

**To identify and validate techniques
for the identification
of optimal parameter-sets for WALRUS
in different catchments in The
Netherlands**



ing. Caspar N.M. Cluitmans

September 2014

MSc thesis
Hydrology and Quantitative Water Management Group
Wageningen University

**To identify and validate techniques for the identification
of optimal parameter-sets for WALRUS in different catchments in The
Netherlands**

MSc thesis

A thesis submitted in partial fulfilment of the
requirements for the degree of

MASTER OF SCIENCE

By

ing. Caspar N.M. Cluitmans

Examiners:
prof. dr. ir. Remko Uijlenhoet
dr. ir. Ryan Teuling

Supervisors:
drs. Paul Torfs
dr. ir. Ryan Teuling

Hydrology and Quantative Water Management Group
Wageningen University, Wageningen
Source cover picture: www.wpm.nl

Abstract

For this study, techniques for parameter estimation of the Wageningen Lowland Runoff Simulator (WALRUS) were identified and validated. The Levenberg-Marquardt calibration algorithm was selected as a suitable technique for parameter estimation. The used Levenberg-Marquardt package required hard boundary values to be set for each parameter that had to be fitted. To obtain suitable values for the lower- and upper-boundaries, so-called 'rough-runs' were performed. The 'rough-runs' prove to be able to give some insight into what suitable boundary values might be. Its successful implementation however depends on the quality of a priori information on the possible optimal parameter values. Validation of the calibration approach showed that it had to be expanded with regularisation, to obtain more realistic simulations. The dependency on a priori information on the possible optimal parameter values however increased, due to the regularisation method's use of centre-parameters. A multi-start method was added to the approach so that local minima could be detected. The fully developed calibration approach was validated on four lowland catchments. One of these catchments had good a priori information on the possible optimal parameter values. For this catchment good results were obtained. For the other three catchments the above mentioned a priori information was of lower quality. For these catchments it was not possible to obtain good results. A sensitivity analysis and the results of the validation showed that c_G , c_V , c_Q and c_S showed behaviour that could indicate that they are sensitive to changes. The sensitivity analysis was also successfully validated as tool for uncertainty analysis. WALRUS simulations were coupled to see if it was possible to model a large catchment without introducing a hydraulic component. This was possible when only a few sub-catchments were coupled.

Keywords: *hydrological model, rainfall-runoff model, WALRUS, lowland catchments, parameter estimation, calibration, Levenberg-Marquardt, sensitivity analysis, regularisation, multi-start method.*

Preface

From the winter until the end of summer of 2014, I worked on my Master thesis. During this time I spent many hours trying to let R do my bidding. This proved often to be difficult and hours were spent sitting at my desk, trying to think of solutions. However, the moments were sweet when solutions were found, with help from my supervisors or by myself. These moments were often the high points of my thesis. Luckily, there were enough of these moments, so that I enjoyed working on this thesis.

My supervisors played a vital role during my thesis. I would thus like to thank Paul Torfs for all his help with the technical difficulties and for his very clear explanations. Ryan Teuling played an important role in keeping the oversight and always had useful feedback. Finally, I would like to thank Claudia Brauer, for helping me with understanding her model.

Caspar Cluitmans

Contents

Abstract	v
Preface	vii
Contents	ix
List of Figures	xi
List of Tables	xii
1 Introduction	1
1.1 Research objective	1
1.2 Research questions	1
1.3 Structure of the report	2
2 About WALRUS	3
2.1 The Model	3
2.2 The Code	5
2.3 An Example	5
2.3.1 Input	5
2.3.2 Results	6
3 Data	9
3.1 Catchments	9
3.1.1 Diesdonk	9
3.1.2 Lieveren	9
3.1.3 Berkel	9
3.1.4 Tungelroyse beek	9
3.2 The Input and Output	9
3.2.1 Precipitation	10
3.2.2 Evapotranspiration	11
3.2.3 Discharge	11
3.2.4 Quality of the data	12
4 Initial Calibration	13
4.1 Calibration algorithms	13
4.1.1 Calibration algorithms in hydrological modelling	13
4.1.2 Choice of algorithm	13
4.1.3 Quasi-Newton and Levenberg-Marquardt	14
4.1.4 Objective Functions	14
4.1.5 The Nash-Sutcliffe coefficient of efficiency	14
4.2 Methodology	15
4.2.1 The a priori information	15
4.2.2 Boundary values and start values	15
4.2.3 Fixed values	16
4.2.4 Scaling the parameters	16
4.3 Results	16
4.3.1 Obtained parameter values	17
4.3.2 Water balance components	17
4.3.3 Validation	17
4.4 Conclusion	17
5 Sensitivity Analysis	19

5.1	Parameter sensitivity	19
5.1.1	Methodology	19
5.1.2	Results	21
5.2	Forcing sensitivity	24
5.2.1	Induced fractional error: methodology	24
5.2.2	Induced fractional error: results	24
5.2.3	Interception: methodology	24
5.2.4	Interception: results	25
5.3	Conclusion	26
6	Regularisation	27
6.1	Regularisation as boundaries for the parameter values	27
6.2	Methodology	27
6.3	Results	29
6.3.1	The influence of the regularisation parameter	29
6.3.2	The obtained parameter values	29
6.3.3	Validation	29
6.4	Conclusion	29
7	Calibration and Validation	31
7.1	Methodology	31
7.1.1	Multi-start method	31
7.1.2	Calibrating with good a priori information: Stadtlohn	31
7.1.3	Calibrating with low quality a priori information: Diesdonk, Lieveren, Tungelroyse beek	32
7.1.4	Coupling of WALRUS models: Berkel sub-catchments	32
7.1.5	Regularisation	33
7.2	Results	33
7.2.1	Calibrating with good a priori information: Stadtlohn	33
7.2.2	Calibrating with low quality a priori information: Diesdonk, Lieveren, Tungelroyse beek	36
7.2.3	Coupling of WALRUS models: Berkel sub-catchments	37
7.3	Conclusion	39
8	Uncertainty Analyses	41
8.1	Methodology	41
8.2	Results	41
8.3	Conclusion	42
9	Discussion	45
9.1	Data	45
9.2	Model	45
9.3	Sensitivity Analyses	45
9.4	Calibration and validation	46
10	Conclusion	47
11	Recommendations	49
Bibliography		51
Appendices		53
11.1	Chapter 4	53
11.2	Chapter 5	55
11.3	Chapter 6	59
11.4	Chapter 7	63

List of Figures

2.1 Schematisation of WALRUS.	3
2.2 Schematisation of the WALRUS R-scripts.	5
2.3 Location and characteristics of the Hupsel catchment, which is used in the example run.	5
2.4 The plotted output of the example simulation of the Hupsel catchment.	7
3.1 The location of the Dutch catchments used in this research.	9
3.2 The Berkel-catchment with the discharge stations and the precipitation measurement stations.	10
3.3 Schematisation of the Berkel sub-catchments.	11
4.1 An example plot of a 'rough-run'.	15
4.2 The plotted results of the initial calibration.	16
5.1 The rainfall event used for the sensitivity analysis.	21
5.2 The total cumulative sensitivity of the parameters.	22
5.3 The fraction of sensitivity of the parameters.	22
5.4 The NS plotted against the percentages of under- and overestimation for precipitation.	24
5.5 The NS plotted against the percentages of under- and overestimation for evapotranspiration.	24
5.6 Example graph to illustrate the insensitivity of the 'top-off' parameter.	25
5.7 Example graph to illustrate the possible solution for the insensitivity of the 'top-off' parameter.	25
7.1 The model run for the Stadtlohn catchment with the fitted parameters that produced the highest NS (0.62) without hitting any boundaries.	34
7.2 The model run for the Stadtlohn catchment with the fitted-parameters and adjusted d_{G0} .	34
8.1 The rainfall event that occurred in the Stadtlohn catchment, which was used for uncertainty analysis.	42
8.2 The total fractional sensitivity of the Stadtlohn parameters, used for uncertainty analysis.	43
8.3 The total cumulative sensitivity of the Stadtlohn parameters, used for uncertainty analysis.	44
11.1 The first validation run that is performed with the fitted parameters obtained with the initial calibration.	54
11.2 The second validation run that is performed with the fitted parameters obtained with the initial calibration.	54
11.3 The cumulative sensitivity of the parameters obtained for the Hupsel catchment, in the period of approximately the 27th of July till the 15th of August in 2012.	56
11.4 The cumulative sensitivity of the parameters obtained for the Hupsel catchment, in the period of approximately the 27th of July till the 15th of August in 2012.	56
11.5 The sensitivity of the fitted Hupsel parameters for the whole time period, represented in fractions.	57
11.6 The first validation run that is performed with the fitted parameters obtained with the regulated calibration run for the Hupsel catchment.	60
11.7 The second validation run that is performed with the fitted parameters obtained with the regulated calibration run for the Hupsel catchment.	61

List of Tables

2.1	Parameter values used in the example run.	6
2.2	The Table produced by WALRUS with the water balance of the example run.	7
4.1	The obtained fitted parameters for the sensitivity analysis.	17
4.2	The Table produced by WALRUS with the water balance of the results of the initial calibration.	18
5.1	The obtained parameters values from the first 19 calibration runs, used for determining the scale values.	20
5.2	The obtained parameters values from the second set of 19 calibration runs, used for determining the scale values.	20
6.1	The centre-parameter, lower- and upper-boundary values used for the six parameters during regularisation.	28
6.2	Eleven different regulated calibration runs that are performed with different α values.	29
7.1	The obtained parameter-sets of 10 regulated Levenberg-Marquardt runs for the Stadtlohn catchment.	35
7.2	The water balance for the Stadtlohn catchment.	35
7.3	The water balance for the Stadtlohn catchment with a higher c_D .	35
7.4	The obtained parameter-sets of 10 regulated Levenberg-Marquardt runs for the Stadtlohn catchment, with the d_0 excluded from the calibration process.	36
7.5	The water balance for the Stadtlohn catchment for the run with the highest NS.	36
7.6	The obtained parameter-sets of 10 regulated Levenberg-Marquardt runs for the Diesdonk catchment, calibrated with a priori information originating from the Stadtlohn catchment.	37
7.7	The obtained parameter-sets of 10 regulated Levenberg-Marquardt runs for the Diesdonk catchment, calibrated with a priori information originating from the Cabauw catchment.	37
7.8	The obtained parameter sets of 10 regulated Levenberg-Marquardt runs for the Lieveren catchment, calibrated with a priori information originating from the Cabauw catchment.	38
7.9	The results of the Tungelroyse beek calibration runs.	38
7.10	The obtained parameter-sets per Berkelland sub-catchment and their respective objective value.	38
7.11	The best obtained parameter-sets for the two categories of catchments that are calibrated in this chapter.	40
11.1	The NS's of the validation runs performed with the ten fitted parameter-sets that are obtained with the first set of calibration runs of the Stadtlohn catchment.	64
11.2	The NS's of the WALRUS runs performed with different values for c_D for the Stadtlohn catchment.	64
11.3	The NS's of the validation runs performed with the ten fitted parameter-sets that are obtained with the second set of calibration runs of the Stadtlohn catchment with the d_{G0} excluded from the calibration process.	64

1 | Introduction

Recently the Wageningen Lowland Runoff Simulator (WALRUS) has been developed at Wageningen University (Brauer et al., 2014a). This rainfall-runoff model will be the main focus of this Master thesis.

The goal of developing a model like WALRUS is to simulate the behaviour of the hydrological system, in this case the rainfall-runoff process in lowlands. With help of such a model, one could get a better idea of the effect of certain natural processes over time. The improved predictions of such a new model could lead to more fitting solutions to problems concerning water management. When one wants to simulate the rainfall-runoff process, it is important to first identify the required complexity of the model structure. Sorooshian and Gupta (1985) defined this as the first of two stages of model development. In this first stage you ask questions like, does the model need a surface water reservoir? Or, is a certain coupling needed? The model structure of WALRUS is thus already designed (Brauer et al., 2014a). The second stage that Sorooshian and Gupta (1985) defined, is parameter estimation. In other words, finding an optimal parameter-set for a certain catchment. Identifying and validating techniques that are able to find these optimal parameter-sets, will be the main topic of this thesis.

The idea behind parameter estimation is to find suitable values for the parameters, so that the simulated output of the model corresponds to the catchment characteristics. This way, the quality of the predictions made by the model will increase. Finding suitable values however proves to be a challenge in environmental modelling. Parameters of conceptual models like WALRUS, often represent not directly measurable characteristics of a catchment (Wagener et al., 2003). This means that finding optimal parameter values cannot only be done by fieldwork. Due to this problem, inverse modelling is often used. With inverse modelling, relatively easy to obtain observations from the catchment in question are used to estimate the parameter values. An algorithm will try to match the simulated output with the observed output as closely as possible by adjusting the parameters of the model. For inverse modelling, WALRUS needs rainfall, evapotranspiration and observed discharge datasets. This way, the model can simulate the discharge (with the rainfall and evapotranspiration) and the algorithm can adjust the model's parameters (with help of the observed discharge) to obtain good results. In other words, WALRUS needs to be calibrated to find an optimal parameter-set.

It is important to identify a good technical approach to parameter estimation. When this is not done, the model can behave unrealistic with respect to its internal model states, but nonetheless simulate the discharge well. When a good technique is identified, the optimised parameter-sets will truly represent the characteristics of the catchment. The output can then be physically explained by looking at the parameter-values. See Kirchner (2006) for a

discussion on this.

WALRUS is already calibrated for two different catchments. Namely the Hupsel catchment in the east of The Netherlands and the Cabauw catchment in the west of The Netherlands (Brauer et al., 2014b). However, by applying WALRUS in a different way (i.e. with a different calibration approach) on other catchments, the workings of the model might be better understood. The found results can lead to the improvement of the applicability of WALRUS.

1.1 Research objective

The following research objective is used for this thesis:

to identify and validate techniques for the identification of optimal parameter-sets for WALRUS in different catchments in The Netherlands.

Like mentioned above, approaching the parameter estimation from a different direction can increase the insight into the workings of WALRUS. The identification of optimal parameter-sets in different catchments, increases the insight into what values the parameters of WALRUS can have. Due to the fact that only a few studies are done in this area, little information on these types of parameter values is available. In this research the Sum of Squares (SS) will be used as indicator for the model performance.

1.2 Research questions

The following research questions are used to pursue the research objective:

1. Can a different calibration approach than the one used in Brauer et al. (2014b), obtain good objective values?
2. Is the used calibration approach efficient?
3. To what extent do boundaries on the parameters have to be set for the calibration algorithm, to be able to identify optimal parameter-sets?
 - 3.1 Is there a way to identify boundary values for the parameters?
 - 3.2 Is it possible to avoid that fitted parameter values equal boundary values?
4. Can WALRUS be calibrated with a limited amount or lacking knowledge on possible optimal parameter values?
5. Does equifinality occur with WALRUS?
6. Do the same parameters as mentioned in Brauer et al. (2014a) prove to be the most sensitive?
7. Is it possible to find optimal parameter-sets for coupled WALRUS simulations without introducing a hydraulic component?

1.3 Structure of the report

This report is written in a more or less chronological manner. Chapter 2 and Chapter 3 can be seen as the preparatory phase of the research. In Chapter 2 the model itself will be discussed. An explanation will be given of its physical and numerical side. The chapter will be concluded with an explanation of an example run. The used catchments will be discussed in Chapter 3. The calculations that are needed to transform the data to the correct format will also be explained in this chapter.

