table 4. Prior parameter bounds for MC simulation of model 9

Name	Formula	Range	Ideal value	Notes
Nash-Sutcliffe model efficiency of the flow (NS)	$O_{NS} = 1 - \frac{\sum_{i=1}^n (Q_{o,i} - Q_{s,i})^2}{\sum_{i=1}^n (Q_{o,i} - \bar{Q}_o)^2}$	$(-\infty, 1)$	1	Focus on the high flows. A value of zero indicates performance no better than using the mean flow. Negative values indicate an even worse performance
Nash-Sutcliffe model efficiency of the logarithm of the flow (logNS)	$O_{\log NS} = 1 - \frac{\sum_{i=1}^n (\log(Q_{o,i}) - \log(Q_{s,i}))^2}{\sum_{i=1}^n (\log(Q_{o,i}) - \log(\bar{Q}_o))^2}$	$(-\infty, 1)$	1	Focus on the low flows. A value of zero indicates performance no better than using the mean flow. Negative values indicate an even worse performance
Correlation coefficient (R)	$O_R = \frac{\sum_{i=1}^n (Q_{s,i} - \bar{Q}_s)(Q_{o,i} - \bar{Q}_o)}{\sqrt{\sum_{i=1}^n (Q_{s,i} - \bar{Q}_s)^2 \sum_{i=1}^n (Q_{o,i} - \bar{Q}_o)^2}}$	$(-1, 1)$	1	Focus on dynamics of the flows

Table 5. Objective functions used to assess the model performance

Chapter 4 Results

4.1 Results of parameter distribution methods

Some of the parameter estimation methods depend on other parameter values, as mentioned in Section 3.1. To investigate the errors in the parameter estimation methods, the other parameters needed to determine the current method parameter were assumed to be known and set to the original value, which was used to produce the discharge data. Therefore, the methods' input consisted of data without uncertainties and original parameters except for the value that needed to be obtained by the method. Only one obtained parameter value and distribution is displayed per model; all other results are presented in Appendix 4. In the example distributions in this section the red diamond represents the “correct” original parameter value of the data set.

Results of C_s fit method

The results of the C_s fit method are provided in Table 6. This method returns the original values in the synthetic experiment for data sets of models 1, 2, 3, 4 and 6. A small standard deviation (std) was observed for parameter distributions of these models. This originates from the solve method in Python. However, the method returned values that substantially deviated from the original values for models 7 and 9. The error in models 7 and 9 could be explained by the influence of the percolation flux, which does not reach zero during recession periods. The unsaturated zone will never be empty since the evaporation and percolation are proportional to the storage in this zone.

	<i>Model 1</i>	<i>Model 2</i>	<i>Model 3</i>	<i>Model 4</i>	<i>Model 6</i>	<i>Model 7</i>	<i>Model 9</i>
<i>Original C_s (h)</i>	200	20	200	100	50	500	470
<i>Fit method C_s (h)</i>	200	20	200	100	50	1964	719
<i>Std</i>	1.56E-06	9.99E-02	8.78E-02	1.51E-03	2.56E-02	1.97E+03	1.41E+03
<i>Relative error (%)</i>	-6.37E-08	3.98E-06	-9.47E-08	7.05E-10	1.07E-06	2.92E+02	5.28E+01

Table 6. Examples of the results of the synthetic experiment using the C_s fit method

Results of C_s S-D method

The C_s S-D method returns parameter distributions that contain the original values for the synthetic experiment data for models 7 and 9, as presented in Table 7. The std decreased by a factor 1000 compared to the fit method. The boxplots are a standardized method to display value distributions. The orange line represents the median in the data set. The interquartile range (IQR), between the first and the third quartile, is the likely range of variation in the data set. The maximum is considered 1.5 x IQR above the third quartile and the minimum is considered below 1.5 x IQR of the first quartile. All other values above the maximum and below the minimum are displayed as outliers. In distribution of the C_s values, the spread in the distribution and extreme outliers are clearly visible for the fit method.

	<i>Model 7</i>	<i>Model 9</i>
<i>Original C_s (h)</i>	500	470
<i>Fit method C_s (h)</i>	1977	620
<i>Std</i>	2004	1274
<i>Distribution C_s fit method</i>		
<i>S-D method (h)</i>	500.6	470
<i>std</i>	0.13	1.12
<i>Distribution C_s S-D method</i>		

Table 7. Examples of the results of the synthetic experiment using the S-D method

Results *D*

The *D* method provided accurate results in the synthetic experiment. Some outliers and a wider spread were observed in the distribution of model 9 (see Table 8). This is explained by an incorrect calculation of the overflow. The percolation flux was not considered in the calculation of the overflow. However, this assumption did not have a significant impact (relative error between the median of the estimated distribution and the original value of 0.18%) on the estimation of *D* distribution for the model 9 data.

	<i>Model 4</i>	<i>Model 9</i>
<i>Original value</i>	0.96	0.94
<i>Number of overflows used to calculate <i>D</i></i>	60	35
<i>Median</i>	0.96	0.941
<i>Relative error [%]</i>	0	0.18
<i>Distribution</i>		

Table 8. Examples of the results of the synthetic experiment using the *D* method

Results of $S_{U,max}$ bound and inter-peak method

The bound and inter-peak method returns distributions that include original parameter values for synthetic data of model 4. However, for models 7 and 9, the distributions of the inter-peak method deviated from the original values; model 9 deviated substantially. The inter-peak method is quite sensitive to small errors. For models 7 and 9, the method was considered too sensitive and could not be used to determine the $S_{U,max}$. The bounds method, however, provided a sufficient result.

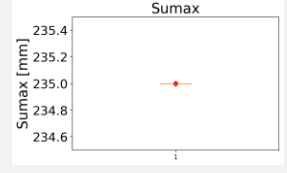
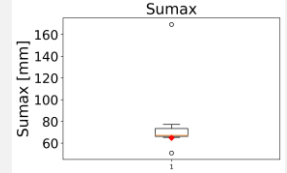
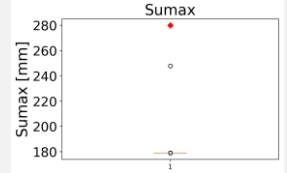
	<i>Model 4</i>	<i>Model 7</i>	<i>Model 9</i>
<i>Original value [mm]</i>	235	65	280
<i>Minimum [mm]</i>	153	44	178
<i>Maximum [mm]</i>	288	86	309
<i>Inter peak method</i>			
<i>Median [mm]</i>	235	67	179
<i>std</i>	3.0e-7	26	12
<i>Distribution</i>			

Table 9. Examples of the $S_{U,max}$ results of synthetic experiment using the inter peak method

Results of P_{max} method

This method had expected results in the artificial world for models 6, 7 and 9. The original values of the P_{max} are included into the distribution. The 75th percentile of the distributions were a lower than the original values of the P_{max} . This underestimation in the “most optimal” value in the distribution could be prevented, but it was determined to be safe and the maximum was not attempted in the calculated percolation rates. The maximum in the distribution is quite sensitive for outliers.

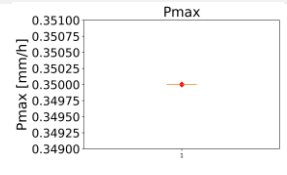
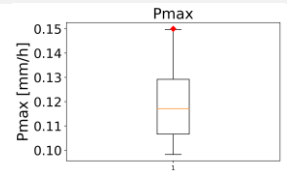
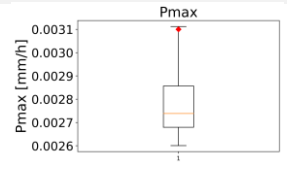
	<i>Model 6</i>	<i>Model 7</i>	<i>Model 9</i>
<i>Original value [mm/h]</i>	0.35	0.15	0.0031
<i>75th percentile [mm/h]</i>	0.35	0.129	0.0028
<i>Relative error [%]</i>	0	-13.85	-9.68
<i>Min. distribution</i>	0.35	0.098	0.0026
<i>Max. distribution</i>	0.35	0.15	0.0031
<i>Distribution</i>			

Table 10. Examples of the results of synthetic experiment using the P_{max} method

4.2 Results synthetic experiments

Some of the parameter estimation methods depend on other parameter values as explained in Section 3.1. However, the methods return parameter distributions. In this synthetic experiment, the missing input parameters to the methods are the “most optimal” number of the previously determined distributions. These “most optimal” parameters are the median for the C_s and D , and the 75th percentile for the P_{max} . These numbers contain errors in the parameter estimation. This section explains the extent to which these errors in the determination of one parameter propagate to other parameter estimations.

4.2.1 Results for model 4

For model 4, three different data sets were tested. A summary of the outcome from the parameter estimations methods is given in Table 11 and Table 12.

Groundwater drainage parameter C_s

The fit-method was used to calculate the C_s value of the data set since there was no flux entering the slow reservoir during recession periods. In this data set a small std was observed due to some outliers. These outliers originated from the solve method used in Python.

Partitioning coefficient D

The partitioning coefficient was calculated using the median C_s value calculated in the previous step. For the synthetic data of model 4, the D can be calculated accurately. All original values of the data sets are included in the estimated parameter distributions.

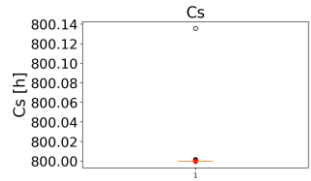
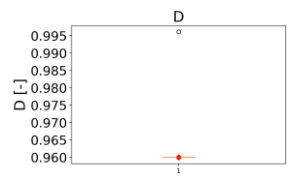
	C_s	D
Relative error	0 %	0 %
Std [-]	0.001 to 0.3	1e-9 to 4e-3
Example distribution		

Table 11. Summary of the results of synthetic experiment model 4 for C_s and D

Maximum unsaturated zone storage $S_{U,max}$

The method to determine the bounds of the $S_{U,max}$ is, in this case, independent of other parameter values since no percolation flux is present in the model. The overflow data from the calculations made for D were used to determine the $S_{U,max}$ distribution with the inter-peak method. The number of calculations, however, were considerably fewer than the number for D (see Appendix 4).

Sometimes, the length of the data series between two overflow moments were less or more than the predefined lengths. In the winter, more overflow moments occur; however, a storage change in the unsaturated zone (evaporation between overflow moments) is not always present. These discharge series were not used to calculate $S_{U,max}$.

$$S_{U,max}$$

Relative error underestimation of $S_{U,max}$	-20% to -35%
Absolute error underestimation of $S_{U,max}$	10-85 mm
Relative error overestimation of $S_{U,max}$	23% to 83%
Absolute error overestimation of $S_{U,max}$	20-85 mm

Table 12. Summary of the results of synthetic experiment for model 4 for $S_{U,max}$

4.2.2 Results for models 7 and 9

The parameter values of three different data sets for models 7 and 9 were used to determine parameter distributions in this section. The approach as described in Section 3.2 (Figure 22) was used to determine the parameter values. First, a summary of the results of model 7 is given (Table 13 and Table 14). Then, the results for model 9 are provided (Table 15 and Table 16). For each parameter, a short description of the results is given.

Model 7

Groundwater drainage parameter C_s

To calculate the C_s value of the data set, the fit method was used for the first estimate, while the S-D method was used for the later iterations. Relative errors between the median of the estimated distribution and the original value were found between -15% and 30%. However, all original values of the data sets were included in the estimated parameter distributions. After three iterations, all three sets converged, and the parameters were stable during the later iterations. The error in the C_s was directly linked to the large overestimation of the maximum $S_{U,max}$.

Maximum percolation rate (P_{max})

The first calculated maximum percolation rate was overestimated; this is related to the first overestimation of the C_s . After three iterations, the value converged. For all three data sets, the P_{max} (75th percentile of the distribution) was underestimated. The relative difference between the original value and the estimated value ranged between -3% and -17%. The same underestimation was observed in the method results (Section 4.1). The estimated parameter distributions include the original values of the data sets.

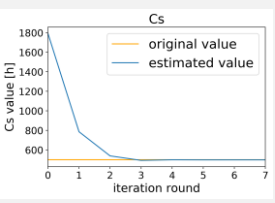
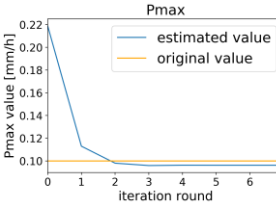
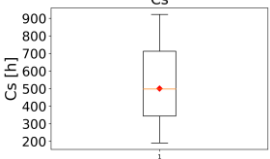
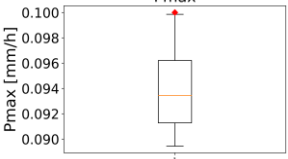
	C_s	P_{max}
Relative error	-15% to 30%	-17% to -3%
Std [-]	50-180	
Example iterations		
Distribution		

Table 13. Summary of the results of synthetic experiment for model 7 for C_s and P_{max}

Maximum unsaturated zone storage $S_{U,max}$

The method to determine the bounds of the $S_{U,max}$ is not independent of other parameter values for model 7. The P_{max} is now also included into the calculations. The larger overestimation of the maximum storage capacity was notable.

	$S_{U,max}$
Relative error underestimation of $S_{U,max}$	-35 % to -30%
Absolute error underestimation of $S_{U,max}$	19-50 mm
Relative error overestimation of $S_{U,max}$	40 to 86%
Absolute error overestimation of $S_{U,max}$	21-63 mm

Table 14. Summary of the $S_{U,max}$ results of synthetic experiment for model 7

Model 9

Groundwater drainage parameter C_s

The fit method was used to calculate the first estimate for the C_s value of the data set. The S-D method was used for the later iterations. After three iterations, all three sets converged. Relative errors between the median of the estimated distribution and the original value under 5% were observed for model 9. All original values of the data sets were included into the estimated parameter distributions.

Partitioning coefficient D

The partitioning coefficient was calculated using the median of the estimated C_s distribution. The calculated results for D did not deviate substantially from the original parameter of the data sets; relative errors with a maximum of 0.3% were observed. However, the estimated parameter distributions did not include the original values.

Maximum percolation rate (P_{max})

The first calculated maximum percolation rate was overestimated; this is related to the first overestimation of the C_s . For all three data sets, the P_{max} (75th percentile of the distribution) was underestimated. The relative difference between the original value and the estimated value ranges between -4% and -10%. The same underestimation was observed in the method results (Section 4.1). The estimated parameter distributions include the original values of the data sets.

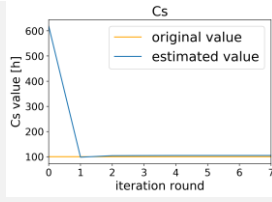
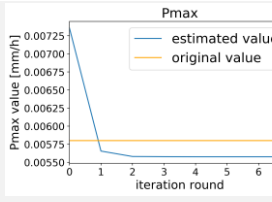
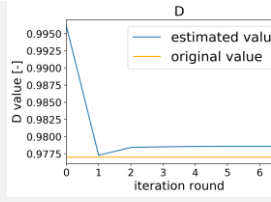
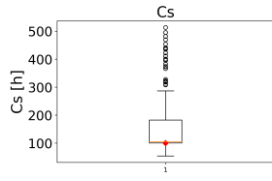
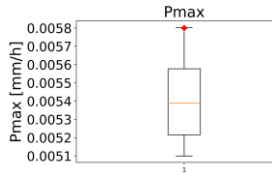
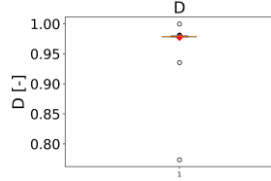
	C_s	P_{max}	D
Relative error	0% to 4%	-4 % to -10%	0% to 0.3%
Std [-]	40-200		0.001-0.009
Example iterations			
Distribution			

Table 15. Summary of the results of synthetic experiment for model 9 for C_s , P_{max} and D

Maximum unsaturated zone storage $S_{U,max}$

The P_{max} parameter was included in the calculations for the bounds of $S_{U,max}$. The method was not independent of other parameter values. The underestimation of P_{max} from the previous analysis did not significantly influence the results. Both the minimum and maximum were correctly estimated.

$S_{U,max}$	
<i>Relative error underestimation of $S_{U,max}$</i>	-40% to -30%
<i>Absolute error underestimation of $S_{U,max}$</i>	70-120 mm
<i>Relative error overestimation of $S_{U,max}$</i>	10% to 40%
<i>Absolute error overestimation of $S_{U,max}$</i>	20-140 mm

Table 16. Summary of the $S_{U,max}$ results of synthetic experiment for model 9

4.3 Results of sensitivity analysis

4.3.1 Results for model 9 including the forcing error

Three data sets were used to test the effect of the forcing data uncertainties on the methods to determine parameter distribution. The same approach described in Section 3.2, Figure 22, was used to find the parameter distributions.

In the bounds method for $S_{U,max}$, the parameter determination method took the precipitation and evaporation data as input. This method was assumed to have the largest effect on the corrupted forcing data. The summary of the over- and underestimation of $S_{U,max}$ is given in Table 17. For all tests, the minimum was always an underestimation of the original $S_{U,max}$ used to produce the data. The maximums for all three tests were larger than the original used $S_{U,max}$ to produce the data. The error for both the minimum and the maximum, however, did not deviate substantially from the error observed in the synthetic experiment without data errors. The absolute difference in minimum and maximum found in the synthetic experiment of Section 3.2 and found in this analysis is equal to the standard deviation of the precipitation measurement error.

	$S_{U,max}$
Relative error underestimation of $S_{U,max}$	-50% to -60%
Absolute error underestimation of $S_{U,max}$	113-187 mm
Relative error overestimation of $S_{U,max}$	10% to 60%
Absolute error overestimation of $S_{U,max}$	58-132 mm

Table 17. Summary of the results for $S_{U,max}$ with precipitation error

Considering the results presented in Table 18, the effect on the other parameter distributions of the error in $S_{U,max}$ is negligible. The C_s , P_{max} and D can still be found with high accuracy.

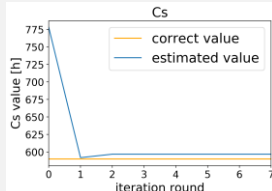
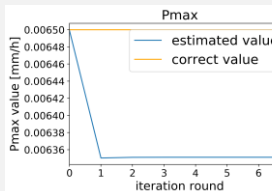
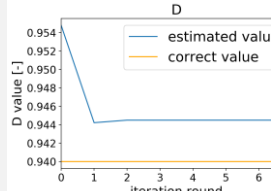
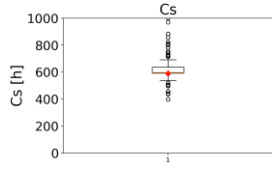
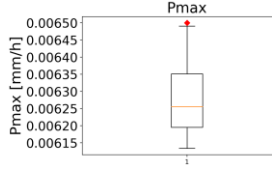
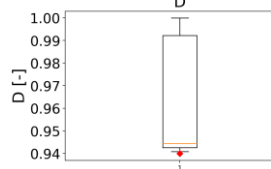
	C_s	P_{max}	D
Relative error	0% to 3%	-2% to -3%	0% to 0.5 %
Std [-]	60-130		0.01-0.05
Example iterations			
Distribution			

Table 18. Summary of the results for C_s , P_{max} and D parameter distributions with precipitation error

4.3.2 Results of model 9 including the data error

Three data sets were used to test the effect of the data error on the methods to find the parameter distributions. The same approach described in Section 3.7, Figure 22, to determine the parameter distributions was used.

The bounds for the maximum unsaturated zone storage were still quite accurate (Table 19). The minimum was underestimated by 30% to 40%. The maximum was overestimated by 10% to 30 % of the original used value to produce the data.

$S_{U,max}$	
<i>Relative error underestimation of $S_{U,max}$</i>	-30% to -40%
<i>Absolute error underestimation of $S_{U,max}$</i>	75-128 mm
<i>Relative error overestimation of $S_{U,max}$</i>	10% to 30%
<i>Absolute error overestimation of $S_{U,max}$</i>	56-78 mm

Table 19. Summary of the results for $S_{U,max}$ with data corruption

The effect of data uncertainty in the discharge became considerable in the P_{max} calculation. In the P_{max} distributions, an overestimation of the original value was observed. Also, in one data set, the minimum distribution was higher than the original P_{max} value used to generate the “true” data set. For the other two sets, the original value was included in the estimated parameter distribution. The overestimation of P_{max} directly translated to the calculation of the C_s distribution, where the median in the distribution became an underestimation.

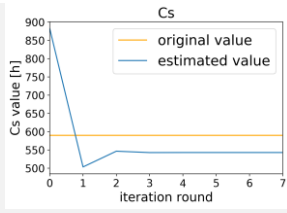
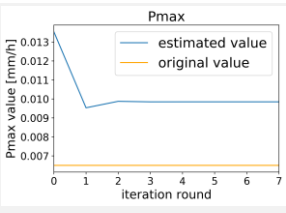
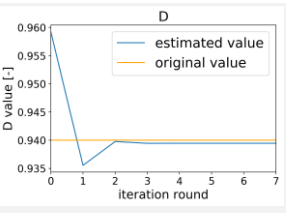
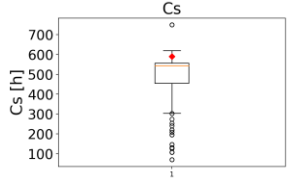
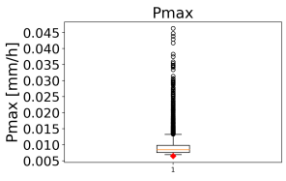
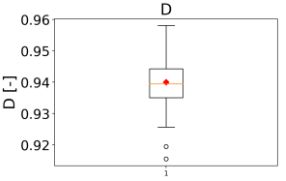
	C_s	P_{max}	D
<i>Relative error</i>	-20% to 0%	0% to 50%	-2% to 1%
<i>Std [-]</i>	40-200		0.01-0.02
<i>Example iterations</i>			
<i>Distribution</i>			

Table 20. Summary of the results for C_s , P_{max} and D parameter distributions with data corruption

4.4 Results of real-world application

The results of the methods to determine parameter distributions for the real-world data set are presented in Table 21. All the methods converged, as the results of the iterations indicate. The small distribution of the D was remarkable. In this case, it was no longer possible to know the “true” value of the data set’s parameters, and a comparison of the error is not shown.

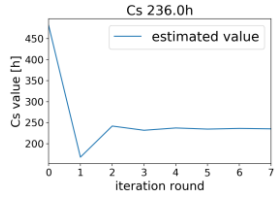
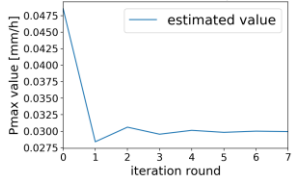
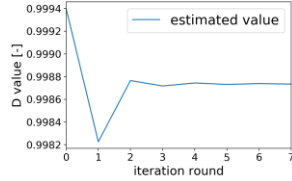
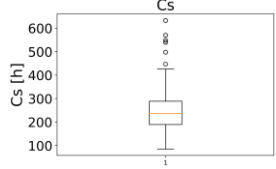
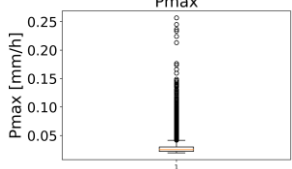
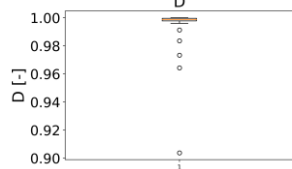
	C_s	P_{max}	D
<i>Estimated value</i>	236 h	0.0299 mm/h	0.9987
<i>Results of iterations</i>			
<i>Std [-]</i>	61.3	0.02026	0.016958
<i>Distribution</i>			
<i>Max boxplot</i>	426	0.041	0.99998
<i>75 percentile</i>	288	0.0299	0.9996
<i>Median</i>	236	0.02518	0.9987
<i>25 percentile</i>	189	0.02201	0.9978
<i>Min boxplot</i>	85	0.01932	0.9951

Table 21. Results of real-world application for C_s , P_{max} and D

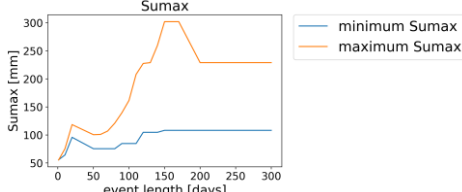
	$S_{U,max}$
<i>Minimum [mm]</i>	108
<i>Maximum [mm]</i>	302
<i>Threshold for different event lengths</i>	

Table 22. Results of real-world application for $S_{U,max}$

4.4.1 Results comparison

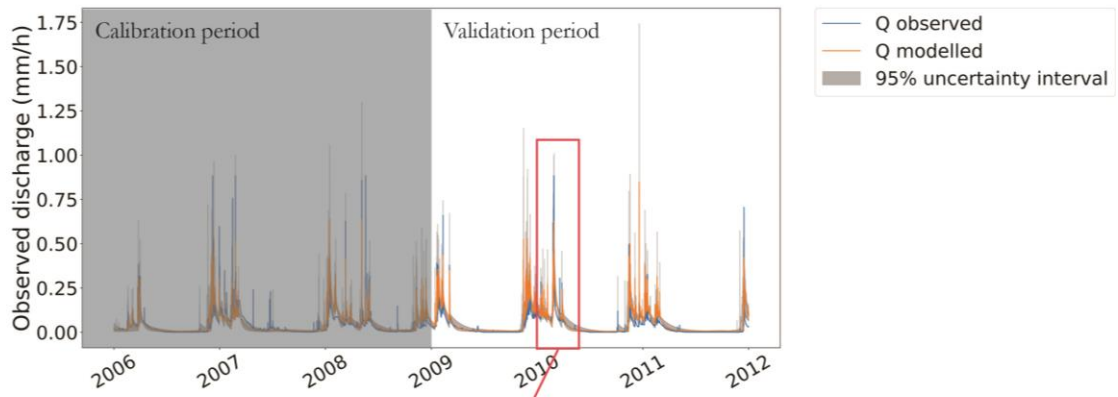
Both the performance of the model using the calibration results of the uninformed prior parameter distributions and the new informed prior parameter distributions found in the methods are presented in Figure 24 and Table 23. In the model performance, no significant difference was observed in the calibration (2006-2008) or validation period (2009-2011). The overall performance (ED) of the models with the uninformed parameter distributions was slightly higher for all periods. A notable difference was observed in performance of the logNS for the parameter distributions of the informed prior parameter distributions: the performance was less than the uninformed parameter distributions. The model uncertainty area spanned from the 2.5 percentile to the 97.5 percentile was more than double for the results of the uninformed prior parameter distributions than for the informed prior parameter results.

Figure 25 illustrates both the prior parameter distributions and the posterior parameter distributions of both the uninformed and informed parameter distributions. The posterior distributions are the distributions of the set parameters originating from behavioral models. Overall, a decreasing trend was observed from the prior uninformed parameter distributions to the informed posterior distribution. The posterior distributions of the informed parameter sets were twice to five times smaller than the posterior distributions of the uninformed parameter sets. This indicates a decrease in equifinality. The posterior distribution of the P_{max} shifted in the informed parameter sets compared to the uninformed sets. For the other parameters, no significant shift was observed.

	<i>Uninformed prior 2006-2011</i>	<i>Informed prior 2006-2011</i>	<i>Uninformed prior 2006-2008</i>	<i>Informed prior 2006-2008</i>	<i>Uninformed prior 2009-2011</i>	<i>Informed prior 2009-2011</i>
<i>Mean ED</i>	0.66	0.730	0.679	0.730	0.648	0.733
<i>Mean NS</i>	0.539	0.555	0.509	0.539	0.565	0.567
<i>Mean LogNS</i>	0.587	0.460	0.614	0.490	0.563	0.434
<i>Mean R</i>	0.782	0.794	0.742	0.753	0.825	0.837
<i>Most optimal ED</i>	0.520	0.680	0.555	0.688	0.438	0.678
<i>Most optimal NS</i>	0.620	0.604	0.563	0.578	0.676	0.631
<i>Most optimal LogNS</i>	0.709	0.481	0.763	0.510	0.751	0.452
<i>Most optimal R</i>	0.797	0.804	0.752	0.764	0.844	0.842
<i>95% uncertainty area [mm]</i>	1887	741	884	341	1003	401

Table 23. Performance results of the uninformed prior parameter distribution and the informed prior parameter distributions over multiple periods. The mean performance represents the mean of all simulated discharges in the uncertainty bound, the most balanced (optimal) performance in the uncertainty bound is also given in this table.

Results of uninformed parameter sets



Results of informed parameter sets

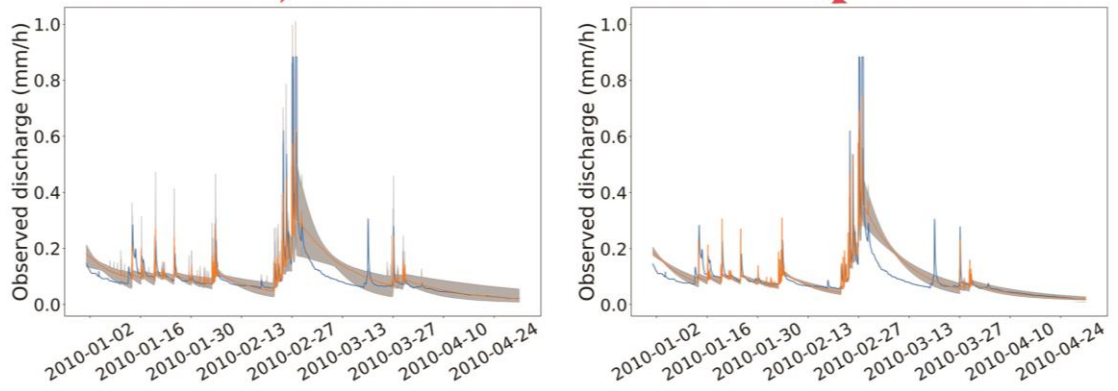
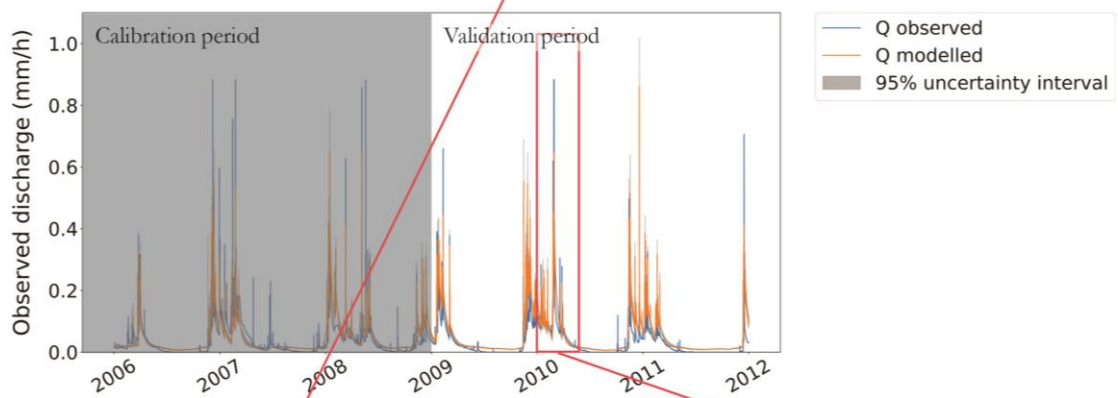


Figure 24. Observed (blue line) and most optimal simulated (orange line) discharges in the calibration and validation periods. The simulated discharge is shown with a 95 percent uncertainty interval (shaded grey area).

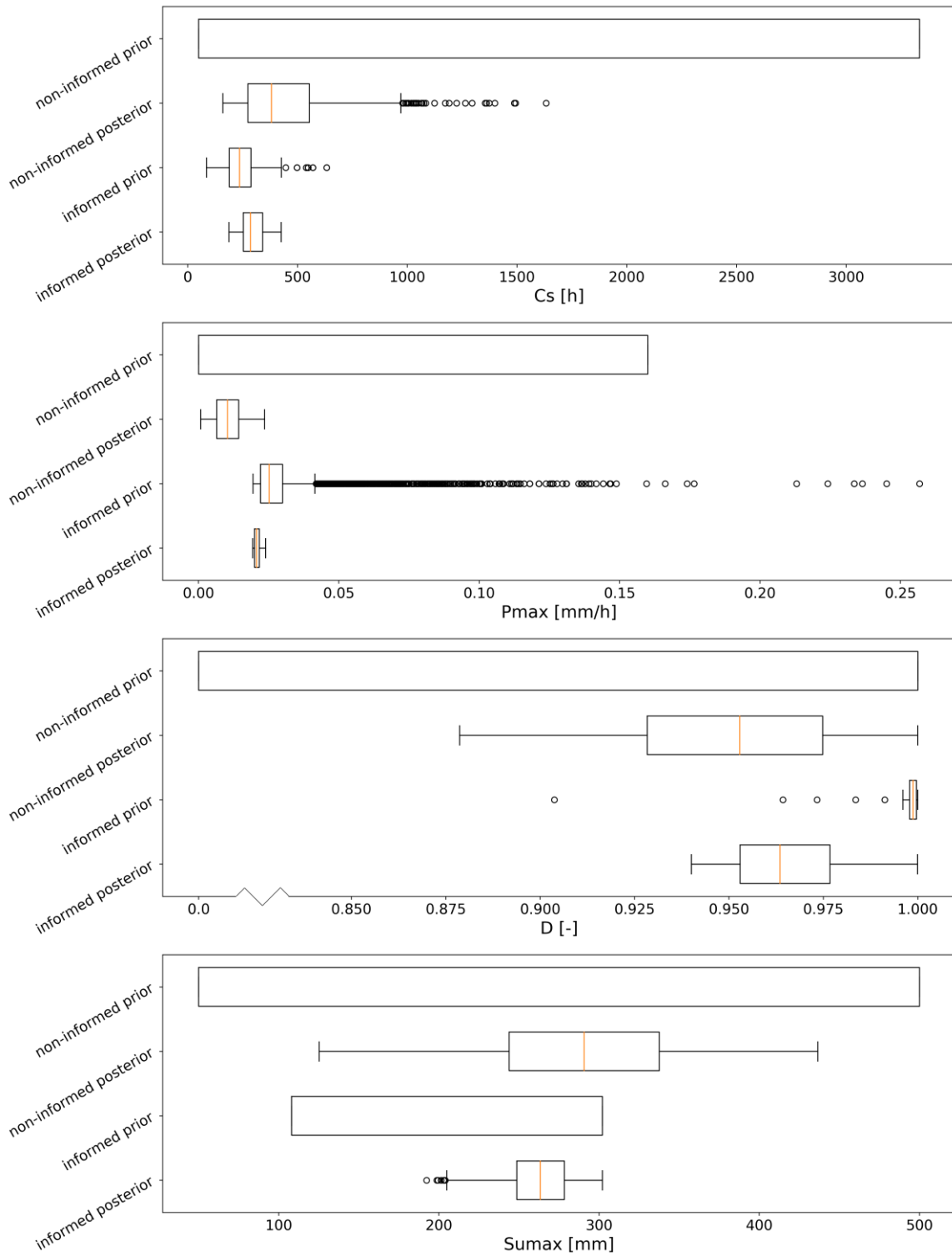


Figure 25. The prior parameter distributions and the posterior parameter distributions of all behavioral models. For both the uninformed parameter distribution and the informed parameter distribution. (boxplots: the dots represents the outliers in de data set, the lower and upper whisker 2.5/97.5th percentiles and, the horizontal orange line the median)

Chapter 5 Discussion

5.1 Methods and synthetic experiment

The chapter reviews the developed methods to obtain parameter distributions and the results of the different research sections. Both the strengths and weaknesses are discussed. In addition, links between the synthetic, sensitivity analysis and real-world application are made.