After the preparatory phase the report continues with the phase which identifies and partially validates the calibration approach. Chapter 4, 5 and 6 are about calibration, sensitivity of the model components and improving the calibration process. The choice for the used calibration algorithm will be explained in Chapter 4. In this chapter an initial calibration run on the test catchment (the Hupsel catchment) will be elaborated. This chapter focusses on the core of the used calibration approach. In Chapter 5 a sensitivity analysis will be performed. With this sensitivity analysis the behaviour of the model is investigated. The main topic of Chapter 6 is regularisation. In this chapter it is investigated if regularisation can improve the calibration process, as applied in Chapter 4.

All the knowledge that is obtained in the first chapters cumulate in Chapter 7. In this Chapter 7 the four catchments are calibrated in different ways. A last addition to the calibration process is also explained in this chapter, namely the use of a multi-start method. Research question number seven will be answered in this chapter. In Chapter 8 the results of one of the calibration runs of the previous chapter will be subjected to an uncertainty analysis. This uncertainty analysis will be performed with help of the same type of sensitivity analysis as explained in Chapter 5.

The Chapters 9, 10 and 11 wrap up the research. In Chapter 9 a critical reflection on the used data and methods takes place. Research questions one to six will be answered in Chapter 10. The recommendations will be elaborated in Chapter 11.

2 | About WALRUS

In this chapter the model will be explained. The first section will focus on how WALRUS simulates the different paths water can take. This includes an elaboration of the couplings and parameters in the model. In section two a short explanation will be given about how the code is built up. The connections between the different scripts will be clarified. Section three discusses a short example to provide insight into the use and application of WALRUS. In the first section of this chapter the most important physical parts of WALRUS will be elaborated. However, some extra physical aspects will be explained in the third section when necessary.

2.1 The Model

In this section the model will be explained on a conceptual level. Figure 2.1 provides a schematisation of the WALRUS model. See the caption of this figure for an extra explanation on the used abbreviations. The state of the content of a catchment is represented in WALRUS by three coupled reservoirs: 1) a soil reservoir, 2) a quickflow reservoir and 3) a surface water reservoir. At least two input variables are needed: precipitation (P) and potential evapotranspiration (ET). The main model output is discharge (Q).

The main paths of the model will now be discussed. In this explanation, precipitation will be seen as the first step in the model. When precipitation falls on the catchment, it is separated between the surface water reservoir (P_S) and the "divider". The P_S is a fixed part of the precipitation and its amount depends on the surface water area fraction (a_S). The divider (W) will separate the precipitation that has fallen on the ground between the quickflow reservoir (P_Q) and the soil reservoir (P_V). The water in the quickflow reservoir will flow at a certain calculated flux into the surface water reservoir (f_{QS}). The quickflow reservoir represents the overland flow that can occur. The rainwater that falls onto the soil reservoir will flow through the vadose zone to the groundwater zone. However, when the rainwater is in the vadose zone it can also evapo(transpi)rate into the atmosphere (ET_V). Evaporation can also take place in the surface water reservoir (ET_S). The discharge flows out of the surface water reservoir and is based on the water level in this reservoir (h_S).

Now that the main paths are discussed the physics will be explained in more detail. As mentioned above, when the precipitation does not fall on the surface water reservoir, it will be divided between the quickflow reservoir and the groundwater reservoir. The amount of water that goes to the groundwater reservoir or the quickflow reservoir is determined by Equation 2.1:

$$W = \cos\left(\frac{\max(\min(d_V, c_W), 0) * \pi}{c_W}\right) * \frac{1}{2} + \frac{1}{2} \quad (2.1)$$

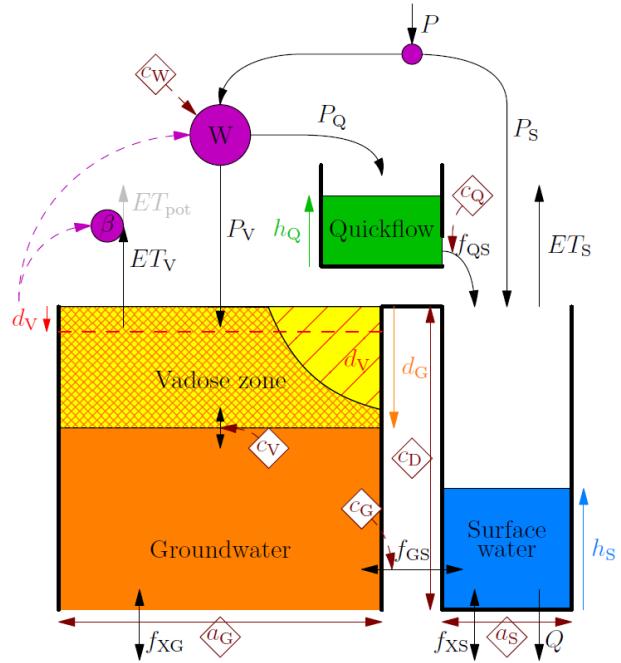


Figure 2.1: Schematisation of WALRUS. The c 's in the figure indicate the model parameters, except for c_D , which forms together with the a 's the supplied parameters. The f 's can be internal or external fluxes. The d 's and the h 's are states. The letters in the subscripts stand for: Q, quickflow; S, surface water; W, the wetness index; V, the vadose zone; G, the groundwater reservoir. X, external reservoirs and D, depth. Source: Brauer et al. (2014a).

where the wetness index is determined by the wetness index parameter (c_W), the storage deficit (d_V) and a default cosine function. When the water is divided by the wetness index, it can go via two reservoirs. The flux via the quickflow reservoir is represented by Equation 2.2:

$$f_{QS} = \frac{h_Q}{c_Q} * a_G \quad (2.2)$$

where the flux from the quickflow reservoir is influenced by the level of the quickflow reservoir (h_Q), the groundwater reservoir constant (c_Q) and the groundwater reservoir area fraction (a_G).

The part of the precipitation that falls on the surface and which does not go via the quickflow reservoir, will go via the groundwater reservoir. In the groundwater reservoir WALRUS explicitly accounts for the coupling between the vadose zone and the groundwater. This means that an increase in groundwater level (or a decrease in groundwater depth, d_G) will lead to a decrease in the size of the vadose zone. The groundwater level acts as a lower boundary condition for the soil moisture profile in the vadose zone. When this level is dynamic it has influence on the soil moisture profile and on processes that are influenced by this soil moisture pro-

file (e.g. ET_V). The change in groundwater depth is calculated with help of the difference between the simulated storage deficit and the equilibrium storage deficit that corresponds to the groundwater level at that specific time:

$$\frac{dd_G}{dt} = \frac{d_V - d_{V,eq}}{c_V} \quad (2.3)$$

where the d_G is the groundwater depth, d_V is the storage deficit, $d_{V,eq}$ the equilibrium storage deficit and c_V the vadose zone relaxation time. There is also a link between groundwater and surface water reservoirs (f_{GS}):

$$f_{GS} = \frac{(c_D - d_G - h_S) * \max((c_D - d_G), h_S)}{c_G} * a_G \quad (2.4)$$

where the flux depends on the height of c_D (i.e., the depth of the channel bottom), the depth of the groundwater table (d_G), the surface water level (h_S) and the parameter c_G . These factors determine if drainage or infiltration can occur between the soil reservoir and the surface water reservoir.

The storage deficit (d_V) influences the wetness index and change in groundwater depth. It also influences the evapotranspiration with Equation 2.5:

$$d_V = \int_0^{d_G} (\theta_S - \theta) dd \quad (2.5)$$

where the θ is the soil moisture content, which corresponds to a certain soil type. The θ_S is the saturated soil moisture profile.

WALRUS also requires a relation between the discharge and water height, the so-called Q,h -relationship. A catchment specific function can be defined for this. However, when no catchment specific function is defined, WALRUS will use a default water height-discharge relationship in the shape of a power-law function:

$$Q = c_S * \left(\frac{h_S - h_{S,min}}{c_D - h_{S,min}} \right)^{x_S} \quad (2.6)$$

where Q is the discharge, h_S is the minimum surface water level and x_S is the exponent which has a default value of 1.5. When the default relation is used, an extra parameter has to be given a value, the bankfull discharge parameter (c_S).

The f_{XG} can represent possible extraction from the groundwater reservoir via seepage to lower layers or via pumps. The f_{XS} can represent pumps or spillways in the surface water system. By assigning simulated upstream discharge to the f_{XS} of the simulation of a downstream catchment, this flux can also be used for the coupling of WALRUS simulations. In such a case it is important, that the simulated upstream discharge is averaged over the downstream catchment.

Internally WALRUS uses the cumulative precipitation. When the amount of rainfall is determined for one time step the cumulative amount of precipitation at the beginning of the time step is subtracted from the cumulative amount at the end of that time

step. This method is used because it works well with the flexible time step approach WALRUS uses.

Gaps in the data are filled in two different ways: missing precipitation values will be set to zero and missing values of potential evaporation, discharge, groundwater level and the external fluxes (f_{XG} and f_{XS}) get an interpolated value.

The script checks if large-scale ponding or flooding occurs. This is the case when the storage deficit (d_V) is smaller than zero and/or the water height (h_S) is larger than the channel depth (c_D). Different rules apply when this is the case. If there is ponding but no flooding all the water in the ponds goes to the surface water. The soil moisture deficit is set to zero. When there is flooding but no ponding the excess surface water will flow into the soil. When there is flooding and ponding WALRUS will simply compute the excess water.

Brauer et al. (2014a) identifies four parameters that show the highest sensitivity to changes. These parameters are thus recommended for calibration. The parameters are i) the soil wetness index parameter (c_W), ii) the groundwater reservoir constant (c_G), iii) the vadose zone relaxation time parameter (c_V) and iv) the quickflow reservoir constant (c_Q). The soil wetness index parameter (i) represents how wet the soil is and thus which part of the water is able to infiltrate and which part will become quickflow. This parameter influences the wetness index, W (Equation 2.1). The vadose zone relaxation time parameter (ii) determines how fast the system will enter a new equilibrium between the vadose zone and the groundwater depth. It is related to the equilibrium storage deficit (Equations 2.3 and 2.5). The groundwater reservoir constant (iii) represents the hydrological characteristics of the soil. The parameter influences Equation 2.4. The quickflow reservoir constant (iv) determines how long it takes for rainfall to travel overland to the surface water reservoir. This parameter influences Equation 2.2. Human influences on the surrounding (e.g., hardening of the top layer, mowing) can influence this parameter. In the example that will be discussed in Section 2.3.1, a more physical interpretation of the parameters will be given. For a full description, I refer to the papers of Brauer et al. (2014a,b).

WALRUS has a default way of determining the initial groundwater depth (d_{G0}). For this, the assumption is made that the initial flux from the groundwater reservoir to the surface water reservoir (f_{GS}), equals the initial discharge. By solving this equation, the d_{G0} is calculated. This method works best when WALRUS starts calculating when the hydrological system in question is in balance. This is around the start of the hydrological year, at the beginning of April. Around this time the system is transitioning from the winter state to the summer state. It has processed the incoming water from the winter precipitation and snow melt and has not yet been influenced by the summer rains. It's also important that at the very beginning of the simulation period no large (exceptional) rainfall peaks occur.

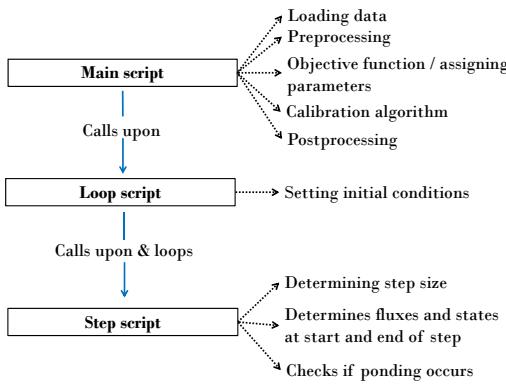


Figure 2.2: Schematisation of the WALRUS R-scripts. In the main-script of WALRUS the necessary data is loaded, the data is preprocessed, the parameters can be assigned and the necessary components for calibration can be written down. The main-script calls upon the loop-script. This script sets the initial conditions and loops for each time-step the step-script. In the step-script the step size is determined, the states and fluxes at the beginning and end of the time step are calculated and a check is performed if ponding occurs.

2.2 The Code

WALRUS is written in R, an open-source statistical programming language. Figure 2.2 provides a schematisation of the WALRUS R-scripts.

From the main-script a full run of the model can be performed. In this script the data is loaded, the preprocessing is done, the parameters are set, the loop-script is called upon and the post-processing occurs. If a calibration run is performed, all the necessary components (the objective function, the calibration algorithm etc.) for such a run, have to be written in this main-script. For the preprocessing, WALRUS calls upon a script specifically created for this task. This script converts the data into functions. For the post processing WALRUS calls upon a script that creates — via other scripts — the figures and tables.

The above mentioned loop-script (WALRUS_loop) sets the initial conditions for the model run and loops the so-called step-script (WALRUS_step). The step-script is run for each time step. In this script the size of the time-step is determined and the fluxes and states at the beginning and end of the time step are determined. With help of these results the script checks if there is ponding.

2.3 An Example

The open source code of WALRUS includes an example run. This run will be elaborated to provide some practical insight into the application of WALRUS.

2.3.1 Input

What follows is an explanation of the forcing and parameters that are needed to perform the WALRUS run. First the area and the data will be discussed. Then the Q,h-relationship that is used will be elaborated. Lastly the predefined parameters will be explained.

The area

The example uses data from the Hupsel catchment. This catchment is located in the eastern part of the Netherlands. See Figure 2.3 for the location and topology of the catchment. The area is above mean sea-level, freely draining but classified as lowland.

The data

The time period used is begin December 2011 till the beginning of June 2012. The data file consists of four columns: the date, the precipitation, the potential evaporation and the observed discharge. Because no calibration takes place in this example, the only value of the observed discharge that is used is the value for the initial discharge. If calibration however would be performed, the observed discharge could be used to calibrate the model.

Q,h-relationship

For the relationship between the discharge and the water height a function specifically created for the Hupsel Brook catchment is called upon (func_Q_hS_hupsel). This function is based on information obtained at the outlet weir of the Hupsel catchment.

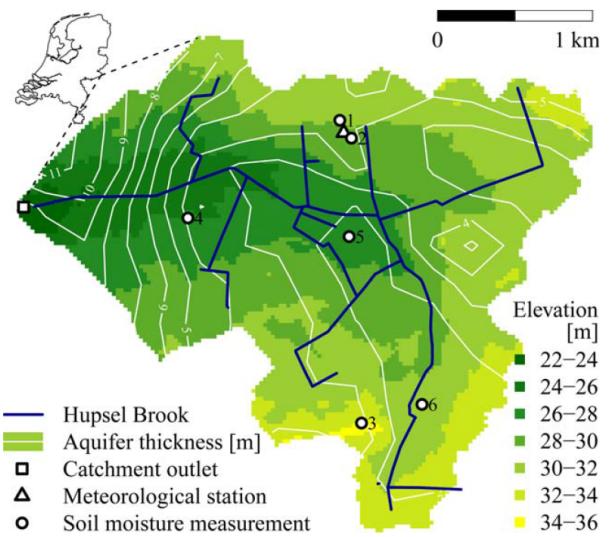


Figure 2.3: Location and characteristics of the Hupsel catchment, which is used in the example run. The elevation shows that there is a relatively big height difference in the catchment, hence the fact that the Hupsel catchment is a freely draining one. The aquifer thickness, ranging from approximately 4 meters to 11 meters is not that thick. The dominating soil type is sand. Source: Brauer et al. (2013).