5.1.1 Methods

In all six methods, to constrain the feasible parameter space, assumptions were made based on expert judgment, which makes the methods less objective. In particular, certain thresholds were defined by visual inspection of the data or by testing different thresholds to assess the method performance. Examples of thresholds based on expert judgment are the minimum duration of no precipitation before a recession period, the maximum discharge at the end of a recession and the multiplication factor in the calculation of D . Whether values determined by expert judgment deviated for other catchments was not tested. Therefore, it is unclear how the methods will perform, and critically assessing the defined parameter distributions is advised.

In the calculation of the C_r value, one major assumption was made in the S-D method. Model runs were performed to calculate the total recharge (P_{tot}). The parameters in the model and the assumed model structure had an impact on the calculated recharge. This recharge affected the calculation of the C_r value, and the S-D method therefore qualified as quite subjective. The method became sensitive to errors in the assumed parameters and the model structure. Although errors were made in the S-D method, the results indicate that the method performs quite well.

In the calculation of the D distribution, it was assumed that the total volume of the fast lateral runoff processes would reach the outflow point of the catchment in one time step. Depending on the size and the time resolution of the measurements, the total volume of fast lateral runoff can reach the outflow point in one time step. However, large catchments are more likely to have some sort of channel routing, resulting in a longer time frame for the total peak volume to reach the outflow point. The value for D will be overestimated if calculations are made with a high time resolution, such as hourly resolution. Catchment size and characteristics should be examined to determine whether the assumptions in the method are applicable before applying this method. This method depends more on expert judgment than the other methods.

For the calculations of the $S_{U,max}$ bounds, an assumption was made in the calculation of the upper bound. The total sum of precipitation was subtracted with a percentage of the maximum percolation volume to decrease the overestimation of the maximum storage capacity. This assumption makes the method sensitive to errors in the calculation of P_{max} , especially if P_{max} is overestimated.

For the calculation of P_{max} , the parameter distribution was determined based on the upper 20% of all calculated percolation rates. The optimal P_{max} was decided, using expert judgment, to be in the 75th percentile of this distribution to filter out outliers. In the synthetic experiments, this result was always an underestimation, which was assumed to be in an acceptable range.

5.1.2 Comparison model run of models 4, 7 and 9

In this synthetic experiment, different data sets were used to find parameter sets for models 4, 7 and 9. The data sets were produced by the models and were not corrupted by any uncertainty, so model 4's output was used to find parameters for model 4. However, in this experiment, no prior knowledge of parameter values was present. This made it possible to discover how the methods cope with errors made in parameter distributions from one method and translate to other parameter distributions. For example, the calculation of the D distribution depend on the calculated C_r value.

In the synthetic experiment data of model 4, the parameter values could be found with high precision. Small errors in the methods did not influence the results. All parameter distributions had a small std, implying limited uncertainties in the estimated parameter values.

However, as mentioned in Section 4.1, the C_r calculations with the fit method were accurate. For model 7 and 9, however, large deviations from the original value were present and the std was quite significant. As explained in Section 3.2, an iterative process was used to determine the parameter distributions.

For all data sets from both model 7 and 9, a fast convergence of the parameter values was observed. In addition, no large deviations of the original value were present in the parameter value estimations; thus, the procedure is considered a suitable approach. In this research, no condition was set to break off the iteration since it was of interest to discover whether the parameter distributions stayed stable. In further development of the methods, the condition can be set to limit the calculation time.

The most remarkable difference between the results of model 7 and 9 is the performance of the C_r method. For model 7, the results differed more than model 9's from the original values. Also, the standard deviation was larger for model 7's data than for model 9's data. No seasonal changes were observed in model 7's data, which could be an indication of a unsuitable model structure. However, this possibility was not investigated in this research. The inability of a system representation could influence the calculations of the parameter distributions.

Another difference in performance between models 4, 7 and 9 is as follows: for model 7 and 9 only an estimation of the lower and upper bound of $S_{U,max}$ could be made. This affected the other parameter distribution calculations, resulting in a larger std of all the distributions.

5.2 Sensitivity analysis

5.2.1 Precipitation error

During this synthetic experiment, the forcing data was corrupted by producing data with forcing data of one point measurement station and using forcing data of another station to determine how the methods to determine parameter distributions behaved. Between the two stations, a difference in precipitation patterns was observed, which indicates a spatial variability in these patterns. Although the error in forcing data was applied, the methods performed quite well and were able to define parameter distributions.

In the last section, a comparison is made between the results of the “clean” data from model 9 and the corrupted data of model 9. However, this comparison is not completely valid since the “clean” data was created using with French forcing data and the precipitation error was created using Dutch forcing data. Nevertheless, the parameter distributions are similar for both the corrupted and “clean” data without large deviations in the original values of the synthetic data.

5.2.2 Data corruption

The second test in the sensitivity analysis was the corruption of both forcing and discharge data. The data was corrupted with unique multiplying constants, which changed both the magnitude and patterns of the data. The most striking effect on the methods was seen in the calculation of the P_{max} distribution. It was found that the original parameter of the data sets was sometimes no longer part of the defined distribution. This indicates that the method does not handle data uncertainty well. The part of the calculated percolation rates could be an inappropriate choice for the parameter distribution, or the method set up might need to be changed to decrease the sensitivity to data uncertainty by examining a longer period to calculate the percolation rates instead of using two subsequent time steps.

Although the P_{max} calculations had some errors, this did not have a significant effect on the calculations of the other parameters. The only affected parameter was the C_r distribution. The original values are in the upper part of the boxplot (Table 20).

5.3 Real world

All methods were applied in a real-world test case with data measured in a catchment in France. It is no longer possible to conclude whether the derived parameter distributions are in the “correct” range since this is not known in the real world. Nevertheless, it is possible to make comparisons with a Monte-Carlo (MC) sampling strategy with uninformed prior parameter distributions.

A change was made in the lower bound of the D to calibrate the model using the informed prior distributions from the methods. The assumption was made that the peak volume calculated in the method was underestimated. The time resolution in the data set was one hour, but arguably, the peak volume of this catchment could not reach the measuring point in one time step.

To confirm that the peak flow is underestimated, an extra test should be performed. In this research, only data uncertainty was added to the synthetic data set; it is also possible to add model structure uncertainty by producing discharge with another model and testing the methods with the model of this research. To investigate the method to determine D , adding a routing function at the fast lateral runoff processes is recommended. This could be accomplished, for example, by adding a fast reservoir.

The most notable difference between the MC strategy with uninformed prior parameter distributions and the informed prior parameter distribution from the methods was found in the results of P_{max} . A large deviation was observed for the behavioral models. The parameter distribution for the uninformed prior parameter distribution was in lower ranges than found in the methods. The same trend of overestimation of P_{max} was also observed in the sensitivity analysis of the methods.

An effect of a too high P_{max} value was observed in the C_s distribution. The majority of the behavioral models (the posterior parameter distribution) were also in a higher range compared to what the methods suggest in the informed prior parameter distributions. In the data uncertainty analysis, this higher C_s value behavior was also observed. This suggests that during calibration the model tried to correct for the unsuitable P_{max} distribution.

A significant decrease in the total uncertainty interval was observed in the comparison of the Monte-Carlo simulation with uninformed prior and informed prior parameter distributions of the methods. This implies that the model was acting with an improved physical realism.

The most important advantage of the methods to find parameter distributions is that the feasible parameter space was decreased by the methods. The model was, therefore, no longer able to compensate for unsuitable model structure assumptions. If the model’s performance decreases significantly compared to uninformed prior calibration results, it should be considered that the assumptions made in the model structure could be inappropriate and essential hydrological processes are left out. This knowledge helps hydrologists and model builders find an appropriate model structure to represent a catchment.

In this analysis, the methods to determine parameters were only tested on one small catchment. For such a simple lumped model with only four parameters, the performance was quite accurate. However, extra care should be taken if the methods are applied to large scale catchments since these catchments are often more heterogenic.

Chapter 6 Conclusions

Answers to the research question and sub-questions are provided in this chapter. The objective of this research was to determine the extent to which it is possible to make an estimation of the parameter values or parameter distribution for given hydrological model structure based on field observations. By extracting parameter information from the data, in theory, the model uncertainty and equifinality in the parameter sets should be reduced. The concept of this research was to use a general conceptual understanding of how hydrological processes in a catchment act. The parameter identification was performed based on sub-periods of different event types. The hypothesis used in this thesis was that each period has different “dominant” processes with associated parameters. Using the sub-periods of the hydrograph to find individual parameters of the model could further increase realism in the models.

The following research question was formulated for this research:

To what degree is it possible to avoid the use of uninformed prior parameter distribution in the calibration of a conceptual hydrological model by using available information from the field observations to generate informed prior distributions?

The sub-questions indicating the crucial issues to be addressed when answering this main question were the following:

- 1. Can we select sub-periods from a hydrograph typical of different hydrological processes which can be coupled to model components, and how can this help to determine model parameters?*
- 2. To what extent are the methods to obtain the parameter distributions sensitive to data errors?*
- 3. To what extent can the methods be used for real rainfall runoff data?*

The synthetic experiment proved that by using the expert-knowledge inverse modelling methods described in this research, we can determine all parameter distributions of the investigated model. Multiple sub-periods, which represent typical hydrological processes, were selected in a hydrograph. Specific model components of the hydrological model are able to describe these processes, and the connected parameter distributions can be found.

In the sensitivity analysis, the P_{max} method to determine the maximum percolation rate parameter distribution was sensitive to data uncertainty, resulting in errors in the determined parameter distributions. Since many of the methods take other parameter values such as the P_{max} as input, other parameter distributions will be affected by the errors due to the sensitivity to data uncertainties. The P_{max} method results in an overestimated parameter distribution. However, although the parameter is overestimated, the other parameters distributions include the original value of the synthetic data, which is considered a positive result.

Regarding the last sub-question, the expert-knowledge inverse modelling methods were applied to real rainfall runoff data. However, it was no longer possible to conclude whether the obtained parameter distributions are in the “correct” range since the parameters of the real world are not known. The comparison between an uninformed prior parameter distribution and the informed prior parameter distributions of the methods in a Monte-Carlo sampling strategy gave insight into the performance of the methods. With the GLUE method, a uncertainty interval was constructed. By using the parameter distributions of the methods, this interval decreased substantially. The total area of the uncertainty interval was less than half the uncertainty interval from the uninformed prior distributions; also, the posterior parameter distributions were two to five times smaller for the informed parameter case.

In this real-world test, a similarity in the performance for both calibrations was observed. These results endorse the importance and the advantages of the methods to obtain parameter information from the data. The uncertainty in the model decreased substantially, and it was possible to be more confident about how appropriate the model structure is for the catchment. With the informed parameter

distributions, the model was no longer able to compensate for inappropriate model structure assumptions, and a meaningful and more realistic representation of the catchment was achieved. The performance test also works the other way around. A significantly decrease in performance could be an indication of an unsuitable model structure where important hydrological processes are left out.

Using the information from the sub-questions, the main question can be answered. This study has demonstrated that by using a data-based techniques, all parameter distributions can be derived for the investigated model. The hypothesis of this study is confirmed.

6.1 Further research

For more general conclusions of the expert-knowledge inverse modelling methods to determine parameter distributions, more tests are needed, both in the synthetic environment and for real-world data. Further research should be directed toward testing whether the methods or approach can be generalized for multiple catchment typologies or whether the experience and expert judgment of the modeler keeps playing an important role.

Additional analyses should be made into the methods sensitive to data uncertainty. This study found that some methods return parameter distributions that did not include the original parameters of the data if data errors were present. Furthermore, synthetic tests including model structure uncertainty should be performed to test the general validity of the expert-knowledge inverse modelling methods.

Literature list

- Ajami, N. K., Duan, Q. & Sorooshian, S., 2007. An integrated hydrologic Bayesian multimodel combination framework: Confronting input, parameter, and model structural uncertainty in hydrologic prediction. *Water Resources Research*, 43(1), W01403.
- Aubert, A. H., Gascuel-Oudou, C. & Merot, P., 2013. Annual hysteresis of water quality: A method to analyse the effect of intraand inter-annual climatic conditions. *Journal of Hydrology*, 478, 29-39.
- Bergström, S. & Forsman, A., 1973. Development of a conceptual deterministic rainfall-runoff model. *Nordic Hydrology*, 4(3), 147-170.
- Beven, K., 1996. Equifinality and Uncertainty in Geomorphological Modelling. *The Scientific Nature of Geomorphology*, 289-313.
- Beven, K., 2006. Searching for the Holy Grail of Scientific Hydrology: $Q_t=H(SR)A$ as closure. *Hydrology and Earth System Sciences*, 10, 609-618.
- Beven, K. & Binley, A., 1992. The future of distributed models: Model calibration and uncertainty prediction. *Hydrological Processes*, 6(3), 279-298.
- Beven, K. J., 2011. Rainfall-runoff modelling: the primer (2nd ed.). Chichester, West Sussex UK: John Wiley & Sons.
- Brauer, C. C., 2014. Modelling rainfall-runoff processes in lowland catchments (doctoral dissertation). Wageningen University. Retrieved from <https://library.wur.nl/WebQuery/wurpubs/452940>
- Brauer, C., Torfs, P., Teuling, A. & Uijlenhoet, R., 2014. The Wageningen Lowland Runoff Simulator (WALRUS): application to the hupsel brook catchment and the cabauw polder. *Hydrology and Earth System Sciences*, 18(10), 4007-4028.
- Butts, M. B., Payne, J. T., Kristensen, M. & Madsen, H., 2004. An evaluation of the impact of model structure on hydrological modelling uncertainty for streamflow simulation. *Journal of Hydrology*, 298(1-4), 242-266.
- Clark, C., 1945. Storage and the unit hydrograph.. *Transactions of the American Society of Civil Engineers*, 110, 1416-1446.
- Clark, M. P. et al., 2008. Framework for Understanding Structural Errors (FUSE): A modular framework to diagnose differences between hydrological models. *Water Resources Research*, 44, W00B02.
- Euser, T. et al., 2013. A framework to assess the realism of model structures using hydrological signatures. *Hydrology and Earth System Sciences*, 17, 1893-1912.
- Fenicia, F., Savenije, H. H. G., Matgen, P. & Pfister, L., 2006. Is the groundwater reservoir linear? Learning from data in hydrological modelling. *Hydrology and Earth System Sciences*, 10(1), 139-150.
- Fleming, P., 1970. A diurnal distribution function for daily evaporation. *Water Resources Research*, 6(3), 937-942.
- Freer, J., McMillan, H., McDonnell, J. & Beven, K., 2004. Constraining dynamic TOPMODEL responses for imprecise water table information using fuzzy rule based performance measures. *Journal of Hydrology*, 291(3), 254-277.
- Gharari, S., 2016. On the role of model structure in hydrological modeling: Understanding models (doctoral dissertation). TU Delft. Retrieved from <https://repository.tudelft.nl/islandora/object/uuid:055795fb-611e-4e04-b431-fd0c377581f1>

- Gharari, S. et al., 2014. Using expert-knowledge to increase realism in environmental system models can dramatically reduce the need for calibration. *Hydrology and Earth System Sciences*, 18, 4839-4859.
- González-Castro, J. A. & Muste, M., 2007. Framework for Estimating Uncertainty of ADCP Measurements from a Moving Boat by Standardized Uncertainty Analysis. *Journal of Hydraulic Engineering*, 133(12), 1390-1410.
- Gupta, H. V. et al., 2012. Towards a comprehensive assessment of model structural adequacy. *Water Resources Research*, 48(8), W08301.
- Gupta, H. V., Sorooshian, S. & Yapo, P. O., 1998. Toward improved calibration of hydrologic models: Multiple and noncommensurable measures of information. *Water Resources Research*, 34(4), 751-763.
- Hrachowitz, M. & Clark, M. P., 2017. HESS Opinions: The complementary merits of competing modelling philosophies in hydrology. *Hydrology and Earth System Sciences*, 21(8), 3953 - 3973.
- Hrachowitz, M. et al., 2014. Process consistency in models: The importance of system signatures, expert knowledge, and process complexity. *Water resources research*, 50(9) 7445-7469.
- Kavetski, D. & Fenicia, F., 2011. Elements of a flexible approach for conceptual hydrological modeling: 2. Application and experimental insights. *Water Resources Research*, 47(11), W11511.
- Kavetski, D., Kuczera, G. & Franks, S. W., 2006. Bayesian analysis of input uncertainty in hydrological modeling: 1. Theory. *Water Resources Research*, 42(3), W03407.
- Lamb, R. & Beven, K., 1997. Using interactive recession curve analysis to specify a general catchment storage model. *Hydrology and Earth System Sciences*, 1(1), 101-113.
- McMillan, H. K. et al., 2011. Hydrological field data from a modeller's perspective: Part 1. Diagnostic tests for model structure. *Hydrological Processes*, 25(4), 511-522.
- Penman, H. L., 1948. Natural evaporation from open water, bare soil and grass. *Proceedings of the Royal Society of London*, 193(1032), 120-145.
- Refsgaard, J. C. & Storm, B., 1995. Mike she. *Computer Models of Watershed Hydrology*, 1, 809-846.
- Renard, B. et al., 2010. Understanding predictive uncertainty in hydrologic modeling: The challenge of identifying input and structural errors. *Water Resources Research*, 46(5), W05521.
- Savenije, H. H. G., 2009. CT4431 - hydrological modelling. Delft, The Netherlands: TU Delft
- Savenije, H. H. G., 2009. HESS Opinions "The art of hydrology". *Hydrology and Earth System Sciences*, 13(2), 157-161.
- Seibert, J., 1997. Estimation of Parameter Uncertainty. *Nordic Hydrology*, 28, 247-262.
- Willems, P., 2009. A time series tool to support the multi-criteria performance evaluation of rainfall-runoff models. *Environmental Modelling & Software*, 24, 311-321.
- Willems, P., 2014. Parsimonious rainfall-runoff model construction supported by time series processing and validation of hydrological extremes – Part 1: Step-wise model-structure identification and calibration approach. *Journal of Hydrology*, 510, 578-590.
- Willems, P. et al., 2014. Parsimonious rainfall-runoff model construction supported by time series processing and validation of hydrological extremes--Part 2: Intercomparison of models and calibration approaches. *Journal of Hydrology*, 510, 591-609.
- Young, P., 2003. Top-down and data-based mechanistic modelling of rainfall-flow dynamics at the catchment scale. *Hydrological processes*, 17(11), 2195-2217.

Appendix 1 Model descriptions

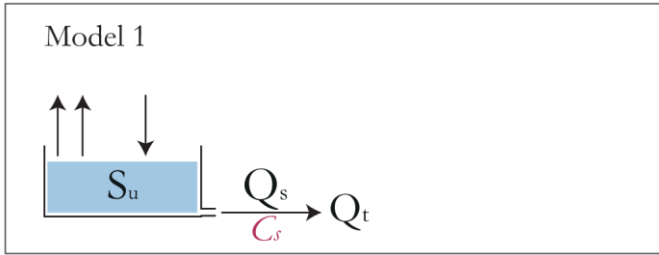
Model 1

This model exists of one linear reservoir (unsaturated reservoir S_u) with only one parameter. The storage change of S_u is formulated in eq. (30). Precipitation P will fill up the reservoir. The predicted evaporation ($E+T$) is at potential rate when there is no rain and sufficient storage eq. (31). The outflow of this reservoir is conceptualized as linear eq. (32) depending on the storage level S_u and on the time-scale of groundwater drainage parameter C_s . This model is considered to be the most simple model.

$$\frac{dS_u}{dt} = P(t) - E(t) - T(t) - Q_u(t) \quad (30)$$

$$\text{If } P = 0: E(t) + T(t) = \min(E_p, S_u(t)) \quad (31)$$

$$Q_u = \frac{1}{C_s} * S_u(t) \quad (32)$$



Model 2

Model 2 exists of one linear reservoir (unsaturated reservoir S_u) with two parameters. The storage change of S_u is formulated in eq. (33). Precipitation P will fill up the reservoir until the maximum storage capacity $S_{u,max}$ is reached. The precipitation that exceeds the storage capacity will form the quick flow Q_f eq. (35). The predicted evaporation ($E+T$) is assumed to be proportional to the potential rate when there is no rain and sufficient storage eq. (34). The slow flow Q_s of this reservoir is conceptualized as linear eq. (36) depending on the storage level S_u and on the time-scale of groundwater drainage parameter C_s .

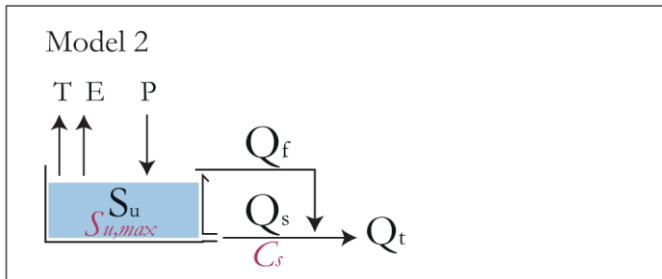
$$\frac{dS_u}{dt} = P(t) - E(t) - Q_s(t) - Q_f(t) \quad (33)$$

$$\text{If } P=0: E(t) + T(t) = \min(E_p \frac{S_u}{S_{u,max}}, S_u) \quad (34)$$

$$Q_f = \max(0, S_u - S_{u,max}) \quad (35)$$

$$Q_s = \frac{1}{C_s} * S_u(t) \quad (36)$$

$$Q_t = Q_s + Q_f \quad (37)$$



Model 3

The third model exists of two reservoirs (unsaturated reservoir S_u and fast runoff reservoir S_f) with two parameters. The storage change of S_u is formulated in eq. (38). Precipitation P will fill up the reservoir until the maximum storage capacity $S_{u,max}$ is reached. The precipitation that exceeds the storage capacity will form Q_{uf} eq. (41). The predicted evaporation ($E+T$) is assumed to be proportional to the potential rate when there is no rain and sufficient storage eq. (40). The storage change of S_f is formulated in eq. (39). The reservoir receives Q_{uf} . The fast flow Q_f eq. (42) is conceptualized as linear to the storage level of S_f and depends on the time-scale of groundwater drainage parameter C_f .

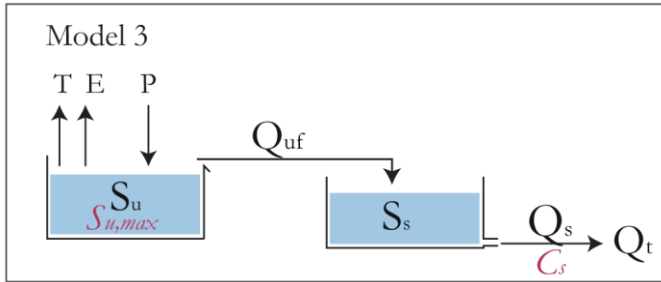
$$\frac{dS_u}{dt} = P(t) - E(t) - Q_{uf}(t) \quad (38)$$

$$\frac{dS_f}{dt} = Q_{uf}(t) - Q_f(t) \quad (39)$$

$$\text{If } P = 0: E(t) + T(t) = \min(E_p \frac{S_u}{S_{u,max}}, S_u) \quad (40)$$

$$Q_{uf} = \max(0, S_u - S_{u,max}) \quad (41)$$

$$Q_f = \frac{1}{C_f} * S_f(t) \quad (42)$$



Model 4

The fourth model exists of two reservoirs (unsaturated reservoir S_u and slow runoff reservoir S_s) with three parameters. The storage change of S_u is formulated in eq. (43). Precipitation P will fill up the reservoir until the maximum storage capacity $S_{u,max}$ is reached. The precipitation that exceeds the storage capacity will form Q_{uf} eq.(46). The predicted evaporation ($E+T$) is assumed to be proportional to the potential rate when there is no rain and sufficient storage eq. (45). The storage change of S_s is formulated in eq. (44). The reservoir receives R_s (47). The slow flow Q_s eq. (49) is conceptualized as linear to the storage level of S_s and depends on the time-scale of groundwater drainage parameter C_s . The fast flow is conceptualized in eq. (48). The total flow (50) is the sum of Q_f and Q_s .

$$\frac{dS_u}{dt} = P(t) - E(t) - Q_{uf}(t) \quad (43)$$

$$\frac{dS_s}{dt} = R_s(t) - Q_f(t) \quad (44)$$

$$\text{If } P = 0: E(t) + T(t) = \min(E_p \frac{S_u}{S_{u,max}}, S_u) \quad (45)$$

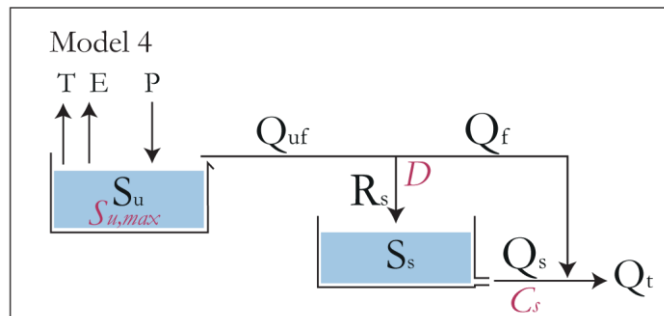
$$Q_{uf} = \max(0, S_u - S_{u,max}) \quad (46)$$

$$R_s = Q_{uf} * D \quad (47)$$

$$Q_f = Q_{uf} * (1 - D) \quad (48)$$

$$Q_s = \frac{1}{C_s} * S_s(t) \quad (49)$$

$$Q_t = Q_s + Q_f \quad (50)$$



Model 6

The sixth model exists of two reservoirs (unsaturated reservoir S_u and slow runoff reservoir S_s) with two parameters. The storage change of S_u is formulated in eq. (51). Precipitation P will fill up the reservoir. The predicted evaporation ($E+T$) is assumed to be at potential rate when there is no rain and sufficient storage eq. (53). The percolation rate $Perc$ is assumed to be constant P_{max} when sufficient storage is present eq. (54). The storage change of S_s is formulated in eq. (52). The reservoir receives $Perc$ and the slow outflow Q_s , eq. (57) is conceptualized as linear to the storage level of S_s , eq. (55), (56) and depends on the time-scale of groundwater drainage parameter C_s .

$$\frac{dS_u}{dt} = P(t) - E(t) - Perc(t) \quad (51)$$

$$\frac{dS_s}{dt} = Perc(t) - Q_s(t) \quad (52)$$

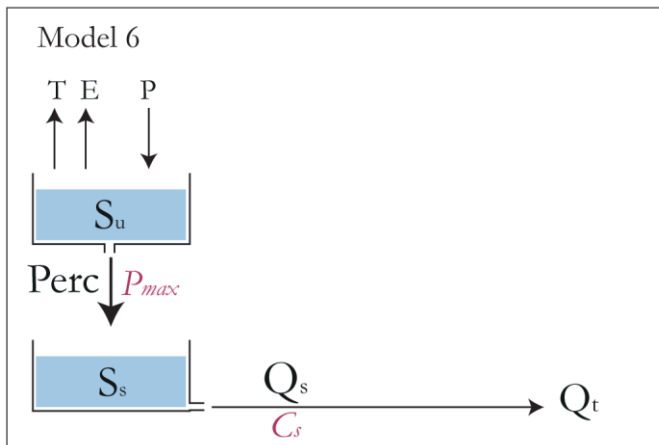
$$\text{If } P = 0: E(t) + T(t) = \min(E_p \frac{S_u}{S_{u\max}}, S_u) \quad (53)$$

$$Perc = \min(P_{max}, S_u) \quad (54)$$

$$Q_{perc} = \frac{1}{C_s} * Perc(t) \quad (55)$$

$$Q_{storage} = \frac{1}{C_s} * S_s(t) \quad (56)$$

$$Q_s = Q_{storage} + Q_{perc} \quad (57)$$



Model 7

The seventh model exists of two reservoirs (unsaturated reservoir S_u and slow runoff reservoir S_s) with three parameters. The storage change of S_u is formulated in eq. (58). Precipitation P will fill up the reservoir until the maximum storage capacity $S_{u,max}$ is reached. The precipitation that exceeds the storage capacity will form the quick flow Q_f eq. (61). The predicted evaporation ($E+T$) is assumed to be proportional to the potential rate when there is no rain and sufficient storage eq. (60). The percolation rate $Perc$ is assumed to be proportional to P_{max} when sufficient storage is present eq.(62). The storage change of S_s is formulated in eq. (59). The reservoir receives $Perc$ and the slow outflow Q_s eq. (65) and is conceptualized as linear to the storage level of S_s , eq.(63), (64) and depends on the time-scale of groundwater drainage parameter C_s . The total flow (65) is the sum of Q_f and Q_s .

$$\frac{dS_u}{dt} = P(t) - E(t) - Perc(t) \quad (58)$$

$$\frac{dS_s}{dt} = Perc(t) - Q_s(t) \quad (59)$$

$$\text{If } P = 0: E(t) + T(t) = \min(E_p \frac{S_u}{S_{u,max}}, S_u) \quad (60)$$

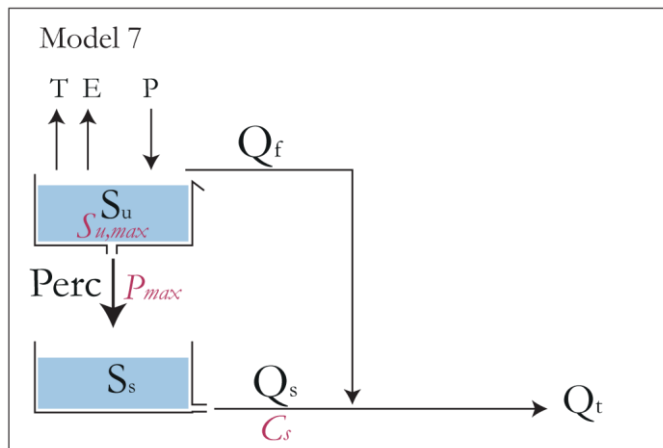
$$Q_f = \max(0, S_u - S_{u,max}) \quad (61)$$

$$Perc = \min(P_{max} \frac{S_u}{S_{u,max}}, S_u) \quad (62)$$

$$Q_{perc} = \frac{1}{C_s} * Perc(t) \quad (63)$$

$$Q_{storage} = \frac{1}{C_s} * S_s(t) \quad (64)$$

$$Q_s = Q_{storage} + Q_{perc} \quad (65)$$



Model 9

The ninth model exists of two reservoirs (unsaturated reservoir S_u and slow runoff reservoir S_s) with four parameters. The storage change of S_u is formulated in eq. (66). Precipitation P will fill up the reservoir until the maximum storage capacity $S_{u,max}$ is reached. The precipitation that exceeds the storage capacity will form the quick flow Q_{uf} eq. (69). The predicted evaporation ($E+T$) is assumed to be proportional to the potential rate when there is no rain and sufficient storage eq. (68). The percolation rate $Perc$ is assumed to be proportional to P_{max} when sufficient storage is present eq. (70). The storage change of S_s is formulated in eq. (67). The reservoir receives $Perc$ and R_s (71). The slow outflow Q_s eq. (75) and is conceptualized as linear to the storage level of S_s eq. (72) (74) and depends on the time-scale of groundwater drainage parameter C_s . The fast flow is conceptualized in eq. (73). The total flow (76) is the sum of Q_f and Q_s .

$$\frac{dS_u}{dt} = P(t) - E(t) - Perc(t) - Q_{uf}(t) \quad (66)$$

$$\frac{dS_s}{dt} = Perc(t) - Q_s(t) \quad (67)$$

$$\text{If } P = 0: E(t) + T(t) = \min(E_p \frac{S_u}{S_{u,max}}, S_u) \quad (68)$$

$$Q_{uf} = \max(0, S_u - S_{u,max}) \quad (69)$$

$$Perc = \min\left(P_{max} \frac{S_u}{S_{u,max}}, S_u\right) \quad (70)$$

$$R_s = Q_{uf} * D \quad (71)$$

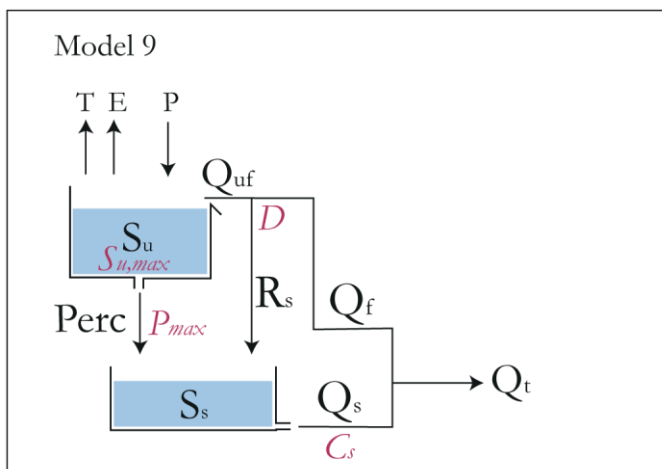
$$Q_{ptot} = \frac{1}{C_s} * Perc(t) + \frac{1}{C_s} * R_s \quad (72)$$

$$Q_f = Q_{uf} * (1 - D) \quad (73)$$

$$Q_{storage} = \frac{1}{C_s} * S_s(t) \quad (74)$$

$$Q_s = Q_{storage} + Q_{ptot} \quad (75)$$

$$Q_t = Q_s + Q_f \quad (76)$$



Appendix 2 Evaporation distribution function

In Table 24 the used sunrise and sunset hours are given for the Kervidy-Naizin catchment in France.