Table 2.1: Parameter values used in the example run. The c_G is low, meaning that the soil is not that permeable (clayey soil). The c_Q is also not that high, underlining the fact that the catchment is fast responding catchment in which the soil reservoir plays a limited role.

c_W [mm]	c_G [mm h]	c_V [h]	c_Q [h]	c_D [mm]	d_{G0} [mm]	st
356	500000	1	3.3	1500	1156	Hupsel

The parameters

For the example a predefined parameter-set is given. The parameters are written down in a text file. Table 2.1 shows the used values. What follows is a short (physical) interpretation of the values of the parameters.

The soil wetness index parameter (c_W) is related to the storage deficit (d_V). This storage deficit can be seen as a representation of the volume of the empty pores. You could interpret the c_W as a measure of how easily the empty pores fill up. However, when WALRUS is used in a more sloping area the topology of that area might influence the c_W , meaning that not only the amount empty pores can influence the c_W . When c_W is equal to d_V the value of the divider (W) should be zero and all the water should go to the soil reservoir. In this case the soil wetness index parameter has a value of 356 mm. The groundwater reservoir constant is 5 000 000 mm*h. Due to its unit, this value is difficult to interpret. However, the parameter represents the resistance that the soil can exert on the water flowing through it. The resistance depends on the soil type and the drainage density (which relates to the channels/rivers). The vadose relaxation time (c_V) is one hour. This means that it takes one hour for the groundwater table to advance towards a new equilibrium. The quickflow reservoir constant (c_Q) has a value of 3.3 hours. This means that a certain amount of quickflow water would take on average 3.3 hours to travel overland to the surface water reservoir. The drainage density is partly represented by the value for the channel depth. In this case the c_D is 1500 mm. The input field 'soil type' (st) also needs to be indicated in WALRUS. A soil type can be for instance loamy sand. In this example the st refers to a equilibrium soil moisture profile which is specifically determined for the Hupsel catchment. The initial value for the groundwater depth is set manually in this example run. In this case the d_{G0} equals 1156 mm. No external fluxes (f_{XG} , f_{XS}) are given for this example run.

2.3.2 Results

The WALRUS model is timed. It takes the model approximately 27 seconds to go through all the steps. This speed depends for an important part on the calculating power of the computer on which WALRUS is run. However, the obtained run time does indicate that WALRUS is computationally efficient.

In Figure 2.4 the results of the example run are shown. When one looks at the top graph of Figure 2.4 one can see that the simulated discharge

follows the peaks of the observed discharge quite well.

Table 2.2 shows the water balance of the WALRUS run for the Hupsel catchment. The external fluxes (f_{XG} and f_{XS}) are both zero, because no external fluxes were given before hand. The flux from the groundwater reservoir to the surface water reservoir is quite low when compared to the flux from the quickflow reservoir to the surface water reservoir. This might be due to the fact that the Hupsel catchment is a fast reacting catchment.

In its standard run WALRUS also calculates the Nash-Sutcliffe coefficient of efficiency (Nash and Sutcliffe, 1970). In this report the Nash-Sutcliffe coefficient of efficiency will be referred to as NS. In Section 4.1.5 the NS will be further elaborated. The NS value for the example run is 0.83. This objective value indicates that the model is able to mimic the values of the observed discharge well. The NS does not say anything about the identifiability of the parameters or the robustness of the model. Validation runs and the application of WALRUS on different catchment are needed to investigate this.

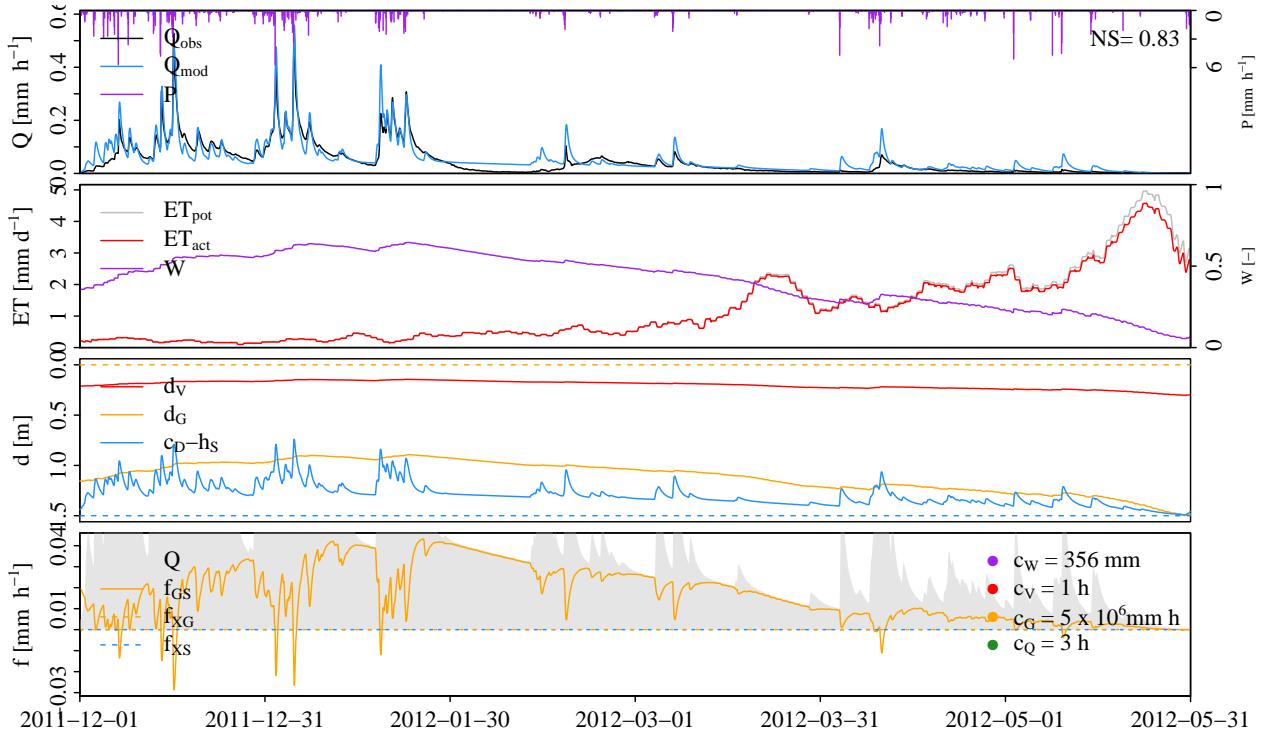


Figure 2.4: The plotted output of the example simulation of the Hupsel catchment. From top to bottom window: the simulated discharge, the observed discharge and the precipitation; the evapotranspiration, with the potential evaporation, the actual evaporation and the wetness index; the states, with the storage deficit, the groundwater depth and the difference between channel depth and water height; the fluxes, with the groundwater-surface water flux, the external flux of the soil reservoir and the external flux of the surface water reservoir. These last two fluxes are both zero because no time series were assigned to these. You can see that the simulated discharge follows the observed discharge quite well. The obtained NS (0.83) is also quite good, expressing the good match between the observed discharge and the simulated discharge.

Table 2.2: The Table produced by WALRUS with the water balance of the example run. The external fluxes are zero because no external fluxes were given before hand. The difference between the f_{GS} and the f_{QS} is quite big. This might be due to the fact that the Hupsel catchment is a fast-reacting catchment. In such a case, the quickflow reservoir often plays a more important role than the groundwater reservoir.

P	ET_{pot}	Q_{obs}	f_{XG}	f_{XS}	ET_{act}	Q	f_{GS}	f_{QS}	d_V	d_G	h_S	h_Q	ndays	check
381	475.1	25.7	0	0	426.5	51.1	7.4	44.2	-96.4	-359.4	-0.2	0	183	0

3 | Data

In the study by Brauer et al. (2014b), two well monitored and measured catchments were used for calibration and validation. It is important to test if WALRUS also performs well in other catchments. In the opinion paper of Gupta et al. (2014) a few benefits are presented for working with more data. i) It leads to an improved understanding of the model. The range of applicability becomes clearer when more data is used. ii) When the model is applied to the data of different catchment, the ability of a model to generalise will be tested. This gives an indication about the robustness of the model. iii) More data can lead to a better understanding of the uncertainty of the model prediction. It can give insight in how much uncertainty to expect when the model is applied to other catchments. It is also interesting to investigate how much the obtained parameters will vary between the catchments.

For this research four catchments have been selected. The choice of the catchments was for a large part driven on the availability of the data. For instance, to perform calibration and validation it was decided that at least a year of data is needed. This way the hydrological behaviour during the four different season might be characterised by the parameter-sets. Because WALRUS works with hourly data, the data has to be available in an hourly-format.

3.1 Catchments

The four Dutch catchments that are selected for this research are: the Diesdonk catchment; the Lieveren catchment; the Berkel catchment and the Tungelroyse beek catchment. Figure 3.1 shows the location of the catchments.

3.1.1 Diesdonk

The Diesdonk catchment is a subcatchment of the river Aa. The river in this catchment is canalized. The elevation of the catchment ranges from 52 m above mean sea level (AMSL) to 14 m AMSL. The size of the catchment is 258,7 square kilometres. The soils in the catchment are mainly fine sands.

3.1.2 Lieveren

The Lieveren catchment is located in the north of the Netherlands. The river is partly canalized. Other parts have been subject to river restoration efforts. The highest part of the catchment is 18 m AMSL and the lowest part 1 m AMSL. The size of the catchment is 106,3 square kilometres. The catchment contains mainly loamy fine sands.

3.1.3 Berkel

The Berkel originates from Germany. Its source is located around Billerbeck. The river flows out into the IJssel at Zutphen. In Germany the river has a

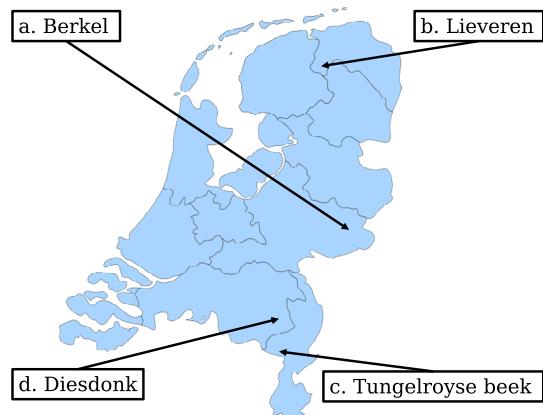


Figure 3.1: The location of the Dutch catchments used in this research. a) The Berkel catchment; b) The Lieveren catchment; c) The Tungelroyse beek catchment; d) The Diesdonk catchment.

slight meandering shape. The human influence on the river in terms of canalisation and drainage is relatively low. In the Netherlands the river is highly canalized and is used for drainage purposes. The elevation of the catchment ranges between 127 m AMSL to 24 m AMSL. The total size of the catchment is 849 square kilometres. The Berkel catchment is mainly dominated by silty sand. For the Berkel catchment, data is available from five different discharge measurement stations. From upstream to downstream these stations are named Lutum (Germany), Stadtlohn (Germany), Rekken (Netherlands, just over the border), Haarlo (Netherlands) and Lochem (Netherlands). The data from these stations allow the Berkel to be divided into five sub-catchments, each named after the discharge measurement station, located at their outlet. Figure 3.2 and Figure 3.3 provide a conceptual overview of the sub-catchments.

3.1.4 Tungelroyse beek

The source of the Tungelroyse beek is located close to Budel in Noord-Brabant. It's about 35 kilometres long and it flows out into Meuse at Neer (where the Tungelroyse beek is called the Neerbeek). Over a length of 30 kilometres the Tungelroyse beek has been redesigned in such a way that it meanders again. The elevation of the research area ranges from approximately 32 m AMSL to 22 m AMSL. The size of the catchment is 152 square kilometres. The catchment of the Tungelroyse beek is mainly dominated by sand soils. However, some loam and clay are present.

3.2 The Input and Output

In Chapter 2 an impression is already given of what is needed for WALRUS as input. To perform a full

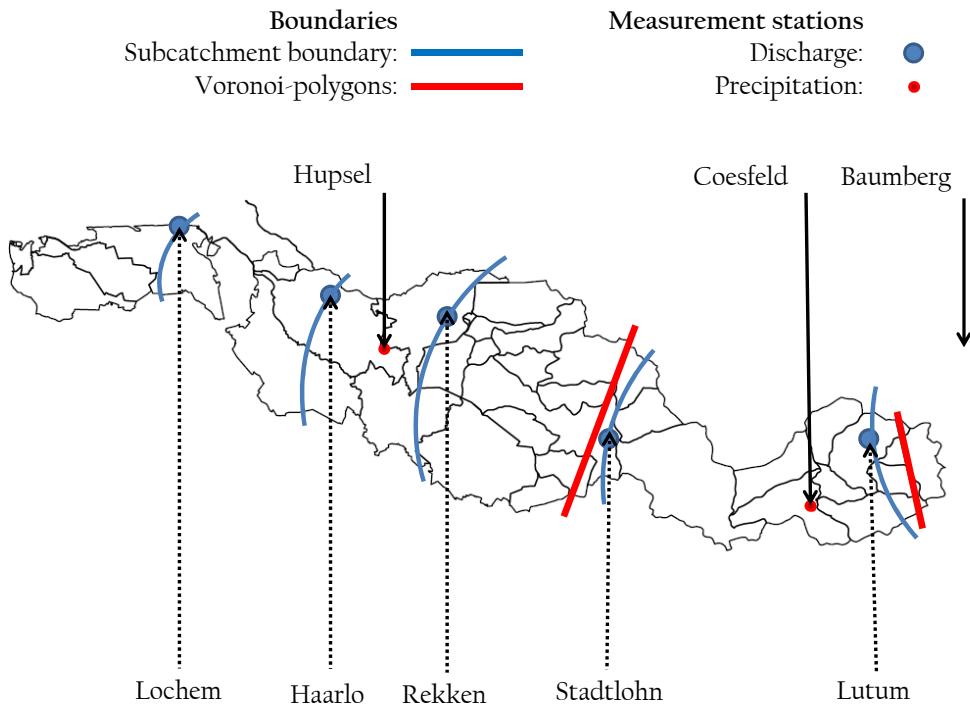


Figure 3.2: The Berkel-catchment with the discharge stations (blue dots) and the precipitation measurement stations (red dots). The blue lines are a rough approximation of the sub-catchment boundaries and the red lines are the Voronoi-polygons. The Baumberg precipitation measurement station is located outside the Berkel-catchment.

run, WALRUS at least needs the hourly data of the precipitation and the potential evapotranspiration of the catchment. To be able to perform calibration and validation runs, the hourly discharge is needed. The groundwater depth can also be used for calibration and validation. Information is also needed on the relationship between discharge and water height. See Section 2.1 for more on the Q,h-relationship.