MONTH	SUNRISE	SUNSET
JANUARY	9:00	18:00
FEBRUARY	8:00	18:00
MARCH	8:00	19:00
APRIL	8:00	21:00
MAY	7:00	22:00
JUNE	6:00	22:00
JULY	6:00	22:00
AUGUST	7:00	22:00
SEPTEMBER	8:00	21:00
OCTOBER	8:00	20:00
NOVEMBER	8:00	18:00
DECEMBER	9:00	17:00

Table 24. Sunrise and sunset time of Naizin (loosely based on <https://www.aroundtheworld360.com/sunrise-sunset/france/naizin/>)

For the catchment only daily potential evaporation is available. However the precipitation data is in hourly resolution, the daily potential evaporation data is therefore transformed to hourly data using (Fleming, 1970).

For sunny days is the evaporation index (EVI) according:

$$EVI = 2.78 - (7.73 + A)^{0.5}$$

$$A = 31.4 D_{\text{HOUR}}^2 - 33.6 D_{\text{HOUR}} + 2.23$$

For cloudy days (days with measured precipitation) is the evaporation index (EVI) according:

$$EVI = 0 \text{ for } D_{\text{hour}} 0 - 0.06$$

$$EVI = 6.75(D_{\text{HOUR}} - 0.06) \text{ for } D_{\text{hour}} 0.061 - 0.26$$

$$EVI = 1.35 \text{ for } D_{\text{hour}} 0.261 - 0.8$$

$$EVI = 1.35 - 6.75(D_{\text{HOUR}} - 0.80) \text{ for } D_{\text{hour}} 0.81 - 1$$

Time index D_{HOUR} =

$$D_{\text{HOUR}} = \frac{\text{time from sunrise}}{\text{daylength}}$$

Average daylight evaporation rate =

$$\frac{\text{total 20 hour evaporation}}{\text{daylength}}$$

Evaporation rate (D_{HOUR}) =

$$EVI * \text{average daylight evaporation rate}$$

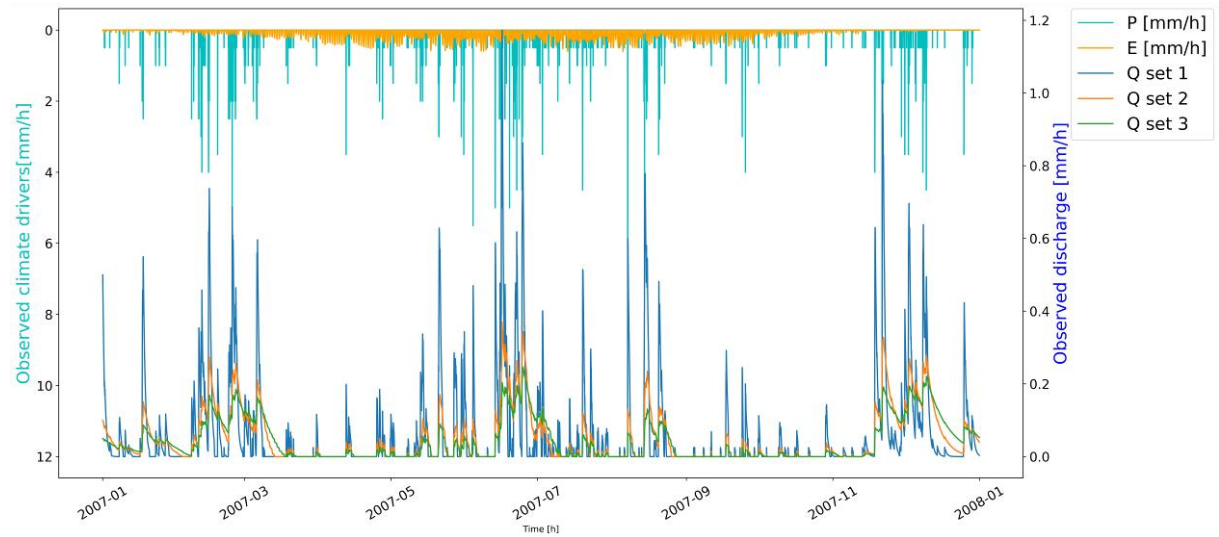
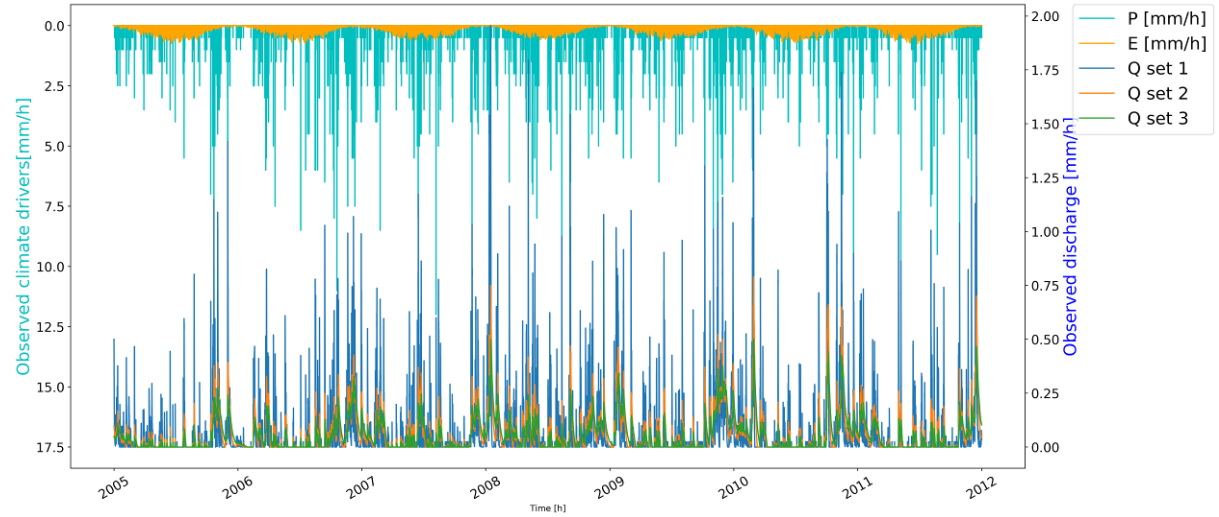
Appendix 3 Parameter values and discharge data synthetic experiment

Model 1

Parameter sets:

	<i>Set 1</i>	<i>Set 2</i>	<i>Set 3</i>
$C_s [h]$	20	100	200

Discharge series:

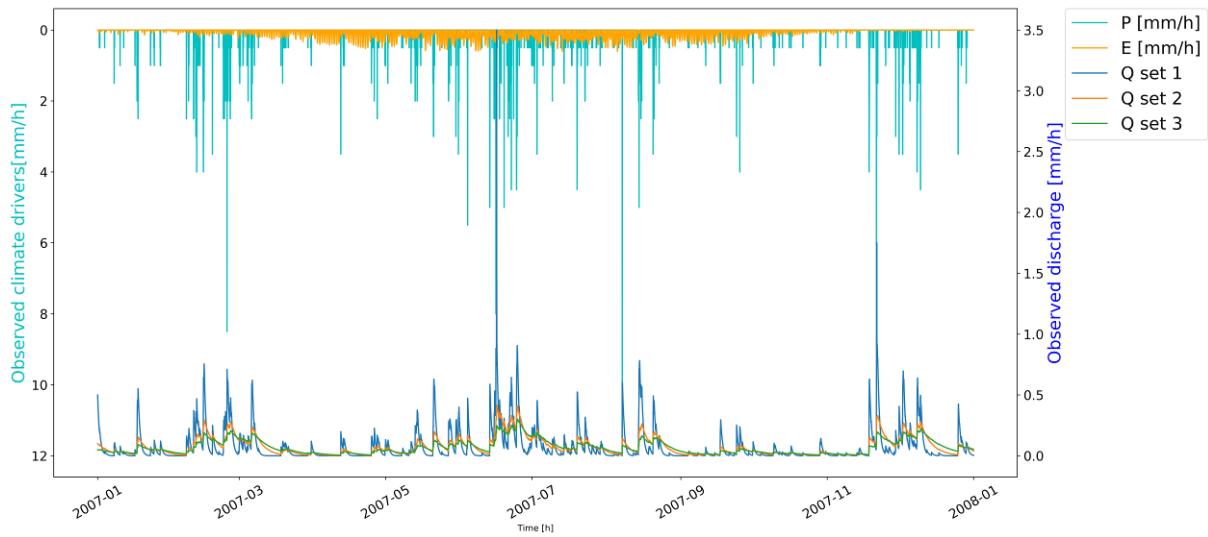
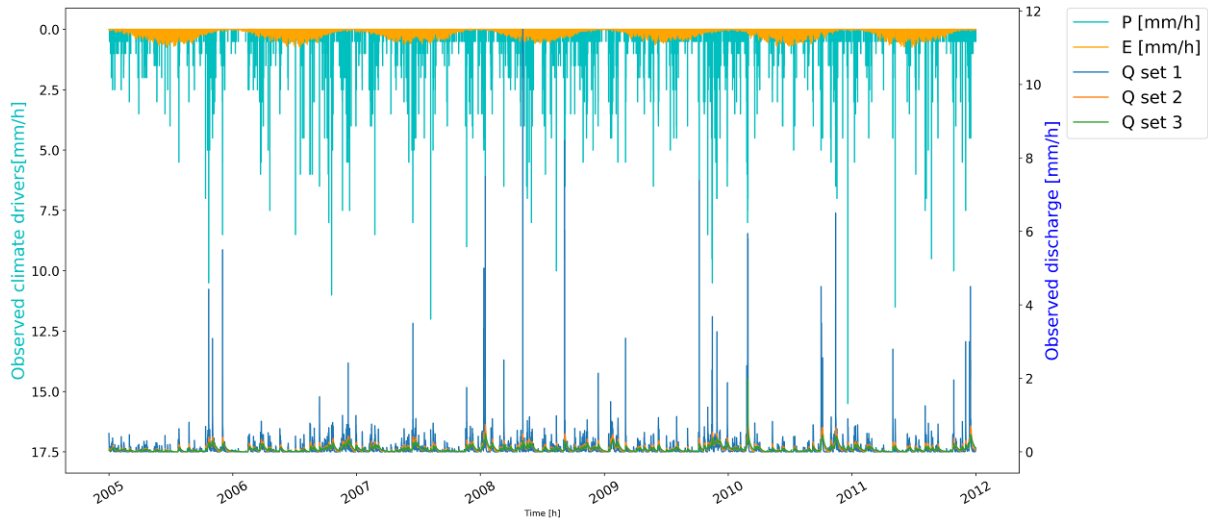


Model 2

Parameter sets:

	<i>Set 1</i>	<i>Set 2</i>	<i>Set 3</i>
$C_s [h]$	20	100	200
$S_{U,max} [mm]$	20	80	100

Discharge series:

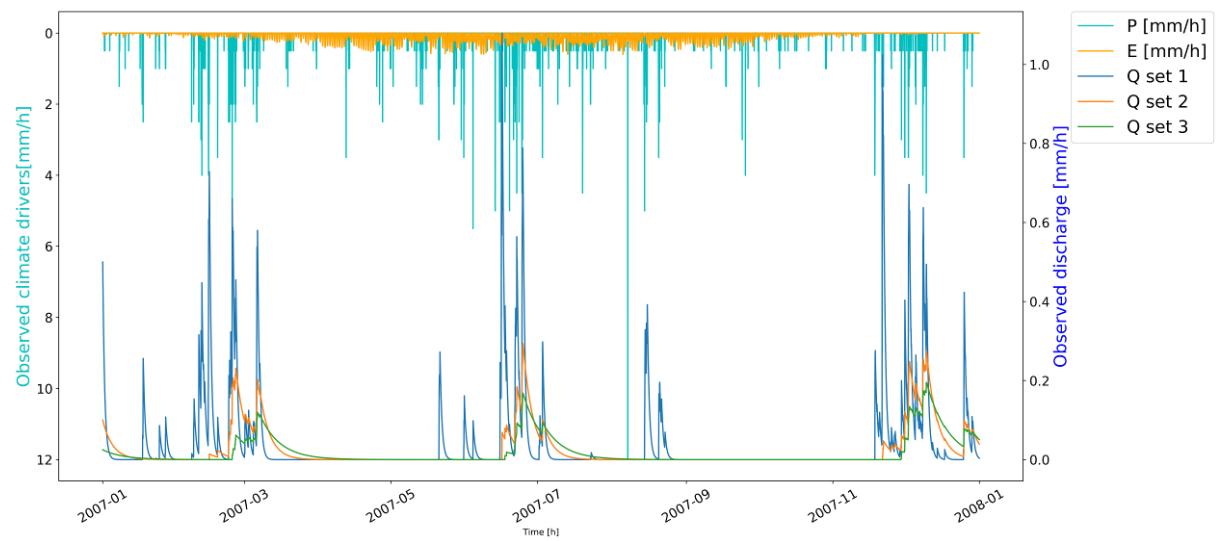
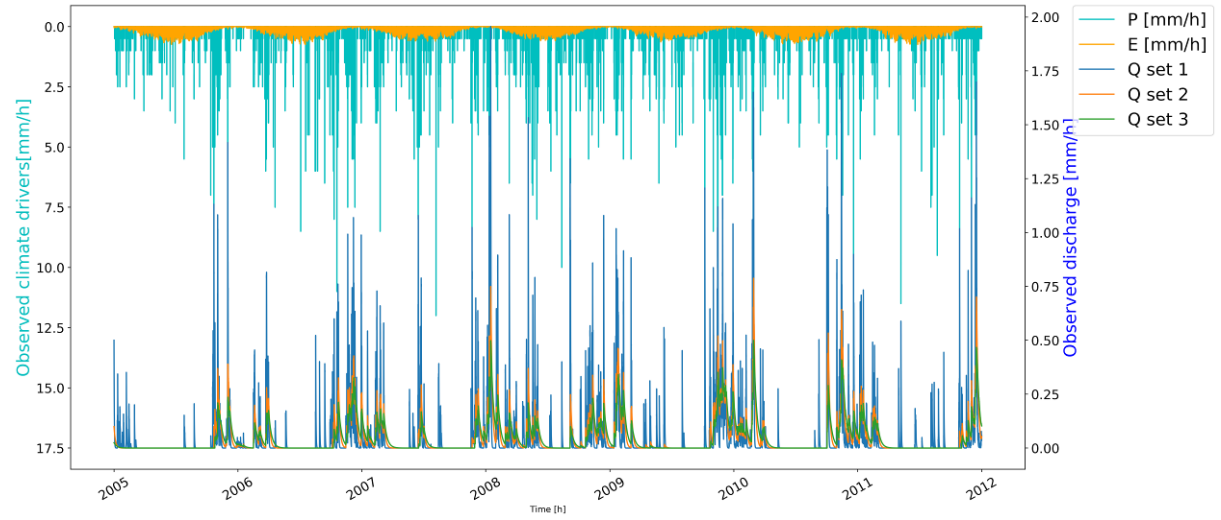


Model 3

Parameter sets:

	<i>Set 1</i>	<i>Set 2</i>	<i>Set 3</i>
$C_s [h]$	20	100	200
$S_{U,max} [mm]$	20	80	100

Discharge series:

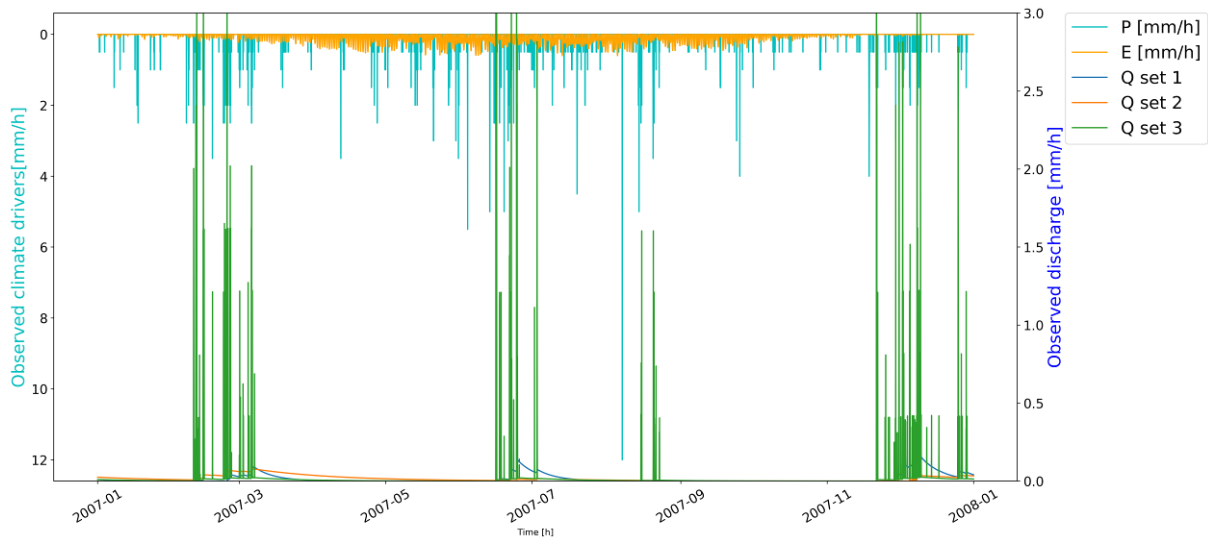
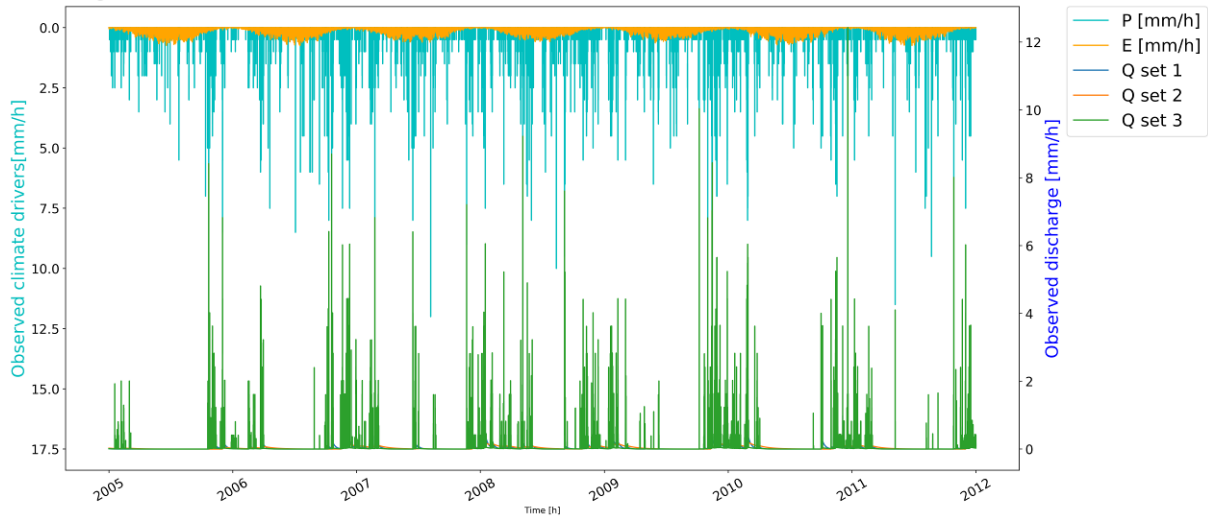


Model 4

Parameter sets:

	<i>Set 1</i>	<i>Set 2</i>	<i>Set 3</i>
$C_s [h]$	200	800	500
$S_{U,max} [mm]$	100	234	40
$D [-]$	0.8	0.96	0.2

Discharge series:

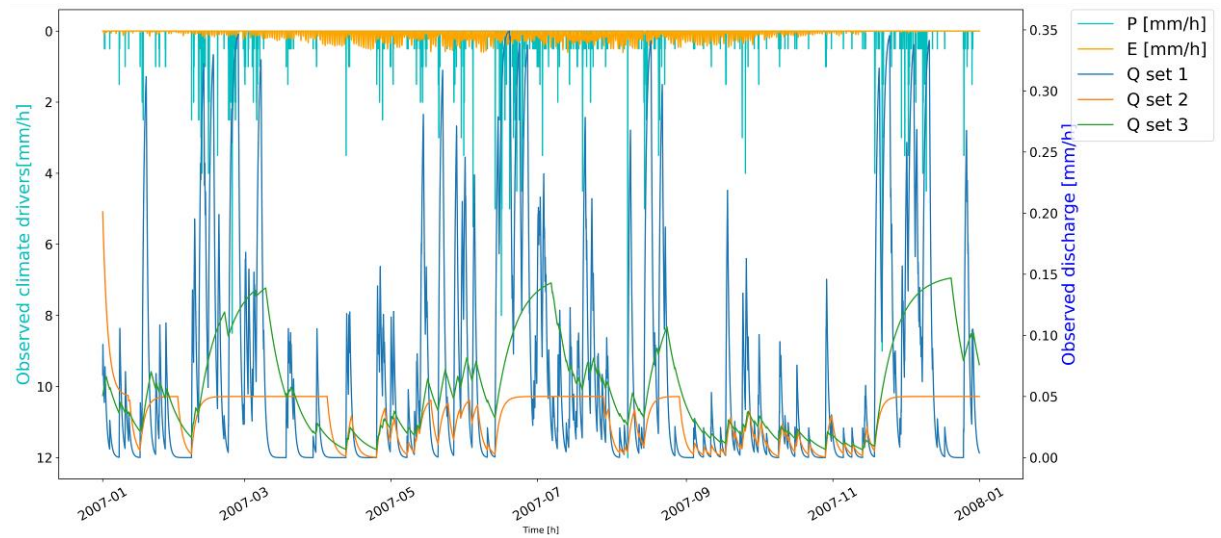
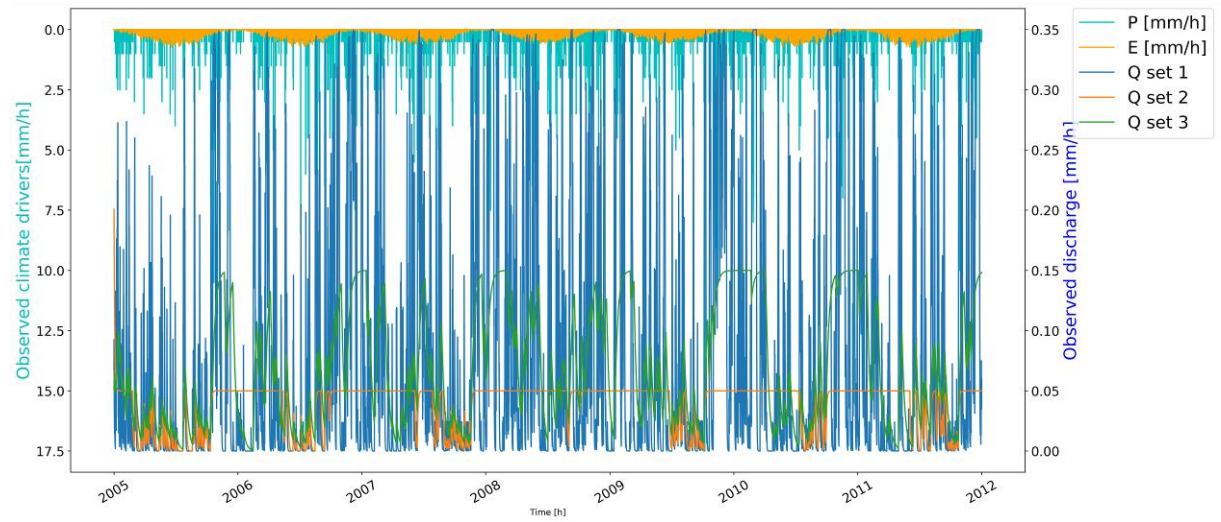


Model 6

Parameter sets:

	<i>Set 1</i>	<i>Set 2</i>	<i>Set 3</i>
$C_s [h]$	20	50	200
$P_{max} [mm]$	0.35	0.05	0.15

Discharge series:

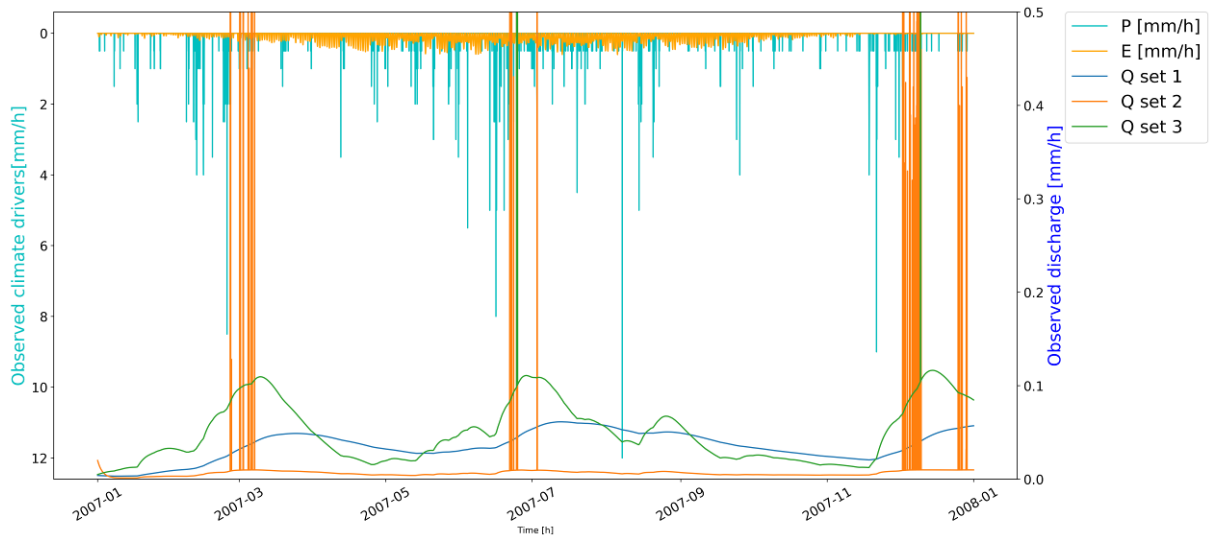
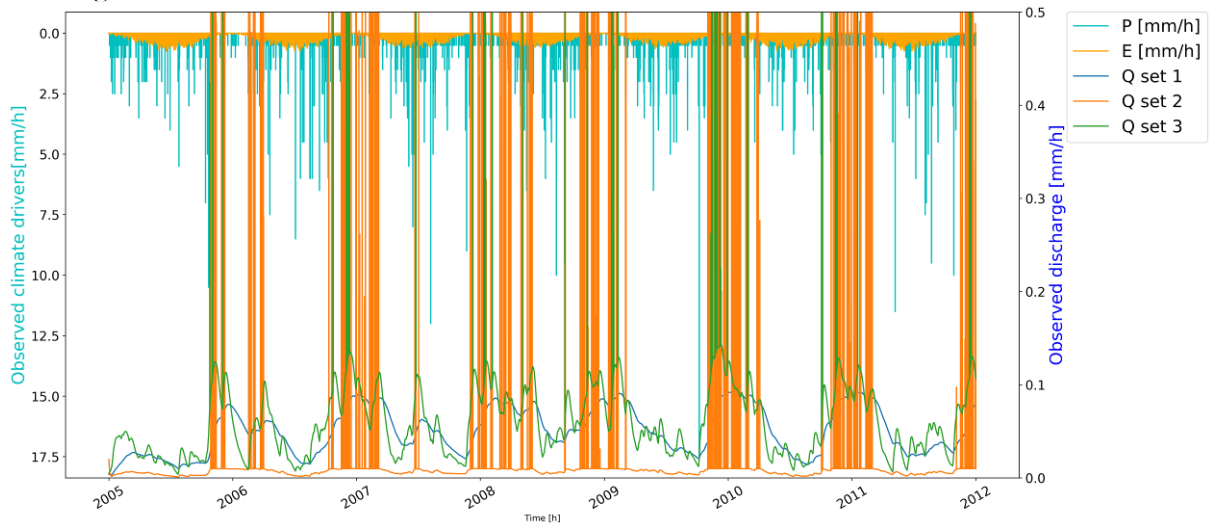


Model 7

Parameter sets:

	<i>Set 1</i>	<i>Set 2</i>	<i>Set 3</i>
$C_s [h]$	500	100	200
$S_{U,max} [mm]$	150	110	65
$P_{mac} [mm/h]$	0.1	0.07	0.15

Discharge series:

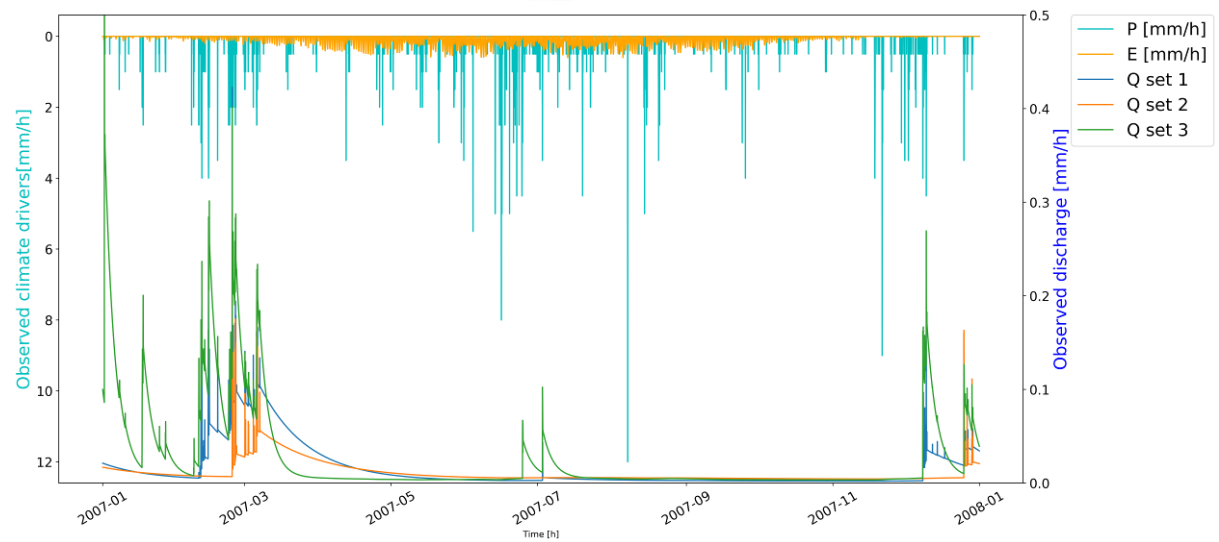
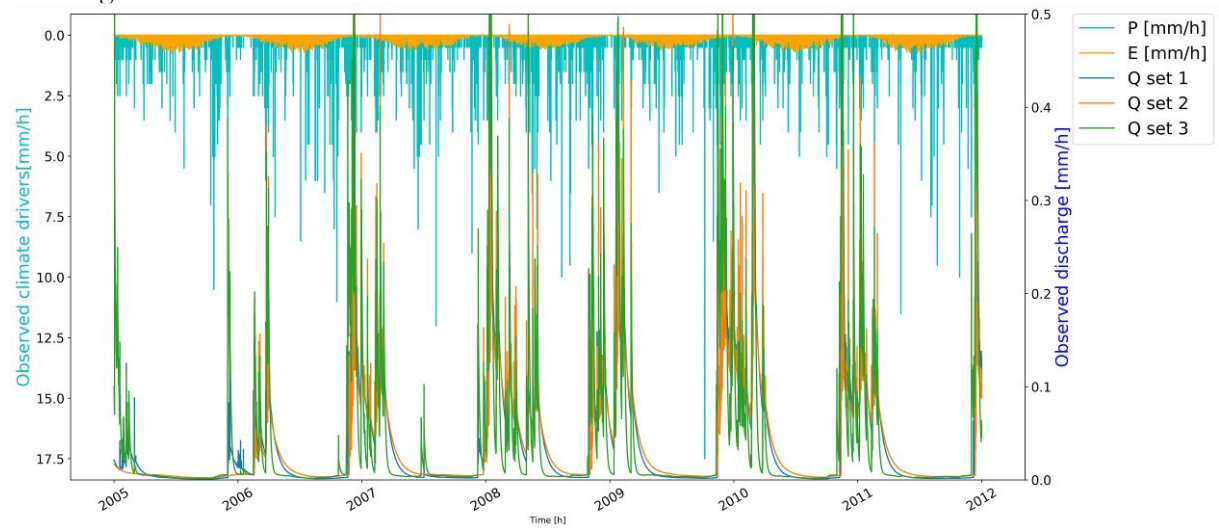


Model 9

Parameter sets:

	<i>Set 1</i>	<i>Set 2</i>	<i>Set 3</i>
$C_s [h]$	470	590	100
$S_{U,max} [mm]$	280	327	220
$P_{mac} [mm/h]$	0.0031	0.0065	0.0058
$D [-]$	0.96	0.94	0.977
<i>Calibration NS</i>	0.55	0.56	
<i>Calibration logNS</i>	0.73	0.69	
<i>Calibration R</i>	0.75	0.75	

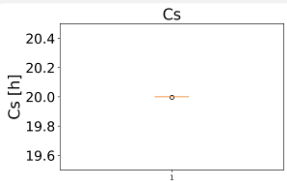
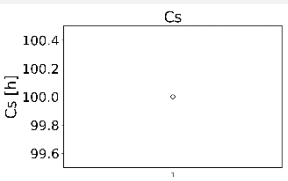
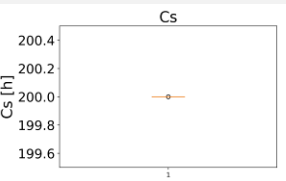
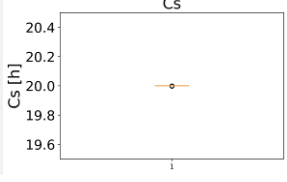
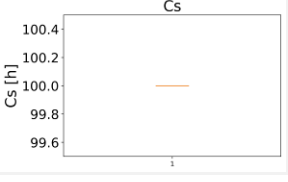
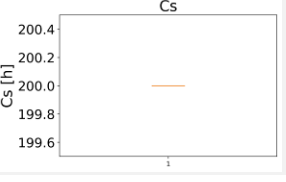
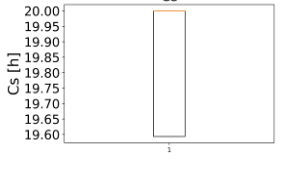
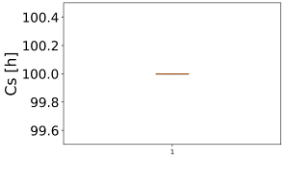
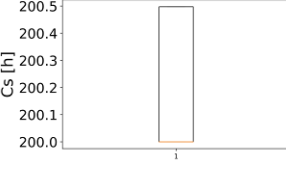
Discharge series:



Appendix 4 Method results

Some of the methods are depending on other parameter values as can be red in 3.1. In this part of the results are the other parameters needed to determine the current parameter of the method assumed to be known and set to the original value. So the input of the methods exists of correct data and correct parameters except for the one which is obtained by the method.