The data is obtained from several master and bachelor theses, the Royal Dutch Meteorological Institute (KNMI) and DinoLoket. The data of the Diesdonk and Lieveren catchment originate from the master thesis research of Lieke Melsen (Melsen, 2012). The data of the Berkel originates from the bachelor thesis research of Femke Jansen (Jansen, 2012). The discharge data of the Tungelroyse beek originates from the PhD research of Joris Eekhout.

The main unit in WALRUS that is used for the precipitation, evapotranspiration and discharge is mm/h. This means that these in- and outputs are averaged over a certain area. They have to be given as the cumulative values for that hour. To obtain the volumetric amount of these in- and outputs, they should be multiplied with the area of the catchment in question. This then would give the cumulative in- or output of that area per hour (e.g. m³/h). When

a volumetric average is needed (e.g. m³/s) of one hour, the cumulative hourly amount should be divided by 3600 seconds.

3.2.1 Precipitation

The Diesdonk precipitation is based on processed radar-data from Hydroweb (Hydronet). The data is calibrated, validated and spatially averaged in the research of Melsen (2012). The Diesdonk precipitation does not need any extra post-processing to be able to use it in WALRUS.

The precipitation data of the Lieveren catchment originates from rain gauges from Water Board Noorderzijlvest. In the research of Melsen (2012) these gauges are compared with data from the KNMI and are deemed reliable.

The rainfall time-series of the total Berkel catchment of Jansen (2012) can be used without post-processing. For each sub-catchment the precipitation has to be calculated. Three data sets of three different precipitation measurement stations are obtained from the research of Jansen (2012). With help of Voronoi-polygons projected on the map of the Berkel catchment the contribution areas can be determined. The use of Voronoi-polygons is based on the idea that the dividing line for the con-

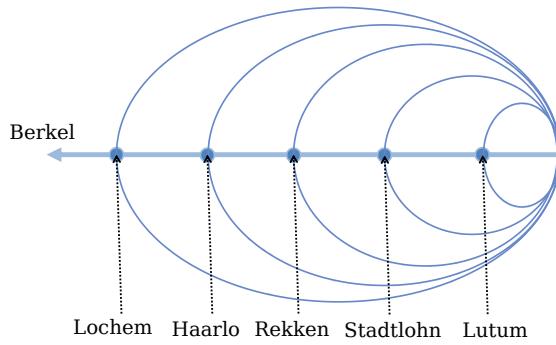


Figure 3.3: Schematisation of the Berkel sub-catchments. The Berkel can be divided in five sub-catchments. The discharge data in Jansen (2012) is averaged over the total area, upstream of the discharge measurement stations.

tribution area of each precipitation measurement stations, is the mid-perpendicular between the measurement stations. The use of Voronoi-polygons is a pragmatic choice. This choice is made because determining the exact contribution area of a rainfall event is complicated and not part of this thesis research. However, it should be noted that the contribution area of the rainfall per measurement station is thus a rough approximation. By inserting the discharge stations the shape of the sub-catchments can be approximated. This, combined with the Voronoi-polygons, leads to estimates of the percentage each precipitation measurement stations contributes to a certain sub-catchment. Figure 3.2 illustrates this.

For the Tungelroyse beek catchment the precipitation data is obtained via the online datacenter of the KNMI. The Eelde measuring station is situated directly below the middle of this catchment. Based on a comparison with other measuring stations surrounding the catchment, it can be assumed that the contribution of rainfall measured by the Ell station is representative for the whole catchment of the Tungelroyse beek.

3.2.2 Evapotranspiration

For the Diesdonk, Lieveren and Tungelroyse beek catchments the hourly potential evapotranspiration has to be calculated. The KNMI station provide the Makkink potential evapotranspiration per day, but not per hour. However, the same stations have datasets with the hourly global radiation. With help of the two previously mentioned datasets and Equation 3.1 the potential evaporation can be calculated per hour. The potential evaporation is calculated by KNMI with the Makkink-formula:

$$ET_0 = 0.65 * \frac{s}{s + \gamma} * \frac{K_{in}}{\lambda} \quad (3.1)$$

ET_0 ($\text{kg} * \text{m}^2 / \text{s}$) is the potential evapotranspiration. s ($\text{kPa} / ^\circ\text{C}$) is the slope of the saturation vapour

pressure curve. γ ($\text{kPa} / ^\circ\text{C}$) is the psychometric constant. K_{in} (W/m^2) stands for the incoming global radiation and λ (J/kg) is the specific heat of vaporization of water.

By dividing the fraction of the daily potential evapotranspiration proportionally over the fraction of the hourly global radiation, the hourly potential evapotranspiration is obtained.

For the Berkel catchment, Jansen (2012) calculated the potential evapotranspiration. These are calculated the same way as the method explained for the other three catchments used in this research.

3.2.3 Discharge

The discharge values have to be converted to mm/h first. The discharge for the Diesdonk and Lieveren catchments are in m^3/s averaged over an hour. This means that the discharge datasets have to be divided by their respective catchment sizes and multiplied by 3600 seconds, to obtain a cumulative discharge with the unit mm/h.

In the bachelor research of Jansen (2012) the discharge is averaged over each sub-catchment and all of the areas upstream of that sub-catchment. Figure 3.3 illustrates this idea. The discharge measured for the Lochem catchment can be seen as the discharge for the whole catchment. The discharge measured for Haarlo, for example, is all of the discharge averaged over the area upstream of the Haarlo discharge measurement station. The discharge-values for the whole Berkel catchment obtained by Jansen (2012) can be used for this research without post-processing. For the discharge for each sub-catchment a few calculations have to be performed to be able to use them. This is because the discharge time series are averaged over the complete area upstream of the discharge measuring stations. Equations 3.2 and 3.3 are used for this.

$$Q_{SubIn} = \frac{Q_u * A_u}{A_{Sub}} \quad (3.2)$$

$$Q_{SubOut} = \frac{Q_{Tot} * A_{Tot}}{A_{Sub}} \quad (3.3)$$

Q_{SubIn} is the discharge that flows into the sub-catchment in question. This is the discharge from the upstream area averaged over the sub-catchment in question. This value can be added to the amount of produced discharge WALRUS calculates for the catchment. Q_{SubOut} can be used for calibrating WALRUS for the sub-catchment. Q_{SubOut} is the out flowing discharge of the sub-catchment in question averaged its area. Q_u is the discharge of the upstream area and A_u is the size of that area. Q_{Tot} and A_{Tot} are respectively the discharge and size of the sub-catchment and its upstream area. A_{Sub} is the area of the sub-catchment in question over which the discharge is averaged.

The discharge data provided by Joris Eekhout is used for the Tungelroyse beek catchment. The discharge measurement location is located in the forest above the village of Haelen. The river is thus not

measured at the outflow location to the east, but approximately four kilometres to the west of this point. The discharge is in m^3/s (averaged over an hour) and has to be averaged over the area of the catchment. The size of the catchment is approximately 152 km^2 . When the original discharge is divided by this value and multiplied by 3600 seconds, a discharge with a unit of mm/h is obtained.

3.2.4 Quality of the data

In general, it is difficult to asses the quality of the data. It is not possible to compare the quality of the different datasets with each other. The only thing that can be noted in this section, is that for the Berkel catchment there appeared to be some errors in the original calculations of the Voronoi-polygons. Because these are done anew in this research, these errors will probably not influence the results.

4 | Initial Calibration

In this chapter a calibration algorithm for parameter estimation will be studied. Using another calibration technique can lead to different optimal parameter-sets and to a different overall optimisation process. By approaching the parameter estimation from a different direction, new behaviour of WALRUS might be revealed. A so-called initial calibration will be performed to validate this calibration algorithm. With help of this validation, the calibration approach can be further developed in subsequent chapters.

Hupsel-data will be used for Chapters 4 and 5. This is because the Hupsel dataset is already thoroughly checked and successful WALRUS runs have been performed with this set (Brauer et al., 2014b). An example of a benefit of using this data is that in these runs no local minima were detected. This lowers the chance that wrong results are obtained for the initial calibration. Local minima are more extensively treated in Chapter 7. The Hupsel-data is the same data which is used in the example run in Chapter 2. For the Hupsel catchment, the Q,h-relationship is known. However, this relationship is often not known for other catchments. The calibration algorithm that will be identified has to be broadly applicable. The catchment specific Q,h-relationship will thus be excluded during the identification and validation process to improve the applicability of the identified calibration approach. Because of this, the bank full discharge parameter (c_s , see Equation 2.6), has to be included in the calibration process. To make the selected calibration approach even more broadly applicable a second extra parameter will be included in the calibration process. Not all the datasets start at the beginning of the hydrological year. This means that the default method in WALRUS will not always be the best method to determine the d_{G0} . This parameter will thus be included in the calibration. In contrary to the research of Brauer et al. (2014b), not four but six parameters will thus be calibrated in this research. The four parameters advised by Brauer et al. (2014a) (c_w , c_G , c_V , c_Q) and the two parameters mentioned above (d_{G0} , c_s).

First a small background introduction will be provided about the calibration algorithm that will be used. A few aspects that are important for calibration will also be elaborated in this part. After this part, the identified method by which WALRUS is calibrated for the Hupsel-catchment will be explained. Lastly the results of the validation — the initial calibration — are discussed.

4.1 Calibration algorithms

In the calibration approach of a hydrological model, two key components can be identified. An objective function, which calls upon the model in question and returns a certain indicator of the model performance, and of course calibration algorithm itself.

First the algorithms will be the the point of focus. After this, the objective functions will be shortly discussed.

4.1.1 Calibration algorithms in hydrological modelling

To provide a better understanding of the use and development of calibration algorithms in hydrological modelling, some background information will be provided. Roughly three important algorithm groups for hydrological models can be identified, namely, classical methods, evolutionary methods and random methods. Classical optimisation methods have long been used, and are still used, for calibrating conceptual hydrological models. Examples of these algorithms are the simplex method (e.g., Spendley et al., 1962) and the Newton-like methods (e.g., Fletcher and Powell, 1963; Marquardt, 1963). These are gradient-based methods and originally deterministic in nature (Wright and Nocedal, 1999). In a deterministic system, a change in input leads to a unique change in ouput. Throughout the time a few drawbacks of these methods have come to light. Due to the possible form of the parameter landscapes, they can have difficulties with global convergence, depending on the starting conditions. Because they both are local search methods, there is a risk that they get stuck in a local minimum and thus miss the optimal value (i.e., the global minimum). This also means that these methods perform less when there is more than one optimal solution (Kavetski et al., 2006). Because of the drawbacks of the simplex and Newton-like methods, the use and further development of other methods have increased throughout the time. An example of an algorithm group that received a lot of attention are the evolutionary optimisation algorithms. These methods are often non-gradient based and can have stochastic elements. In a stochastic system a change in input indicates that there is a chance that the output might fall in this or that region. These evolutionary optimisation methods sometimes make use of particles, or groups of particles, that scan the parameter-space (e.g., Eberhart and Kennedy, 1995). The hydroPSO algorithm (Zambrano-Bigiarini and Rojas, 2013) used in the research of Brauer et al. (2014b) is an example of such a technique. One can also use purely random methods like Monte Carlo. A drawback of the last two mentioned methods is that they can be more time consuming when compared to the classical methods. The more parameters one has, the slower these techniques become ("the curse of dimensionality"). A lot of derivative algorithms exist on these three categories.

4.1.2 Choice of algorithm

There are many different types of optimisation algorithms present in R, which can be applied to WAL-

RUS. The general problem with optimisation algorithms for hydrological models is that none of the earlier mentioned methods can be designated as the single best algorithm. See the paper of Beven and Young (2003) for a discussion on these type of difficulties. Selecting an algorithm on a purely objective basis can prove to be challenging. It can be even debated if it is possible or realistic to come up with some sort of objective selection method. In this research, the selection of practical algorithms for WALRUS is done on basis of pragmatic reasoning.

Because an evolutionary optimisation technique (hydroPSO) is already applied to WALRUS, the choice is made to test the more classical techniques. The main benefit of these techniques over such evolutionary optimisation technique is the computation speed. The environment in which the calibration takes place, namely the R programming-language, also influences the choice of calibration techniques. For this research the optim-package and the minpack.lm-package are tested.

4.1.3 Quasi-Newton and Levenberg-Marquardt

First the standard optimisation package in R is tested, the so-called optim-package. This package is designed in such a way, so that it is broadly applicable. The specific method that is used, is the "L-BFGS-B" method. This method is based on the work of Byrd et al. (1995). It is a quasi-Newton method and allows hard boundaries for the parameter values (also called 'box-constraints'). After two weeks of testing we did not succeed in letting the algorithm function properly. During the successive iterations no convergence occurred. Because the parameters do not get adjusted during the process, something is possibly not correctly assigned for the algorithm via the 'control' commands. However, because these packages are not transparent it was difficult to identify the exact source of the problem. The problem might be partly caused by the fact that the optim-package is a broadly applicable package, it can work with many different objective functions. Because of this, the package is quite complicated and thus more difficult to handle in some cases. It was decided to change to another package.

The minpack.lm-package is a downloadable package for R. It enables the user to use the Levenberg-Marquardt algorithm for optimisation problems. The Levenberg-Marquardt algorithm (Marquardt, 1963) is one of the most widely used optimisation algorithms. It is a deterministic and gradient-based method. It shares an important drawback with other methods of its type, namely that it can get stuck in local minima. However, one of the major benefits of the Levenberg-Marquardt algorithm is its speed. This speed is often higher than many stochastic methods. The problem of getting stuck in local minima can be partially overcome by adapting a multi-start method (Piotrowski and Napiorkowski, 2011). An example of a multi-start method: a particle swarm optimisation algorithm performs a few iterations with a certain num-

ber of particles. For each particle a Levenberg-Marquardt run can then be performed, with the found values of that particle as starting values for the Levenberg-Marquardt run. When two clear groups of parameter-sets are obtained, there are local minima.

The first tests with Levenberg-Marquardt showed convergence.

The Levenberg-Marquardt algorithm was selected due to the fact that successful runs can be performed with it, and because it is a well documented (classical) technique.

4.1.4 Objective Functions

The objective function is the other key component of the calibration approach. It is used by calibration algorithms as indicator of the difference between the model output and the observations. These differences are called residuals. There are different types of objective functions which calculate these residuals. In the research of Brauer et al. (2014b) a test is performed with a few different single-objective functions for WALRUS. For example, an objective function that is better suited for the simulation of peak discharges leads to a more suitable parameter-set for these type of simulations. The use of different objective functions will lead to the emphasize of different aspects in the hydrograph (Wagener and Wheater, 2004). Just like the fact that there is no single best calibration algorithm for hydrological models, there is no single best objective function. The choice of objective function thus depends on the goal of the model run and on the used algorithm.

One of the most used methods is called the least squares method. The goal of this method is to get a small as possible value for the so-called 'sum of squares'. The sum of squares is the sum of the squared residuals. The residuals are squared to avoid problems with negative values. The sum of squares has the following notation:

$$SS = \sum_{t=1}^T (Q_O^t - Q_M^t)^2 \quad (4.1)$$

where Q_O^t are the observations and Q_M^t is the model output.

A way for the Levenberg-Marquardt algorithm to find optimal values for the parameters, is thus to minimise the objective function. In this case the objective function is a function in which a parameter-set is prepared and a WALRUS run is performed. The output of this objective function is a vector containing the differences between the observed values and the simulated values. The vector could also be replaced with the above mentioned Equation 4.1.