C_s fit method results

Model 1	<i>Data set 1</i>	<i>Data set 2</i>	<i>Data set 3</i>
<i>Original value [h]</i>	20	100	200
<i>Number of recessions used</i>	23	31	31
<i>Median [h]</i>	20	100	200
<i>Std [-]</i>	9.85e-4	4.13e-7	1.56e-6
<i>Relative error [%]</i>	0	0	0
<i>Distribution</i>			
Model 2			
<i>Original value [h]</i>	20	100	200
<i>Number of recessions used</i>	31	31	31
<i>Median [h]</i>	20	100	200
<i>Std [-]</i>	0.099	3.55e-7	4.43e-8
<i>Relative error [%]</i>	0		
<i>Distribution</i>			
Model 3			
<i>Original value [h]</i>	20	100	200
<i>Number of recessions used</i>	206	201	260
<i>Median [h]</i>	20	100	200
<i>Std [-]</i>	0.188	0.001	0.23
<i>Relative error [%]</i>	0	0	0
<i>Distribution</i>			

Model 4

<i>Original value [h]</i>	200	500	800
<i>Number of recessions used</i>	42	66	64
<i>Median [h]</i>	200	800	500
<i>Std [-]</i>	0.0038	0.0166	0.24
<i>Relative error [%]</i>	0	0	0
<i>Distribution</i>			

Model 6

	Data set 1	Data set 2	Data set 3
<i>Original value [h]</i>	20	50	200
<i>Number of recessions used</i>	43	30	33
<i>Median [h]</i>	20	50	200
<i>Std [-]</i>	0.1	72	2.47e-8
<i>Relative error [%]</i>	0	0	0
<i>Distribution</i>			

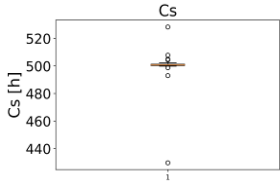
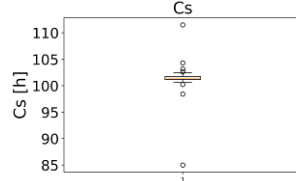
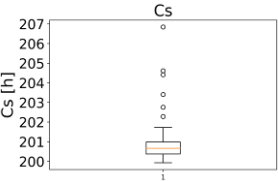
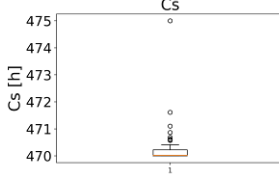
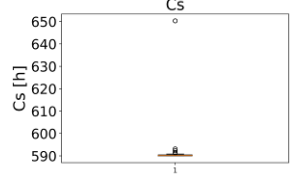
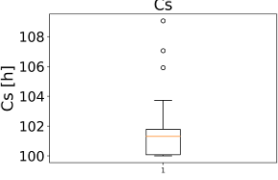
Model 7

<i>Original value [h]</i>	500	100	200
<i>Number of recessions used</i>	40	31	41
<i>Median [h]</i>	1977	613	499
<i>Std [-]</i>	2004	474	415
<i>Relative error [%]</i>	295	513	149
<i>Distribution</i>			

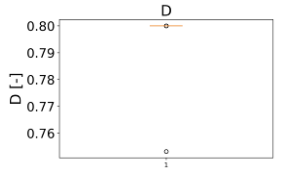
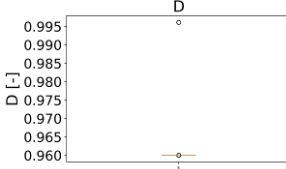
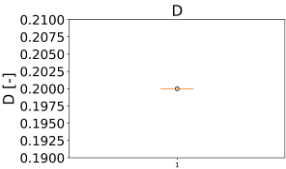
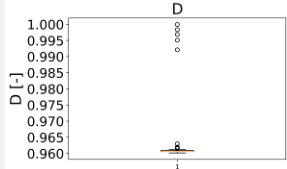
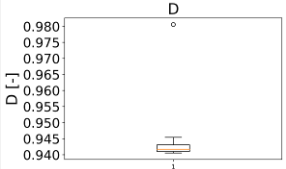
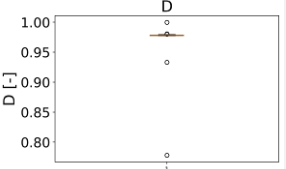
Model 9

<i>Original value [h]</i>	470	590	100
<i>Number of recessions used</i>	34	31	44
<i>Median [h]</i>	620.5	894	324
<i>Std [-]</i>	1274	1488	1016
<i>Relative error [%]</i>	32.02	51.6	224.9
<i>Distribution</i>			

Cs S-D method

Model 7	<i>Data set 1</i>	<i>Data set 2</i>	<i>Data set 3</i>
<i>Original value [h]</i>	500	100	200
<i>First estimate C_s [h]</i>	1977	613	499
<i>Median [h]</i>	500.64	101.42	200.66
<i>Std [-]</i>	13	3.19	1.24
<i>Relative error [%]</i>	0.13	1.42	0.33
<i>Distribution</i>			
Model 9			
<i>Original value [h]</i>	470	590	100
<i>First estimate C_s [h]</i>	620	894	324
<i>Number of recessions used</i>	81	46	51
<i>Median [h]</i>	470	590.1	101.3
<i>Std [-]</i>	1.12	8.77	1.74
<i>Relative error [%]</i>	0	0.02	1.3
<i>Distribution</i>			

D method

Model 4	<i>Data set 1</i>	<i>Data set 2</i>	<i>Data set 3</i>
<i>Original value</i>	0.8	0.96	0.2
<i>Number of peak flow used</i>	103	60	452
<i>Median [-]</i>	0.8	0.96	0.2
<i>Std [-]</i>	0.00459	0.00458	2.65e-14
<i>Relative error [%]</i>	0	0	0
<i>Distribution</i>			
Model 9			
<i>Original value</i>	0.96	0.94	0.977
<i>Number of peak flow used</i>	68	35	33
<i>Median [-]</i>	0.961	0.941	0.9977
<i>Std [-]</i>	0.009	0.006	0.035
<i>Relative error [%]</i>	0.06	0.18	0.07
<i>Distribution</i>			

SU_{,max} bounds method

Model 3

<i>Original value [mm]</i>	20	80	100
<i>Minimum [mm]</i>	18	60	63
<i>Relative error %</i>	-10	-25	-37
<i>Maximum [mm]</i>	31	136	183
<i>Relative error %</i>	55	70	83

Model 4

<i>Original value [mm]</i>	100	235	40
<i>Minimum [mm]</i>	64	153	32
<i>Relative error %</i>	-36	-35	-20
<i>Maximum [mm]</i>	183	288	60
<i>Relative error %</i>	83	23	50

Model 7

<i>Original value [mm]</i>	150	110	65
<i>Minimum [mm]</i>	98	71	44
<i>Relative error %</i>	-35	-35	-32
<i>Maximum [mm]</i>	207	171	86
<i>Relative error %</i>	38	55	32

Model 9

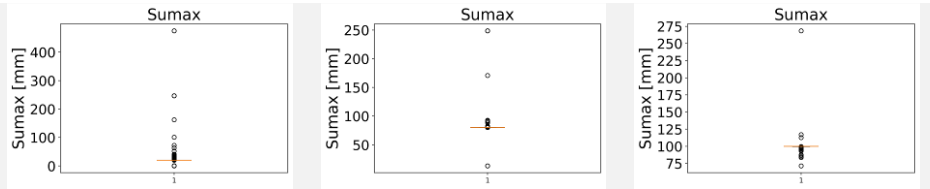
<i>Original value [mm]</i>	280	327	220
<i>Minimum [mm]</i>	178	206	146
<i>Relative error %</i>	-36	-37	-34
<i>Maximum [mm]</i>	309	466	276
<i>Relative error %</i>	10	43	25

SU_{max} inter-peak method

Model 3

<i>Original value [mm]</i>	20	80	100
<i>Number calculations</i>	460	364	344
<i>Median [mm]</i>	20	80	99.99
<i>Std [-]</i>	42	22.8	9.5

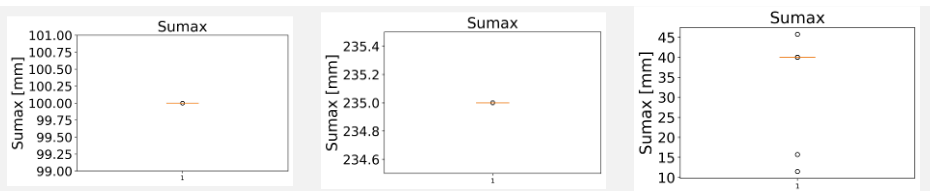
Distribution



Model 4

<i>Original value [mm]</i>	100	235	40
<i>Number calculations</i>	43	21	185
<i>Median [mm]</i>	100	235	40
<i>Std [-]</i>	6.5e-8	3.0e-7	2.77

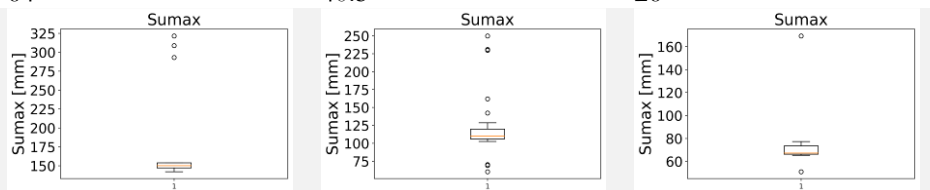
Distribution



Model 7

<i>Original value [mm]</i>	150	110	65
<i>Number calculations</i>	15	42	14
<i>Median [mm]</i>	150	110	67
<i>Std [-]</i>	64	40.5	26

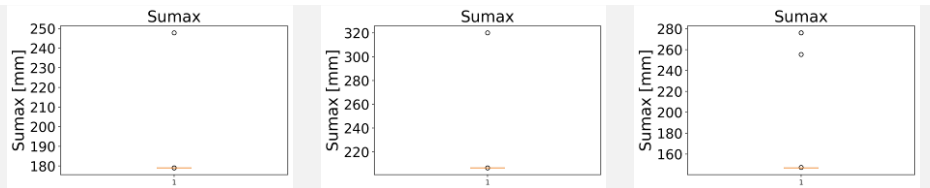
Distribution



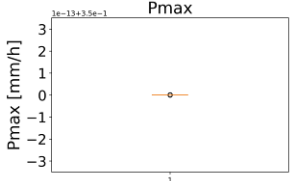
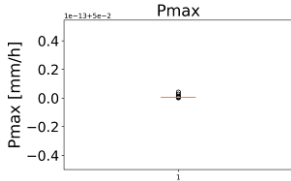
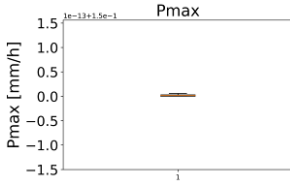
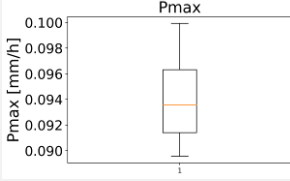
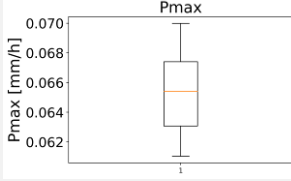
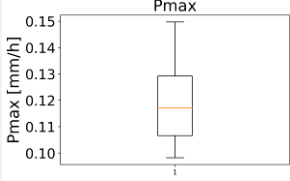
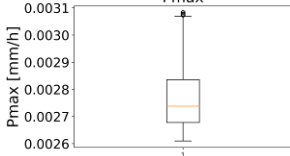
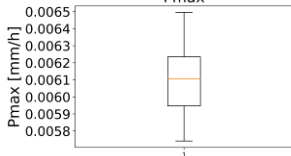
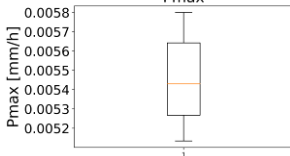
Model 9

<i>Original value [mm]</i>	280	327	220
<i>Number calculations</i>	28	10	27
<i>Median [mm]</i>	179	206	146
<i>Std [-]</i>	12	34	44

Distribution



P_{max} method

Model 6	<i>Data set 1</i>	<i>Data set 2</i>	<i>Data set 3</i>
<i>Original value [mm/h]</i>	0.35	0.05	0.15
<i>75 percentile [mm/h]</i>	0.35	0.05	0.15
<i>Relative error [%]</i>	0	0	0
<i>Min. distribution</i>	0.35	0.05	0.15
<i>Max. distribution</i>	0.35	0.05	0.15
<i>Distribution</i>			
Model 7			
<i>Original value [mm/h]</i>	0.1	0.07	0.15
<i>75 percentile [mm/h]</i>	0.096	0.067	0.129
<i>Relative error [%]</i>	-3.69	-0.70	-13.85
<i>Min. distribution</i>	0.089	0.061	0.098
<i>Max. distribution</i>	0.1	0.07	0.15
<i>Distribution</i>			
Model 9			
<i>Original value [mm/h]</i>	0.0031	0.0065	0.0058
<i>75 percentile [mm/h]</i>	0.0028	0.0062	0.0056
<i>Relative error [%]</i>	-9.68	-4.62	-3.45
<i>Min. distribution</i>	0.0026	0.0057	0.0051
<i>Max. distribution</i>	0.00307	0.0065	0.0058
<i>Distribution</i>			

Appendix 5 Results synthetic experiments

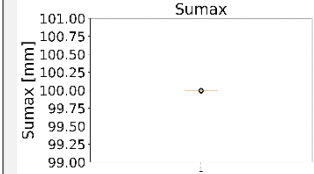
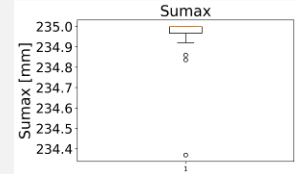
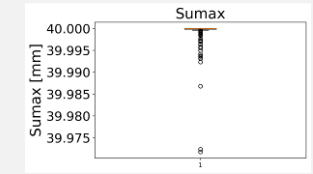
Some methods are depending on other parameter values to determine the current parameter. However, the methods contain errors in the parameter estimation. In this paragraph it is possible to see to what extent these errors in the determination of one parameter propagate to other parameters estimations.

Model 4

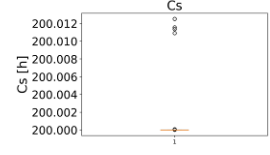
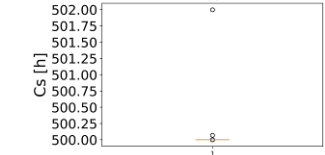
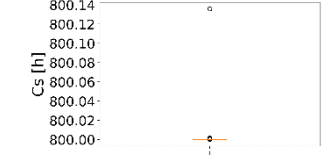
$S_{U,max}$ bounds method

	Data set 1	Data set 2	Data set 3
Original value [mm]	100	235	40
Minimum [mm]	64	153	32
Relative error	-36	-35	-20
Maximum [mm]	183	288	60
Relative error	83	23	50

$S_{U,max}$ inter-peak method

	Data set 1	Data set 2	Data set 3
Original value [mm]	100	235	40
Number calculations	43	21	182
Median [mm]	100	235	40
Std [-]	0.0024	0.13	0.0033
Distribution			

C_s fit method

	Data set 1	Data set 2	Data set 3
Original value [h]	200	500	800
Number of recessions used	42	66	64
Median [h]	200	800	500
Std [-]	0.0038	0.0166	0.24
Relative error [%]	0	0	0
Distribution			

D method

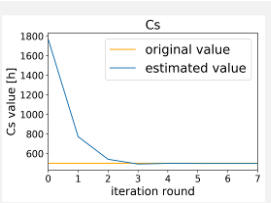
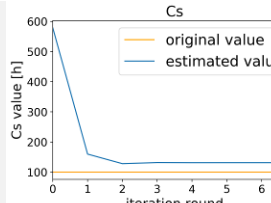
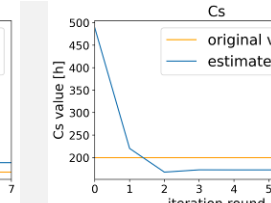
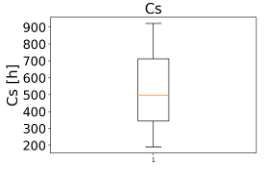
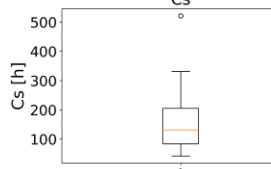
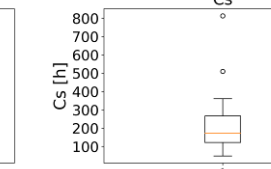
	<i>Data set 1</i>	<i>Data set 2</i>	<i>Data set 3</i>
<i>Original value</i>	0.8	0.96	0.2
<i>Number of peak flow used</i>	103	61	452
<i>Median [-]</i>	0.8	0.96	0.2
<i>Std [-]</i>	0.00459	0.00458	4.44e-9
<i>Relative error [%]</i>	0	0	0
<i>Distribution</i>			

Model 7

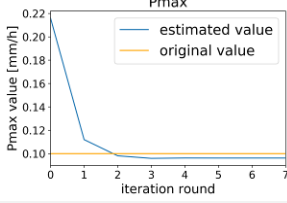
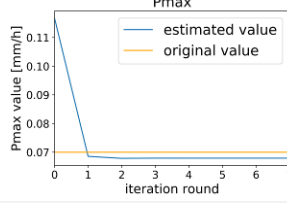
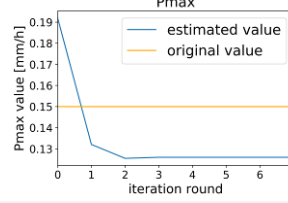
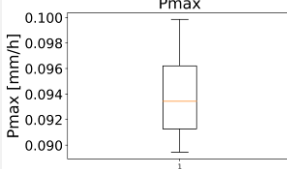
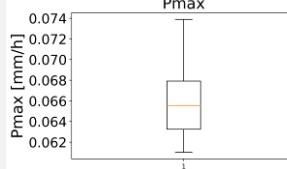
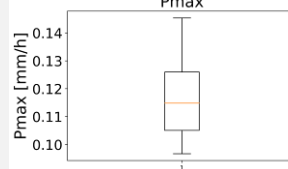
$S_{U,max}$ bounds method

	Data set 1	Data set 2	Data set 3
Original value [mm]	150	110	65
Minimum [mm]	99	73	46
Relative error	-34	-34	-29
Maximum [mm]	210	173	121
Relative error	40	57	86

C_s S-D method

	Data set 1	Data set 2	Data set 3
Original value [h]	500	100	200
First estimate C_s [h]	1770	578	488
Results of iterations			
Number of recessions used	24	11	28
Median [h] After 7 iterations	497	131	172
Std [-]	180	86	103
Relative error [%]	-0.42	31.36	-13.65
Distribution			

P_{max} method

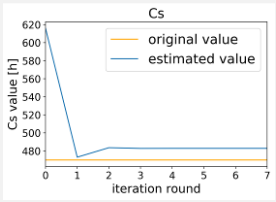
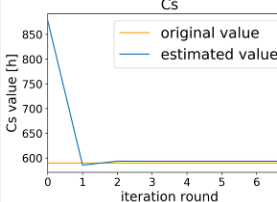
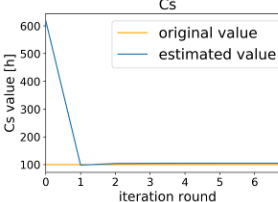
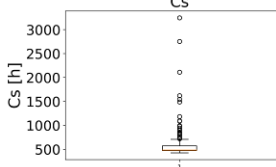
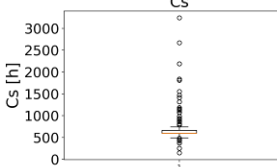
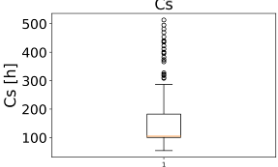
	Data set 1	Data set 2	Data set 3
Original value [mm/h]	0.1	0.07	0.15
Results of iterations			
75 percentile [mm/h]	0.096	0.067	0.12
Relative error [%]	-3.79	-2.98	-16.02
Min. distribution	0.089	0.061	0.096
Max. distribution	0.998	0.073	0.15
Distribution			

Model 9

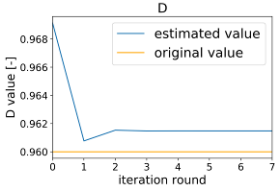
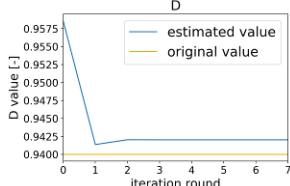
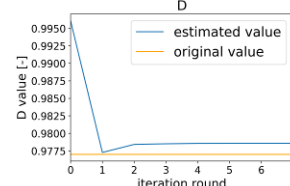
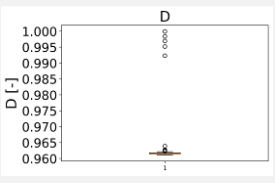
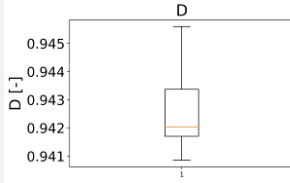
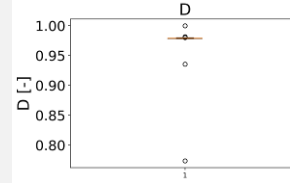
$S_{U,max}$ bounds method

	Data set 1	Data set 2	Data set 3
Original value [mm]	280	327	220
Minimum [mm]	178	206	146
Relative error	-36	-37	-34
Maximum [mm]	309	466	276
Relative error	10	43	25

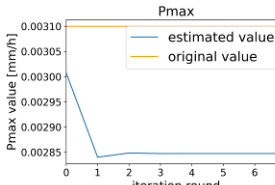
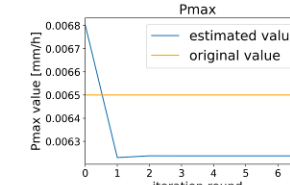
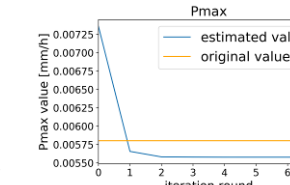
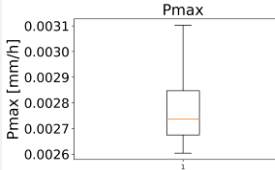
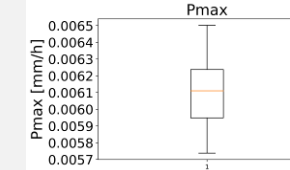
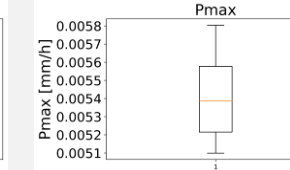
C_s S-D method

	Data set 1	Data set 2	Data set 3
Original value [h]	470	590	100
First estimate C_s [h]	615	877	618
Results of iterations			
Number of recessions used	42	42	54
Median [h] After 7 iterations	482	593	105
Std [-]	385	350	104
Relative error [%]	2.74	0.6	4.68
Distribution			

D method

	Data set 1	Data set 2	Data set 3
Original value	0.96	0.94	0.977
Results of iterations			
Number of peak flow used	68	28	33
Median [-]	0.962	0.942	0.979
After 7 iterations			
Std [-]	0.009	0.001	0.04
Relative error [%]	0.15	0.22	0.16
Distribution			

P_{max} method

	Data set 1	Data set 2	Data set 3
Original value [mm/h]	0.0031	0.0065	0.0058
Results of iterations			
75 percentile [mm/h]	0.0028	0.0062	0.0056
Relative error [%]	-8.15	-4.04	-3.85
Min. distribution	0.0026	0.00573	0.0051
Max. distribution	0.0031	0.00650	0.0058
Distribution			

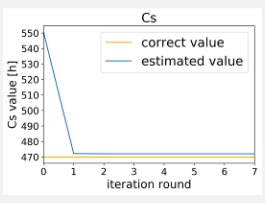
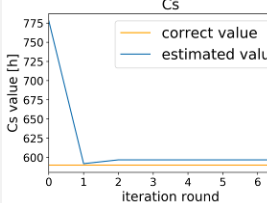
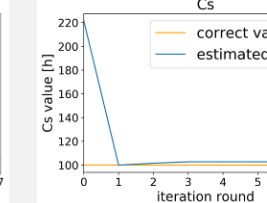
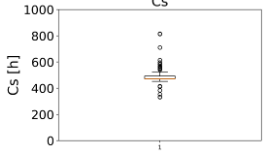
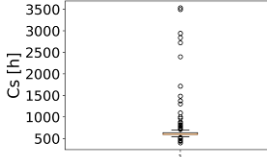
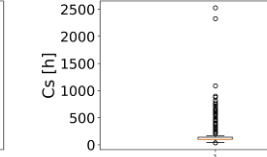
Appendix 6 Results data uncertainty

Precipitation error

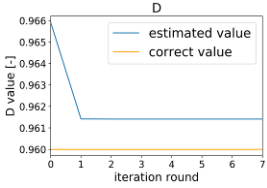
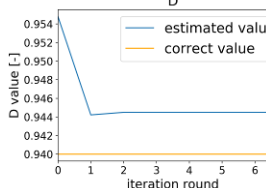
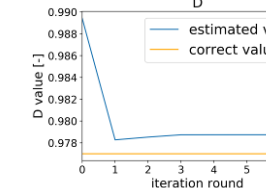
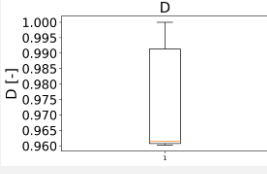
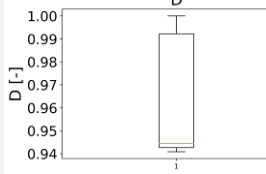
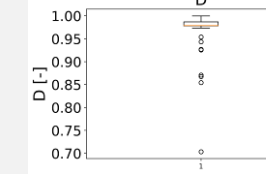
$S_{U,max}$ bounds method

	Data set 1	Data set 2	Data set 3
Original value [mm]	280	327	220
Minimum [mm]	114	140	107
Relative error	-59	-57	-51
Maximum [mm]	359	385	352
Relative error	28	18	60

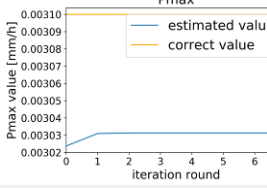
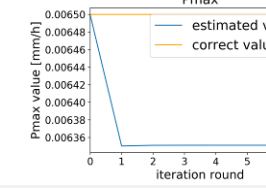
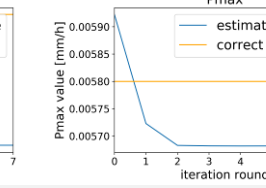
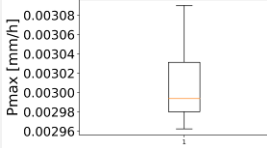
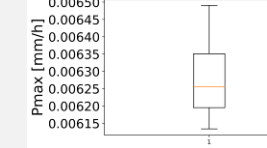
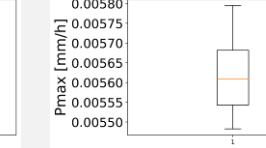
C_s S-D method

	Data set 1	Data set 2	Data set 3
Original value [h]	470	590	100
First estimate C_s [h]	550	778	221
Results of iterations			
Number of recessions used	51	42	45
Median [h] After 7 iterations	472	596	103
Std [-]	525	452	303
Relative error [%]	0.44	1.16	2.74
Distribution			

D method

	Data set 1	Data set 2	Data set 3
Original value	0.96	0.94	0.977
Results of iterations			
Number of peak flow used	100	70	79
Median [-]	0.961	0.944	0.979
After 7 iterations			
Std [-]	0.015	0.024	0.04
Relative error [%]	0.15	0.48	0.18
Distribution			

P_{max} method

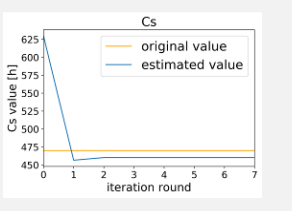
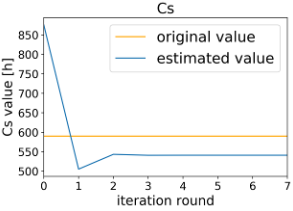
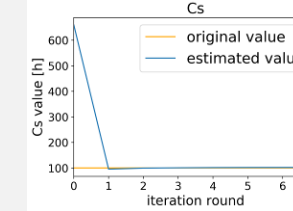
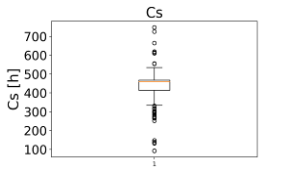
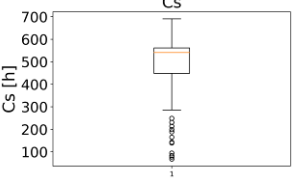
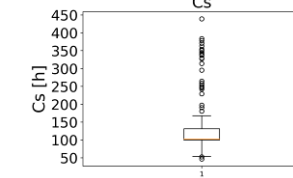
	Data set 1	Data set 2	Data set 3
Original value [mm/h]	0.0031	0.0065	0.0058
Results of iterations			
75 percentile [mm/h]	0.0030	0.0063	0.0056
Relative error [%]	-2.22	-2.29	-2.04
Min. distribution	0.0029	0.0061	0.0054
Max. distribution	0.0031	0.0065	0.0058
Distribution			

Data corruption

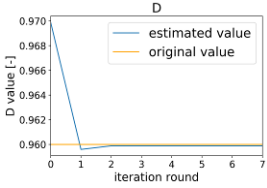
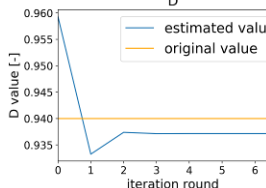
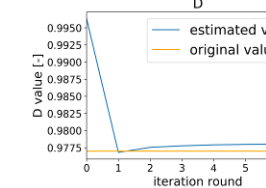
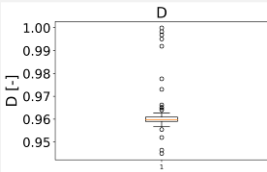
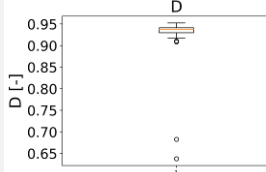
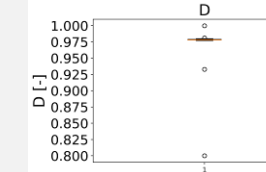
$S_{U,max}$ bounds method

	Data set 1	Data set 2	Data set 3
Original value [mm]	280	327	220
Minimum [mm]	183	203	155
Relative error	-35	-38	-30
Maximum [mm]	323	508	277
Relative error	15	55	26

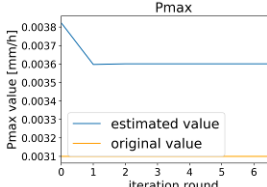
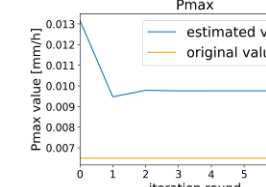
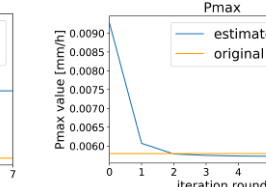
C_s S-D method

	Data set 1	Data set 2	Data set 3
Original value [h]	470	590	100
First estimate C_s [h]	629	875	661
Results of iterations			
Number of recessions used	38	38	43
Median [h]	460	541	102
After 7 iterations			
Std [-]	99	138	80
Relative error [%]	-2.08	-8.27	2.19
Distribution			

D method

	Data set 1	Data set 2	Data set 3
Original value	0.96	0.94	0.977
Results of iterations			
Number of peak flow used	69	28	33
Median [-] After 7 iterations	0.9599	0.937	0.978
Std [-]	0.01	0.07	0.03
Relative error [%]	-0.01	-0.3	0.1
Distribution			

P_{max} method

	Data set 1	Data set 2	Data set 3
Original value [mm/h]	0.0031	0.0065	0.0058
Results of iterations			
75 percentile [mm/h]	0.0036	0.0098	0.0057
Relative error [%]	16.16	50.18	-1.31
Min. distribution	0.0027	0.0069	0.0048
Max. distribution	0.0046	0.130	0.0067
Distribution	