4.1.5 The Nash-Sutcliffe coefficient of efficiency

For an indication of the performance of a model the Nash-Sutcliffe coefficient of efficiency (NS) is often used. This coefficient is developed by Nash and Sutcliffe (1970).

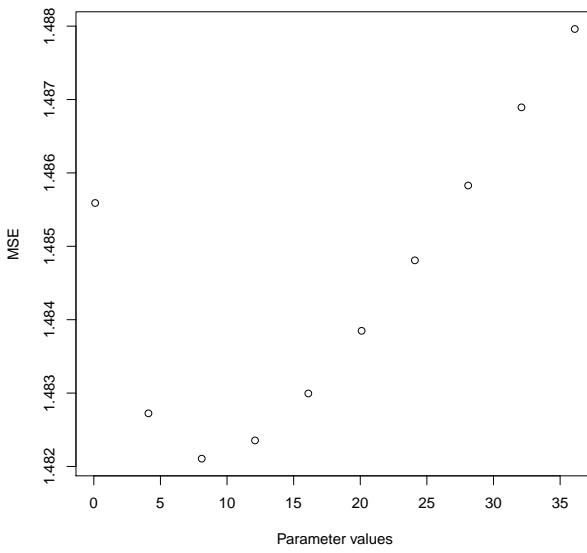


Figure 4.1: An example plot of a 'rough-run'. One parameter is changed for each of the ten WALRUS runs, while the rest of the parameters are fixed throughout the ten runs. This process is repeated for each parameter. The obtained mean sum of squares (MSE) is on the y-axis. For readability, the value is multiplied with one thousand. The parameter values are on the x-axis. The maximum value and minimum values of the parameter values form respectively the lower- and upper-boundary of the calibration algorithm.

$$NS = 1 - \frac{\sum_{t=1}^T (Q_O^t - Q_M^t)^2}{\sum_{t=1}^T (Q_O^t - \bar{Q}_O)^2} = 1 - \frac{SS}{\text{VAR}(Q_O)} \quad (4.2)$$

where Q_O stands for the observed discharge, Q_M is the simulated discharge and \bar{Q}_O is the mean of the observed discharge. As one can see, the NS is simply one minus the sum of squares divided by the variance of the observed discharge.

When the NS is 1 the model results are exactly the same as the observed values. If the NS is 0 the simulated results are as good as the mean of the observed data. When the NS is negative, the simulated results are worse than the mean of the observed data.

4.2 Methodology

Like explained in Section 4.1.3, the Levenberg-Marquardt algorithm will be used. The calibration period starts at the beginning of April, 2012 and ends at the start of October, 2012. This period is selected because it is a relatively wet period, which is quite normal for the area in question. In the research of Brauer et al. (2014b), the period November 2011 - October 2012, was used for the calibration of the Hupsel catchment.

4.2.1 The a priori information

The algorithm in the minpack.lm-package needs to be provided with parameter boundaries and start values. However, it is difficult to select correct lower- and upper-boundary values and start values. WALRUS is a new model and only two fitted parameter-sets yet exist (Brauer et al., 2014b). This means there is very little a priori information on what the correct boundary values for the parameters might be. The fact that WALRUS is a conceptual model contributes to the problem of choosing correct values. The parameters represent larger-scale processes that are governed by smaller-scale properties or processes. They are thus more difficult to link to specific, measurable environmental characteristics. This makes the physical interpretation of the parameter values more difficult. It is thus important that there is some a priori information regarding what the possible parameter values might be. This information can then be used to determine boundary values and start values.

4.2.2 Boundary values and start values

For the validation of the calibration approach, so-called 'rough-runs' are performed to have at least a global idea of what the most efficient boundary values and starting values for the algorithm might be. To be able to perform these 'rough-runs', first a set of start values for the parameters have to be selected. These start values should be based on the a priori information on the possible optimal parameter values of the catchment in question. Besides these single (fixed) start values, a sequence is created with ten values for each individual parameter. In each sequence the values are equally distanced from each other. After this, the 'rough-runs' can be performed. This is done by fixing all the parameters in WALRUS, except for one. WALRUS will now be run ten times, and for each time the non-fixed parameter will have the next value of its respective sequence. When the mean sum of squares of each WALRUS run is plotted against the parameter values of the non-fixed parameter, an idea can be formed of what might be the optimal starting value of that parameter for Levenberg-Marquardt. When the plot shows no clear minimum, the sequence that is created can be adjusted and the 'rough-run' can be repeated. The 'rough-run' method is repeated for each parameter. Figure 4.1 shows an example of such a plot. The lower- and upper-boundaries for the calibration run are formed by the respective minimum and maximum parameter values of the 'rough-run'. It is thus important that the values of the above mentioned sequences are selected carefully. When a 'rough-run' of a parameter shows a clear minimum, the minimum value is taken. When the 'rough-run' only shows a line that descends in one direction, the mean is taken.

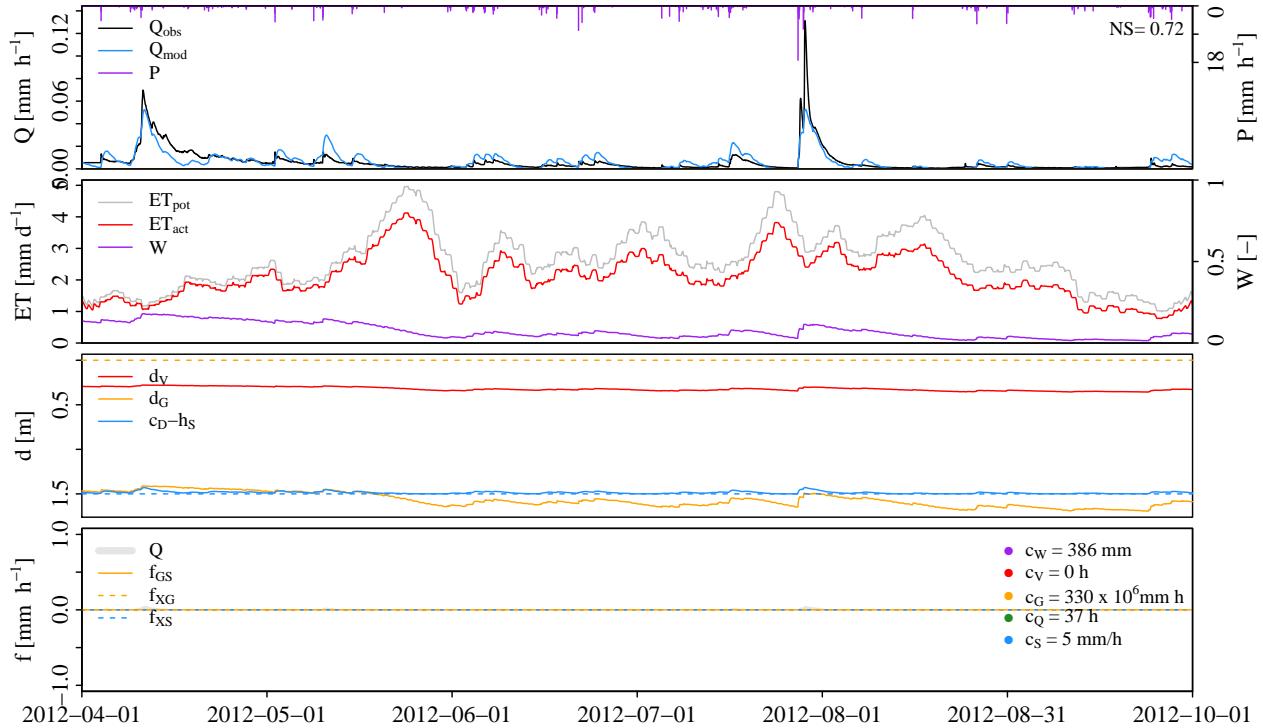


Figure 4.2: The plotted results of the initial calibration. The plot is created by the WALRUS R-script. Judging from the NS (0.72), it could be said that the model performs relatively well. The simulated discharge seems to follow the observed discharge quite good, judging from the top window. However, it is prone to underestimation during large peaks. The parameter values in the lowest window show that the obtained parameters are not that realistic. Comparison with the values Brauer et al. (2014b) obtained, shows that the c_G and the c_Q might be too high.

4.2.3 Fixed values

The parameter-set in the objective functions contains nine parameters, including the six (variable) parameters that have to be calibrated. The other three parameters are the channel depth (c_D), the surface water area fraction (a_S) and the soil type (st). These three parameters are fixed with the same values as in Chapter 2.

4.2.4 Scaling the parameters

The process of the calibration algorithm getting closer to a solution, is called convergence. This process of convergence can be speeded up by letting the Levenberg-Marquardt only handle parameters that have the same scales, e.g. between 0 and 1. This is done because of the differences in sensitivity of each parameter towards changes imposed by the algorithm. The algorithm imposes these changes because it wants to find an optimal value for the parameters. An example of this type of sensitivity: the groundwater reservoir parameter is a few times bigger than the wetness index parameter. However, the algorithm is trying to change these values with the same amount. It thus could take a lot longer to find an optimal value for the groundwater reservoir parameter than for the wetness index parameter due to its value-size. Two functions are needed to scale to parameter values back and forth. A function that scales the parameters from

the Levenberg-Marquardt size to the (normal) WALRUS size, and vice versa. An example of these functions is given for the parameter c_W :

$$\text{from_WALRUS_to_LM} = \text{LB}_{c_W} + \text{UB}_{c_W} * \text{VAR}_{c_W} \quad (4.3)$$

$$\text{from_LM_to_WALRUS} = \frac{\text{VAR}_{c_W} - \text{LB}_{c_W}}{\text{UB}_{c_W}} \quad (4.4)$$

LB_{c_W} is the imposed lower boundary of the parameter c_W . UB_{c_W} is the upper boundary of this parameter. VAR_{c_W} is the variable value for c_W for which the algorithm is trying to find an optimal value.

4.3 Results

Table 4.1 shows the fitted parameters. When a WALRUS run is performed with the fitted parameters, a NS of 0.72 is obtained. The graphs produced by the WALRUS R-script for this run are shown in Figure 4.2. You can see in this figure that the simulated discharge seems to follow the observed discharge quite well. Small peaks are however slightly overestimated and large peaks underestimated.

Multiple calibration runs with the same starting values and boundaries lead to the same parameter-values. This indicates that the Levenberg-Marquardt algorithm is stable and the convergence is independent of the starting point.

Table 4.1: The obtained fitted parameters for the sensitivity analysis. The parameter values that equal boundary values are printed in bold font. When compared with the obtained parameters of Brauer et al. (2014b), the c_G and c_Q appear to be too high. The parameters c_G and c_V both hit an imposed boundary.

	c_W [mm]	c_G [mm h]	c_V [h]	c_Q [h]	d_{GO} [mm]	c_S [mm/h]
Fitted parameters	386.2	330e6	0.2	36.53	1467	5.156

4.3.1 Obtained parameter values

A NS of 0.72 is a relatively good value. However, what one cannot deduct from this value is that two parameter values equal boundary values. This happened for the groundwater reservoir parameter and for the vadose zone relaxation time. These boundary values are often unrealistic. This is because they have to be broad enough to be able to capture all the possible realistic values between them. It might also be the case that the upper- and lower-boundary values are selected wrong. For the c_V , Brauer et al. (2014b) found a value of 0.2 hour. This means that the value found with this calibration could be close to the optimal value. At least one side note has to be made however, because the c_V hit an imposed boundary. Because the other parameters have other values than Brauer et al. (2014b) found, the numerical optimal value might not be 0.2. The difference between the values of Brauer et al. (2014b) and this calibration run might be produced by the different calibration algorithms that are used. The quick flow reservoir coefficient is also too high, but it did not hit a boundary. In the research of Brauer et al. (2014b) a c_Q of 1 hour is found for the Hupsel. In this case the algorithm found a c_Q of approximately 36.5 hours, which seems unrealistic for a fast-reacting catchment like the Hupsel catchment. The value for the soil wetness index parameter, approximately 386.2 mm, is quite close to the value Brauer et al. (2014b) found for this parameter (356 mm).

4.3.2 Water balance components

Table 4.2 shows the water balance of the WALRUS run with the fitted parameters. The f_{GS} , the flux from the groundwater reservoir to the surface water reservoir is equal to zero. The f_{QS} , the flux from the quick flow reservoir to the surface water reservoir is equal to 26.5. The fact that f_{GS} equals zero can be explained by looking at the value for c_G . This value is really high, which leads to the direct drainage of all the water entering the ground reservoir. There is thus no possibility for the water in the ground reservoir to enter the surface water reservoir. This also means that the quick flow reservoir has to compensate for the not properly functioning ground reservoir with its own flux. Like mentioned earlier, the c_Q value is also really high. When all the water would go via the quick flow reservoir in a really fast manner (the c_G would then be low), the simulated discharge would become too high at certain moments. Because the calibration algorithm wants to make the difference between the observed discharge and the simulated discharge as small as possible, this

slowing down effect — which is not provided by the ground water reservoir — has to be simulated in another way. By making the c_Q high, the water will flow slowly through the "quickflow" reservoir.

4.3.3 Validation

Validation runs are performed to be able to investigate the performance of WALRUS with the fitted parameter-set in other years. The two periods that are used are: from the beginning of April till the beginning of October, 2011; from the beginning of April till approximately half of September, 2013. The year of 2011 is a dry year but at 3/4 of the used period there is a large rainfall peak. The period in 2013 is wetter than that of 2011, but the beginning of spring is marked by a prolonged dry period. The NS's that are obtained are respectively 0.83 and 0.44. Figure 11.1 and Figure 11.2 in the Appendix show the plots that the WALRUS R-script created of these runs. The NS indicates that in 2011 WALRUS seems to perform even better with the fitted parameters than for the period for which WALRUS is calibrated. However, in 2013 the performance considerably drops. This can indicate that with this fitted parameter-set, WALRUS is not able to perform that well in different situations.

4.4 Conclusion

It is possible to obtain fitted parameter-sets with the Levenberg-Marquardt algorithm for WALRUS. However, the results show that using Levenberg-Marquardt with hard boundaries is not sufficient to obtain realistic results. The 'rough-run' method is able to give some insight into what efficient values might be for the lower- and upper-boundary values for the parameters, and for the start values of the algorithm. The method is however really depended on the a priori information on the possible parameter values.

In Chapter 6 regularisation will be investigated. With help of this technique, problems like hitting boundaries and obtaining unrealistic values might be solved. For the sensitivity analysis in Chapter 5, the fitted parameters obtained via the initial calibration will be used.

Table 4.2: The Table produced by WALRUS with the water balance of the results of the initial calibration. The flux from the groundwater reservoir to the surface water reservoir is zero. This means that the groundwater reservoir is not completely functioning. This can be explained by looking at the c_G value, which is really high. All the water in the reservoir will directly go away to a deeper layer, without being able to flow to the surface water reservoir. The flux from the quick flow reservoir to the surface water reservoir is high. This is probably because this reservoir compensates for the not properly functioning groundwater reservoir. The external fluxes are zero because no external fluxes were given before-hand.

P	ET_{pot}	Q_{obs}	f_{XG}	f_{XS}	ET_{act}	Q	f_{GS}	f_{QS}	d_V	d_G	h_S	h_Q	$ndays$	$check$
381	475.1	25.7	0	0	387.6	26.6	0	26.5	-33	-119.8	0	-0.1	183	0

5 | Sensitivity Analysis

In this chapter a sensitivity analysis will be performed for WALRUS with the fitted parameter-set that is obtained in the previous chapter. In the research of Brauer et al. (2014b) a sensitivity analysis is already performed. The analysis in that research is done for the calibration runs on two specific catchments. It identified, among others, the parameters that showed the most sensitivity. These parameters are part of the fitted parameter-set obtained for the Hupsel catchment in the previous chapter. Because a different calibration technique is used, more information on the behaviour (sensitivity) of the model parameters might prove to be useful. The extra insight into the sensitivity of the parameters, can be used so that the calibration algorithm can be applied in the most effective way.

The above mentioned sensitivities relate to the model structure. An analysis of these type of sensitivities bring the identifiability of the parameters to light (Shin et al., 2013). Besides the sensitivities in the model structure itself, WALRUS can also be sensitive for changes in input (i.e. forcing). A research into the sensitivity of the model towards forcing, mainly gives an indication of the amount of uncertainty produced by the possible noise in a dataset. An insight in this behaviour will add an extra dimension to the interpretation of the results.

The behaviour of the WALRUS-model towards changes in parameter values and forcing, will thus be the focus of this chapter. The period that will be used throughout this chapter is from the beginning of April 2012 (start of the hydrological year) until the beginning of October.

5.1 Parameter sensitivity

In this section the methods and results of the research into the parameter sensitivity will be elaborated.

5.1.1 Methodology

The sensitivity of the model towards each parameter is calculated with help of Equation 5.1:

$$\text{sens_parameter} = \frac{\text{adjusted_run} - \text{base_run}}{\text{eps}} \quad (5.1)$$

where the sens_parameter is the derivative which gives an indication of the sensitivity of each point in the time series. The adjusted_run is a WALRUS run with a fitted parameter-set in which one parameter is adjusted. The adjusted parameter equals the unadjusted parameter plus eps. The base_run is a normal WALRUS run with the unadjusted fitted parameters. eps is the amount by which the parameter is increased and the value by which the difference between the adjusted_run and the base_run is divided. The value of eps for each parameter is based on its lower- and upper-boundary, which are based on a priori information

(see Section 4.2). The difference of the lower- and upper-boundary values is divided by a hundred to obtain the parameter-specific eps. The sens_parameter is calculated for each of the six calibrated parameters. In literature this method is also often referred to as "One-At-a-Time", abbreviated as OAT. To obtain the total sensitivity of the six parameters, Equation 5.2 is used.

$$\begin{aligned} \text{sens_tot}^2 = & (\text{scale}_cW * \text{sens}_cW)^2 + (\text{scale}_cG * \text{sens}_cG)^2 \\ & + (\text{scale}_cV * \text{sens}_cV)^2 + (\text{scale}_cQ * \text{sens}_cQ)^2 \\ & + (\text{scale}_dG0 * \text{sens}_dG0)^2 + (\text{scale}_cS * \text{sens}_cS)^2 \end{aligned} \quad (5.2)$$

where sens_cW is the specific sensitivity of the parameter c_W (the sens_parameter in Equation 5.1). The scale_... is needed to be able to compare the sensitivity of the parameters with each other. The squares in Equation 5.2 are taken to avoid negative values. For each parameter a scale value thus has to be created. These scale values need to have two properties: 1) they need to be able to scale the sens_parameter values in such a way that the sensitivities of the parameters can be compared with each other. For example, the physical dimensions of the groundwater reservoir parameter and the wetness index parameter are difficult to compare. The difference in the size of their respective units is a factor 1e6. 2) The second property the scale values need to reflect, is the size of the variation that calibrated parameter values can have for different time periods in the same catchment. When they can have relatively large differences, they are more sensitive to changes than values that don't have these large differences. This could mean that when a validation run is performed on calibrated values, the results could wildly differ because of the parameters that show these large variations in values. To get an idea of this size of variation, multiple calibration runs are performed. The calibration method used for these runs is the same one used for the initial calibration (see Section 4.2 for an elaboration). The time period of each of these runs is six months. For each run the period will be shifted with one month. To speed up the calibration process, the obtained fitted parameter-set of each run is used as the starting parameter values for the next run. When a boundary is hit in a run, the obtained parameters that hit the boundary will be put ten percent away from the respective boundary. This way it will become clear if the algorithm constantly wants to converge to that boundary, or that the calibration algorithm will search new values within the imposed boundaries throughout the changing time periods.

The scales are calculated in two different ways (Equations 5.3 and 5.4).

$$\text{scale}_cW = \frac{cW_{\max} - cW_{\min}}{cW_{\text{mean}}} \quad (5.3)$$

Table 5.1: The obtained parameters values from the first 19 calibration runs, used for determining the scale values. Parameter values that equal boundary values are printed in bold font. The obtained values can be used to determine the scale values for the sensitivity analysis of the parameters. The runs are performed with a moving time-window between the beginning of April 2012 till the beginning of October of that same year. Each calibration period is six months. The period of each successive run is shifted with one month. The obtained parameters from each run are used as starting parameters for the next one to increase the speed. When fitted parameters are equal to a boundary, they are supposed to be put ten percent away from that boundary. However, this does not work because some parameter values just fall outside a border. This problem is solved and a second set of 19 calibration runs is performed (see Table 5.2 for the results).

Run	c_W [h]	c_G [mm]	c_V [h]	c_Q [mm/h]	d_{G0} [mm]	c_S [mm/h]
1	47.937	36725760	10.03	24.06	100.0	18.2
2	103.581	48185198	0.20	50.05	427.1	18.2
3	95.216	38809975	0.20	44.45	511.0	18.2
4	1.000	300000000	0.20	0.01	573.3	18.2
5	159.907	330000000	0.20	44.36	481.7	18.2
6	142.401	330000000	36.30	47.10	806.7	18.2
7	123.070	330000000	36.30	39.86	917.4	18.2
8	148.425	330000000	36.30	41.77	966.8	18.2
9	138.814	330000000	36.30	45.21	668.9	18.2
10	133.975	330000000	36.30	47.49	715.5	18.2
11	118.057	330000000	36.30	50.33	600.0	18.2
12	116.240	330000000	36.30	50.38	651.0	18.2
13	103.866	330000000	36.30	49.76	441.8	18.2
14	106.351	330000000	36.30	72.02	371.4	18.2
15	97.642	330000000	36.30	72.02	397.2	18.2
16	48.432	330000000	36.30	72.02	409.5	18.2
17	3.333	330000000	36.30	72.02	177.3	18.2
18	1.508	330000000	36.30	72.02	267.7	18.2
19	1.061	330000000	36.30	72.02	222.6	18.2

Table 5.2: The obtained parameters values from the second set of 19 calibration runs, used for determining the scale values. The parameter values that equal boundary values are printed in bold font. When parameters have a value in between five percent above or below a boundary, they will be put ten percent away from that boundary. This solved the problem that occurred with the first set of 19 calibration runs. Compared to Table 5.1 the new values show a lot more variation. The values give an insight in the size in variations for parameter values can have in the same catchment.

Run	c_W [h]	c_G [mm]	c_V [h]	c_Q [mm/h]	d_{G0} [mm]	c_S [mm/h]
1	47.94	36725760	10.0348	24.06	100.0	18.200
2	106.70	47062304	0.2000	49.46	435.3	15.270
3	91.68	40126620	0.2000	44.04	494.5	13.539
4	33.85	30000000	0.2000	0.01	559.7	18.200
5	33.85	30000000	0.2000	7.21	557.0	18.200
6	144.38	109579158	0.2000	41.01	815.4	18.200
7	116.63	103676498	0.2000	32.84	904.8	18.200
8	141.05	60763984	0.2000	31.91	963.5	18.200
9	137.00	64364883	0.2000	34.03	684.6	18.200
10	123.29	54486233	0.2000	33.34	706.5	18.200
11	150.90	45045018	36.3000	33.48	688.5	18.200
12	154.11	30000000	0.2000	27.73	802.3	18.200
13	139.94	40129470	0.2746	31.16	564.7	18.200
14	145.45	65542186	0.3767	72.02	528.8	18.200
15	117.64	92484204	0.2584	72.02	487.7	10.167
16	47.55	129858598	0.2000	29.80	427.2	6.094
17	41.66	94060723	0.2000	25.30	431.9	4.658
18	1.00	119742283	0.2000	35.42	283.5	4.600
19	117.81	240638336	0.2647	32.14	652.2	18.200

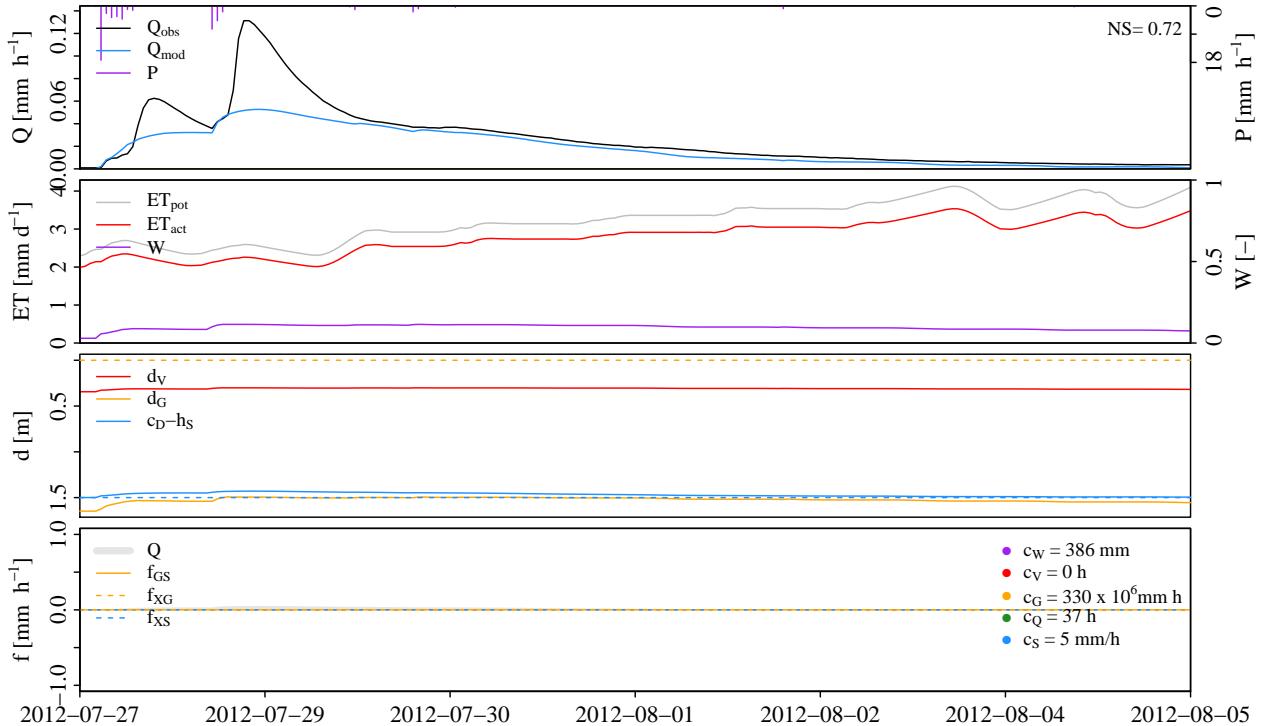


Figure 5.1: The rainfall event used for the sensitivity analysis. This rainfall event occurred between the 27th of July and the 5th of August in 2012. The purple lines show the precipitation, the black line shows the observed discharge and the blue line shows the discharge modelled by WALRUS. With this plot the purely mathematical/black-box approach of the sensitivity analysis can be linked to the physical processes.

$$\text{scale_cW} = \frac{\text{sd(cW)}}{\text{cW_mean}} \quad (5.4)$$

where cW_{max} and cW_{min} are respectively the maximum and minimum value found for c_w with the multiple calibration runs. cW_{mean} is the mean of all the c_w values found with the multiple calibration runs. $\text{sd}(cW)$ is the standard deviation of the multiple values found for c_w .

In Equation 5.3 the difference between the maximum and minimum found values is used to determine the scale. This value is divided by the mean to scale the values so that the different scale values can be compared. In Equation 5.4 the difference is replaced with the standard deviation. Both are used to see what the effect is of different scale values.

The total sensitivity can be represented in a cumulation of the different sensitivities and in the fraction each parameter adds to the total sensitivity. The square in Equation 5.2 has as drawback for the cumulative way of representing the number. It enlarges the peaks. This way it can become unreadable when there are a few peak sensitivities in the data. Because of this an extra formula is created in which the square root is taken from the terms in Equation 5.2 after the square taken.

5.1.2 Results

Table 5.1 shows the results of the first 19 calibration runs that are performed for the determination of the scale values. The used method brings to

light that there is a problem with some of the fitted parameter values that Levenberg-Marquardt finds. Some of these values fall slightly outside of the imposed upper- and lower-boundary. This leads to problems when the scripts checks if the found fitted-parameters hit a boundary, and if they then need to be put ten percent away from that boundary before they are used as starting values for the next run. When an if-statement is used to check if a parameter value is equal to a boundary, it will not work because of the slightly different values. For example, for c_w Levenberg-Marquardt found in a test run a value of 722, while the upper boundary actually has a value of 721. For the next run this found value will not be put ten percent away from the boundary because the values are not equal. Because of this problem, the if-statement that checks if parameters from a previous calibration are equal to a boundary, is changed. When parameters have a value in between five percent above or below a boundary, they will be put ten percent away from that boundary. Table 5.2 shows the new set of 19 calibration runs. The newly adopted approach seems to have fixed the problem.

When one looks at both sets of 19 calibration runs, one can see there is quite some difference between the values of both sets. In the old set, the parameters that hit a boundary are used for the next run without adjustment due to the fault in the script. Apparently it is then difficult for Levenberg-Marquardt algorithm to diverge from such a boundary, when the starting value (approximately) equals

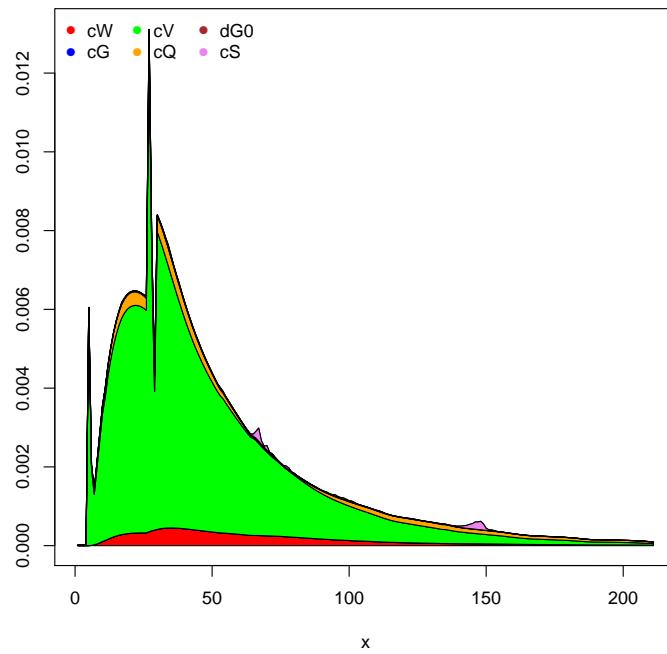


Figure 5.2: The total cumulative sensitivity of the parameters in the period of approximately the 27th of July till the 15th of August in 2012. On the y-axis the amount sensitivity is given. On the x-axis the time steps are shown. The scale values in Equation 5.2 are determined with help of Equation 5.4. Due to the amount of uncertainty concerning the correctness of the scale values, this plot should only be used to get an impression of the shape of total cumulative sensitivity.

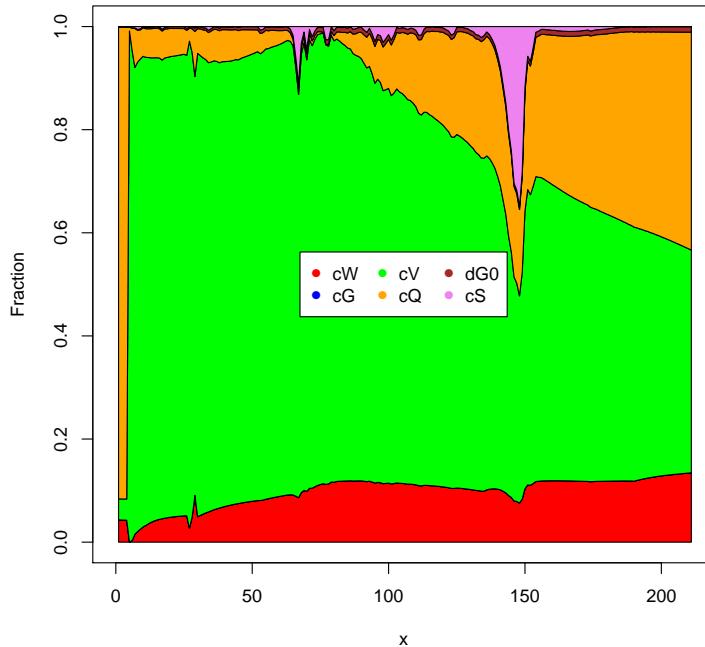


Figure 5.3: The fraction of sensitivity of the parameters in the period of approximately the 27th of July till the 15th of August in 2012. On the y-axis the fraction of sensitivity is given. On the x-axis the time steps are shown. The scale values in Equation 5.2 are determined with help of Equation 5.4. Because this plot portrays the amount of sensitivity in fractions it is less depended on the correctness of the scale values. This is due to the fact that all the values are divided by Equation 5.2.

the boundary. It also means that the calibration algorithm starts changing the other parameters, because it wants to have the residuals as small as possible. These other parameters thus are selected in such a way that they compensate for the parameter that hit the boundary. This sort of behaviour does not lead to physically realistic values. This idea is underlined by the second set. Table 5.2 shows that it is possible for the calibration algorithm to find new values, once the starting values are placed sufficiently away from the boundary. This can indicate that these type of hard borders are not that practical, or maybe even slow down the calibration process.

In the old set of runs, c_G hit the upper boundary early on and did not diverge from that boundary in the later runs. In the new set of runs c_G shows a lot of variation. The variation in c_V also improved slightly. However, this parameter is clearly prone to hitting boundaries. One of the biggest differences in c_V values between the old set of runs and the new set, is that c_V hits the lower boundary instead of the upper boundary. c_S also showed some more variation, however, it is still prone to hitting the upper boundary. c_W , c_Q and d_{G0} do not constantly hit certain boundaries and show a certain degree of variation.

The next step is to see what the influence is of the difference between Equation 5.3 and Equation 5.4. There is only a slight difference in values when Equation 5.3 or Equation 5.4 is used. Some peaks in sensitivity seem to be more pronounced when the scale values are determined by the standard deviation. When the difference is used to calculate the scale values, all the lines seem to be a bit smoother. In the appendix two zoomed-in plots of the fractions of the sensitivities are placed. The time period that is used for these plots is from the 15th of August 2012 till the 25th of July 2012. One is calculated with Equation 5.3 (Figure 11.3) and one with Equation 5.4 (Figure 11.4). Because the sensitivities calculated with Equation 5.4 show the differences better due to the more pronounced sensitivities, these are used for further analysis.

A plot of the cumulative sensitivity expressed in fractions per time step for the whole time-period is placed in the appendix (Figure 11.5). This plot gives an impression on how the sensitivity can vary throughout a few months. With help of a plot that expresses the cumulative sensitivity (not in fractions) a moment with a peak sensitivity is selected and zoomed-in upon. The peak in sensitivity occurs between approximately the 27th of July and the 15th of August in 2012. When this moment is compared with plots of the WALRUS run (Figure 4.2) it is clear that this peak in sensitivity is connected to the rainfall event that occurred around that time. Figure 5.2 shows the cumulative sensitivity of all the parameters. Because there is a certain degree of uncertainty regarding the correctness of the scale values, this plot is mainly useful for investigating the shape of the sensitivity peak. Figure 5.3 shows the fraction of the sensitivities of each parameter. It is important to realise that this plot shows the fraction of

the amount of sensitivity a parameter has at a certain moment, as compared to the sensitivity of the other parameters at that same moment. It is difficult to compare peaks in a plot like this. When one looks at the two peaks of sensitivity of c_S that occur during the recession period in Figure 5.2, one could say that this parameter shows the same amount of sensitivity at those two moments. However, when one looks at Figure 5.3 the second peak of c_S is larger than the first one. This because at the second moment a larger fraction of the total sensitivity is accounted for by the c_S . You thus cannot conclude from Figure 5.3 that at the second moment c_S shows more sensitivity than at the first moment.

The groundwater reservoir coefficient c_G is not visible in Figure 5.3. In Section 4.2 it is mentioned that the groundwater reservoir does not directly contribute to the discharge. The water balance (Table 4.2) shows that the f_{GS} is equal to zero. When one looks at Equation 2.4 in Chapter 2, one can see that the f_{GS} is, among other variables and constants, influenced by the c_G . Because the f_{GS} equals zero, this means that changes in the groundwater reservoir coefficient (c_G) do not influence the results. Because the c_G has a less important role, the c_V might have gotten a far more prominent role. The c_V influences the wetness index (Brauer et al., 2014a). The wetness index in turn determines how much goes through the quick flow reservoir. Because the soil reservoir is disabled, a large part of the precipitation goes via the quick flow reservoir. This also means that the c_V might become more sensitive to changes because it influences the wetness index, that controls the input of the now most important reservoir.

The drop in amount of sensitivity of the c_V , and the increase in sensitivity of the parameters c_W , c_Q and c_S , might be caused by the filling up of the reservoirs due to the rainfall event. The c_W seems to react slowly to the rainfall peak. This parameter has an important influence on the wetness index (see Equation 2.1 in Chapter 2, which has an increasingly more important role when it starts raining). c_Q mainly appears to be sensitive to the recession period of the rainfall event. When one looks at Equation 2.2 in Chapter 2, one can see that the c_Q influences the f_{QS} . Because of the importance of the quick flow reservoir in this case, it might not be that strange that the c_Q is sensitive to changes. Because the system might become satisfied just after the rainfall event, the c_Q can become more sensitive to changes. The quick flow reservoir has to hold the water longer because the surface water reservoir also fills up. This way it can avoid that the surface water reservoir becomes too full. This in turn avoids the discharge from becoming too high.

After the large rainfall peak a few smaller rainfall moments occur during the recession of the peak. When these moments are compared with Figure 5.3 it is clear that the c_S is sensitive to these kinds of moments. Like mentioned above, it is possible that due to the rainfall event, the surface water reservoir is completely filled. When smaller rainfall events follow the c_S becomes more important because this

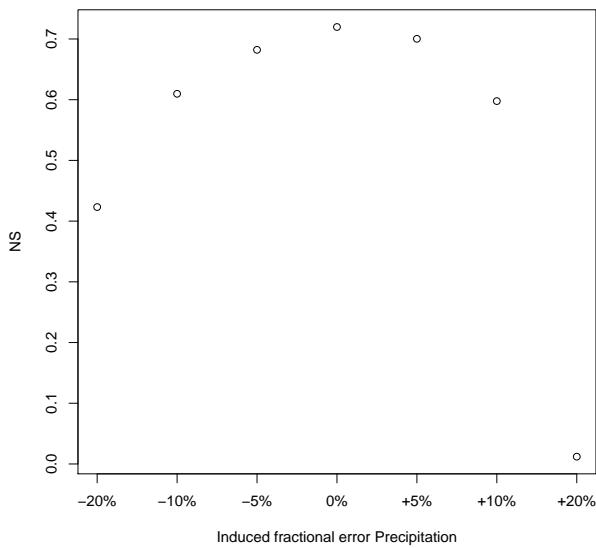


Figure 5.4: The NS plotted against the percentages of under- and overestimation for precipitation. The graph is clearly not symmetrical. The slope close to the zero percent on the negative side is smaller than the one on the positive side. When moving away from the zero, the slope clearly increases. This graph is the mirror image of the evapotranspiration induced fractional error graph. A reason for this is that WALRUS uses the effective precipitation, i.e. the precipitation minus the evapotranspiration.

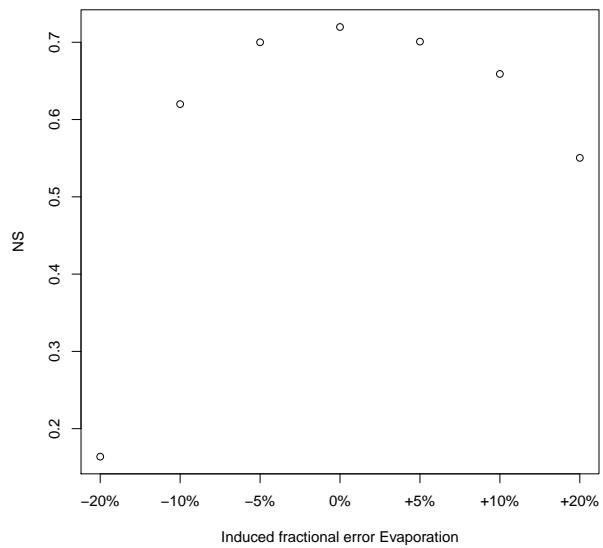


Figure 5.5: The NS plotted against the percentages of under- and overestimation for evapotranspiration. The graph is clearly not symmetrical. The slope close to the zero percent on the positive side is smaller than the one on the negative side. When moving away from the zero, the slope clearly increases. This graph is the mirror image of the precipitation induced fractional error graph. A reason for this is that WALRUS uses the effective precipitation, i.e. the precipitation minus the evapotranspiration.

parameter represents how much the surface water reservoir can handle.

The initial groundwater depth, d_{G0} shows a relatively constant amount of sensitivity when one looks at Figure 5.3. This can indicate that the initial groundwater depth is not directly linkable to the different process within the system, but exerts a certain amount of influence to the system as a whole.

5.2 Forcing sensitivity

The sensitivities towards the forcing give an indication on what the impact is of forcing uncertainties. For instance, when interception is not accounted for. The sensitivities can give an indication of where some weaknesses of the model lie, with respect to the model input.

5.2.1 Induced fractional error: methodology

Measurement devices are prone to (systematic) mistakes. For instance, with rain gauges this could manifest in systematic under or over catch. A rain gauge might be badly placed, thus not being fully representative for the catchment rainfall.

To investigate the effect of an error concerning forcing, runs with WALRUS will be performed in which the precipitation or the evapotranspiration

will be adjusted with a certain percentage. The selected percentages are five percent, ten percent or twenty percent over- or under-estimation.

5.2.2 Induced fractional error: results

Figures 5.4 and 5.5 show the results of the calculations that are made to investigate the effect of errors in forcing. The NS is plotted against the amount of under- and over-estimation of precipitation and evapotranspiration. When one looks at the results of the precipitation it is clear that the graph is not symmetrical. The slope close to the zero percent on the negative side is smaller than the one on the positive side. When moving away from the zero, the slope clearly increases. This graph is the mirror image of the evapotranspiration induced fractional error graph. A possible reason for this is that precipitation and evapotranspiration influence the same reservoir in WALRUS. Because the precipitation graph is the mirror-image of the evapotranspiration graph, the same observations hold for the evapotranspiration as for the precipitation.

5.2.3 Interception: methodology

WALRUS does not take interception of precipitation into account. To investigate the sensitivity of WALRUS towards the effect of interception, the precipitation part of the data will be adjusted to incorpo-

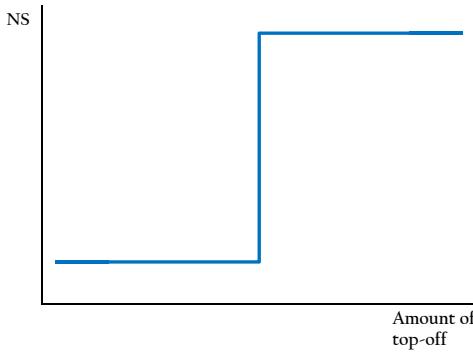


Figure 5.6: Example graph to illustrate the insensitivity of the 'top-off' parameter. When the interception script only contains if-statements, small changes in the 'amount of top-off'-parameter will only lead to (abrupt) changes in NS at specific locations. At other locations no changes in NS will occur. In such case, it can be difficult for the calibration algorithm to diverge from its starting value.

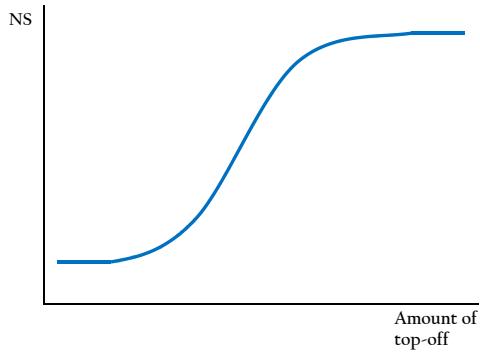


Figure 5.7: Example graph to illustrate the possible solution for the insensitivity of the 'top-off' parameter. When the interception scripts would consist of for instance sinus-like or power-law functions, small changes in the amount of top-off parameter would lead to (maybe small) changes in NS. The calibration algorithm would then possibly be able to find a more optimal value for this parameter.

rate a hypothetical interception. Two conditions are used for the adjustment of the precipitation dataset. Condition one (1) is that two days before the rainfall event it has to be dry. The second condition (2) is that the interception only will be subtracted from the first hour. 1) If it has not been substantially dry before the rainfall event the interception reservoirs might still be filled. Interception thus has no effect on a specific rainfall event. All the rainfall water will immediately flow out of the already full interception reservoirs. When these interception reservoirs are emptied (mostly through evaporation) interception can take place. It is assumed that all the reservoirs are emptied after two days of no rain. Another condition is that 2) the interception is only subtracted from the first hour. This is done purely for practical reasons. When one wants to spread the interception over the first two or three hours of a rainfall event a host of extra conditions (if-statements) need to be implemented. Like mentioned earlier, this can decrease the identifiability in the sensitivity analysis of the different effects interception will have. When less rain has fallen in one hour than the "top-off" amount, the rainfall for that hour will be set to zero.

The interception script, created to adjust the precipitation data, will be included in a new objective function. The parameter-set in the new objective function gets an extra parameter, namely the 'amount of top-off'-parameter. The value of this parameter represents how much precipitation is intercepted (i.e. topped-off of the rainfall peak), when both above mentioned conditions are met. The interception will thus be made part of the calibration process. With help of this calibration process it can be investigated if the 'amount of top-off'-parameter is sensitive. When the algorithm adjusts the parameter, it means that changes in the difference between observed and modelled discharge

occur, because of changes in the 'amount of top-off'-parameter. The amount of sensitivity of this parameter can say something about the influence of interception.

5.2.4 Interception: results

The basic script that adjusts the precipitation data for interception, is tested first. This script has the 'amount of top-off'-parameter as a constant. A few tests with this script indicate that including the interception script with a WALRUS-run can influence the results. When WALRUS is run with the fitted parameters and the amount of top-off is set on seven millimetres, the NS becomes 0.68. This means that the NS drops, when compared to a WALRUS run without interception (NS = 0.72).

When the calibration is performed the parameter for the amount of top-off does not diverge from its starting value. When the starting value is changed (for instance, from two millimetres to seven millimetres) the amount of top-off still does not diverge from its new starting value. However, the NS shows the same drop as with the tests of the basic interception script. The parameter is thus not changed during calibration, but the interception script can clearly influence the results. The fact that the calibration algorithm does not change the 'amount of top-off' parameter might be caused by local unsensitivity of that parameter. The interception script uses a lot of if-statements. These if statements can only be "on" or "off", meaning that small changes created by the calibration algorithm might not influence the results at all. Only when the 'amount of topoff'-parameter passes a certain threshold, something might happen. These thresholds are however very local.

This problem can be solved by changing the if-statements into functions with for instance sinus or power-law shapes. These functions have a certain width in which the system gradually changes from one state to the next when the amount of top-off is adjusted. This is in contrary to the if-statements, where the change can only happen abruptly, due to their "on" or "off" nature.

Figure 5.6 shows a cartoon version of a graph that could represent the change of NS, when the amount of top-off is increased when using only the if-statements. It is clearly visible that the change is very abrupt and small changes in the amount of top-off done in the larger part of the "horizontal" area will not produce any changes in NS. The calibration algorithm will thus not find a better NS when it is changing the 'amount of top-off'-parameter. Figure 5.7 shows how it could be when the if-statements are replaced by functions with a different shape (e.g. sinus, power-law). The change in NS is far less abrupt, because the area in which the change is taking place has a certain width. Small changes in NS would occur, when the calibration algorithm is changing the 'amount of top-off'-parameter. When these changes occur the algorithm will keep searching for values that will lead to a better NS. One of the drawbacks of this alternative for the if-statement, is that it introduces an extra parameter. This is because the width of the transition between one state and the next also needs to be defined. In general, over-parametrisation should be avoided.

A same sensitivity analysis like in Section 5.1 could be performed. However, because the interception already shows insensitivities through the calibration runs, this would not give any extra insight. The results would depend heavily on the starting values. If the starting values will be far away from a point where an abrupt change can occur, the sensitivity analysis will not show any sensitivity towards the 'amount of top-off'-parameter. Because this subject is not the main-topic of this master thesis and due to time-constraints, these functions will not be implemented.

5.3 Conclusion

Sensitivity analysis is needed to understand how parameters react to changes, and thus possibly to the changes induced by the calibration algorithm. However, the used calibration approach in combination with this sensitivity analysis, is clearly not sufficient on its own to identify optimal parameter-sets. The calibration approach thus has to be expanded.

The c_V , c_Q and c_S often showed the largest contribution to the total sensitivity at specific moments. The c_W was relatively constant.

Apparently it is difficult for the Levenberg-Marquardt algorithm to diverge from a boundary when the start value equals a boundary value. The Levenberg-Marquardt algorithm can find different values, when the start value is manually put far enough away from the boundary value.

There are methods to investigate the errors regarding the forcing of the model. In this case it is clear that the speed with which the NS drops, increases when the induced fractional error increases linearly.

The research into the influence of interception showed that it can change the objective value of a WALRUS simulation. The interception script would have possibly functioned better when the if-statements were replaced with for instance a power-law function. The drawback of such a function however is that it introduces an extra parameter.

6 | Regularisation

The initial calibration in Chapter 4 shows that it is difficult to find efficient values using the Levenberg-Marquardt routine. The calibration algorithm will search for parameter values that make the residuals as small as possible, sometimes regardless of physical realism. The question is, in which part of the numerical processes lie the problems. As stated in Section 4.3, multiple calibration runs lead to the same parameter values. This means that the algorithm appears to be stable in its performance. The problem might be caused by WALRUS itself. However, it is not in the scope of this Master thesis to investigate if this is intrinsic to these type of models or that it is a solvable issue within WALRUS. There is however a way to solve this problem without adjusting the WALRUS script. This actually means that the problem can be solved by applying Occam's Razor. The solution that is applied, is based on the least assumptions (the solution is not based on assumptions about the model or the calibration algorithm). The method is called regularisation, which is the main topic of this chapter.

6.1 Regularisation as boundaries for the parameter values

The idea of imposing hard boundaries, as done with the initial calibration, is that the most physically absurd values are excluded from the process. However, imposing hard boundaries on the calibration process can have its problems. See Section 4.2 for a more elaborate explanation on the physical problems regarding imposing hard boundaries. There is also a numerical problem. The two different sets of nineteen calibration runs, described in Section 5.1.2, have brought this to light. Apparently the Levenberg-Marquardt calibration algorithm can have difficulties moving away from a boundary once it has hit that boundary. Like mentioned above, regularisation might be able to solve this problem. A soccer field with an invisible goal can be used as an analogy to describe the difference between imposing hard boundaries and regularisation. One's target is to hit the invisible goal of which the location is of course unknown. The direction in which one thinks the ball has to be kicked, is like the algorithm its way of determining which direction it should search. Imposing hard boundaries is like putting a fence around the field. This makes sure that when one kicks the ball in the wrong direction, it doesn't leave the field. This increases the chance that one will eventually hit the invisible goal. Regularisation is like attaching an elastic string between the ball and the area in which one thinks the goal might be. One's idea of where the invisible goal is situated, is based on a priori information. When one kicks the ball, it still has a certain degree of freedom in its movement. However, the elastic string does not allow the ball to stray too far away from the expected location. This is because then the pulling

force on the ball would become too high. This idea of putting a fence around a field versus exerting a certain pulling force on the ball towards the suspected location of the goal, can be seen as the main difference of hard boundaries (box-constraints) versus regularisation. Regularisation can thus be seen as imposing soft boundaries. In this case regularisation will be used to overrule or influence the numerical processes with help of a priori information on the possible optimal parameter values. The a priori information gives an idea of where the invisible goal on the field might be and thus where the ball approximately should end up.

To describe it in more technical terms, with regularisation an extra term is added to the objective function. In this regularisation term the approximate expected parameter value is given (which is depended on the a priori information) and a weight is assigned to this expected parameter value. When the calibration algorithm tries to increase a certain parameter value with a large weight, it is then possible that the objective function value quickly becomes too high. The calibration algorithm will then search for a value closer to the expected parameter value.

The total amount of contribution of a regularisation term to the objective function, can be seen as the amount of trust that is put in the a priori information. An example of using regularisation in hydrology, can be found in the article of Wiese and Nützmann (2011). For a more technical description of regularisation in inverse modelling, and the different choices that can be made with respect parameter choices, I would like to refer to the paper of Bauer and Lukas (2011).

6.2 Methodology

The objective function is changed for the regularisation process (Equation 6.4). For the initial calibration a vector containing the differences between the observed discharge and the simulated discharge is used as objective function (see Section 4.2). Equation 6.1 shows the original objective function for the Levenberg-Marquardt calibration algorithm.

$$OF = \sum_{t=1}^T (Q_O^t - Q_M^t)^2 = \sqrt{SS} \quad (6.1)$$

where OF stands for Objective Function. Q_O^t is the observed discharge at moment t and Q_M^t is the simulated discharge at moment t . It is clear that this objective function is equal to the square root of the sum of squares (see Equation 4.1 for the sum of squares). To make the values of the new objective function better interpretable, Equation 6.1 will be made similar to the equation for the NS (Equation 4.2). In Section 4.1.5 it is explained that part of the nominator of the fraction term in Equation 4.2 is equal to Equation 4.1. This also means that the

Table 6.1: The centre-parameter, lower- and upper-boundary values used for the six parameters during regularisation. The centre-parameters are roughly based on the values Brauer et al. (2014b) found for the Hupsel-catchment. The values are rounded off, because they only need to be approximations. The Levenberg-Marquardt algorithm still has to be able to find its own value. The lower- and upper-boundary values are a bit different with respect the ones used during the initial calibration (Section 4.2). This is due to increased insight in where the optimal parameter values might be.

	c_W [h]	c_G [mm]	c_V [h]	c_Q [mm h]	d_{G0} [mm]	c_S [mm/h]
Center-parameters	350	5e6	1	3	2000	0.08
Lower-boundaries	1	3e6	0.1	0.01	100	0.01
Upper-boundaries	721.0	273e6	36.1	72.0	4600	18.1

nominator is similar to Equation 6.1. The first step in making these formulas equal, is to square the difference of the observed discharge values and simulated discharge values in Equation 6.1. This will also avoid problems with possible negative values. Equation 6.1 will be divided with the variance of the observed discharge values. This way the objective function becomes dimensionless. This makes the formula better comparable with the regularisation terms that will be added later on (which will also be made dimensionless). The variance by which Equation 6.1 will be divided, is simply the mean of the denominator of the fraction term in Equation 4.2. Because we want to make Equation 6.1 similar to Equation 4.2, the mean then also has to be taken from the nominator in the new objective function. Equation 6.2 shows the normalised objective function that will be used for the regularisation:

$$\text{Norm_OF} = \frac{\frac{1}{N} * \sum_{t=1}^T (Q_O^t - Q_M^t)^2}{\text{VAR}(Q_O^t)} \quad (6.2)$$

where the N equals the length of the dataset that is used for the calibration run. The only difference now between Equation 4.2 and Equation 6.2 is that the first term from which the fraction is subtracted is left out of the new objective function. With the objective function written in this way, the values are more comparable with the NS.

With the objective function rewritten, the regularisation terms can be added. Equation 6.3 shows the regularisation function:

$$\text{RF} = \alpha * \sum_{i=1}^I \frac{(p_{i,\text{mod}} - p_i^*)^2}{\text{VAR}(p_{i,\text{runs}})} \quad (6.3)$$

where RF stands for regularisation function. The $p_{i,\text{mod}}$ is the parameter value the calibration algorithm finds for each iteration. $p_{i,\text{runs}}$ are the parameter values of the parameter in question, found with multiple calibration runs with a moving time-window (similar to the nineteen calibration runs, see Section 5.1.2). The variance of $p_{i,\text{runs}}$ represents the amount of variation a parameter value can have in a certain area. The larger this variation is, the smaller the regularisation penalty will be. This way the calibration algorithm will have a larger range in which it can search for a fitting value. When the variation is small, the regularisation penalty will be large. This will force the calibration algorithm too converge closer to the suspected

parameter value. In this case, the parameter values that are used for $p_{i,\text{runs}}$ are the same ones as obtained in Section 5.1.2 and written down in Table 5.2. α and p_i^* are the values that have to be selected before the calibration process starts. The α is the regularisation parameter. This means that for all the different parameters within the summation of Equation 6.3, there is only one regularisation term. The p_i^* is the parameter specific centre-parameter. This is the value around the calibration algorithm should find the parameter value. Both α and p_i^* depend on a priori information. The value for α could be interpreted as the amount of uncertainty with respect to the a priori information available for the centre-parameters. When the uncertainty about the a priori information is high, the α will have a low value and when the uncertainty is low the α will have a high value. The values for p_i^* depend on the a posteriori information of an earlier research. Equation 6.4 shows the new objective function:

$$\text{New_OF} = \text{Norm_OF} + \text{RF} \quad (6.4)$$

where one can clearly see that the new objective function consists of the normalised objective function and the regularisation function.

This regularisation method is Bayesian in nature, meaning that the relationship between the a priori and a posteriori information is used to find the right values.

Table 6.1 shows the used centre-parameters, lower- and upper-boundary values. The centre-parameters are based on the values found for the Hupsel-catchment in the research of Brauer et al. (2014b). The values are rounded-off, this is because they only have to be approximations. Levenberg-Marquardt still has to find its own values. The lower- and upper-boundaries used for the regularisation are a bit different than the ones used for the initial calibration. The lower-boundary of c_V is lower than it first was. This is because during the initial calibration Levenberg-Marquardt hit that specific boundary several times. This might indicate that the boundary value was wrong and that the real value might be lower than the lower-boundary. Levenberg-Marquardt will be run several times with different values for the regularisation parameter α . This way the effect of the amount of contribution of the regularisation term to the objective function can be investigated.

Table 6.2: Eleven different regulated calibration runs that are performed with different α values. From α values 2 to $1e-2$ the regularisation term dominates the output of the objective function. During calibration all the parameter-values are eventually pulled towards the centre-parameters. From α values $1e-5$ to $2e-7$ the algorithm is able to find values within the hard boundaries (except for $\alpha=5e-7$), not too far away from the centre-parameters. From α values of $1.5e-7$ and smaller the regularisation term has too little influence. For the α values of $1.5e-7$ and $1e-7$ the algorithm hit one boundary (for c_V). For an α value of $8e-8$ the algorithm hit three boundaries (for c_G , c_V and c_Q). The parameters that hit a boundary are printed in bold letter type. The α value of $8e-8$ also leads to the lowest NS of all the runs.

α	NS	c_W [h]	c_G [mm]	c_V [h]	c_Q [mm h]	d_{GO} [mm]	c_S [mm/h]
2	-0.31	349.7	4.9e6	1.0	3.0	1996.8	0.09
0.75	-0.31	349.6	5.3e6	1.0	3.0	1996.9	0.08
0.25	-0.31	349.5	5.0e6	1.0	3.0	1997.0	0.08
$1e-2$	-0.31	349.5	5.1e6	1.0	3.0	1997.0	0.10
$1e-5$	0.08	391.4	11.2e6	0.5	0.2	1936.6	2.14
$1e-6$	0.63	412.1	15.8e6	0.8	4.2	1641.2	1.64
$5e-7$	0.67	415.4	19.0e6	0.1	4.2	1611.5	1.29
$2e-7$	0.68	413.0	21.1e6	0.2	4.2	1570.9	1.22
$1.5e-7$	0.69	417.2	43.8e6	0.1	3.6	2189.4	0.97
$1e-7$	0.69	412.0	60.2e6	0.1	4.2	2186.2	1.17
$8e-8$	-0.89	385.2	276e6	0.1	72.0	1227.9	0.02

6.3 Results

6.3.1 The influence of the regularisation parameter

First we will shortly look on the influence of the different α values. Table 6.2 shows the results of different calibration runs, each with a different value for the regularisation parameter α . The first few runs with α 's ranging between 2 and $1e-2$ show that the regularisation term is too dominant in the objective function. The found parameter values are almost equal to the centre-parameters. For the α values of $1e-5$ and $1e-6$ Levenberg-Marquardt is able to find more or less its own values without hitting any boundaries. Table 6.2 shows that for an α value of $5e-7$ one boundary is hit. For an α of $2e-7$ no boundaries are hit. Of all the runs, this alpha value is the highest without having fitted parameter values that equal boundary values. When the α value is lower than $2e-7$, the influence of the regularisation term on the objective function significantly drops and Levenberg-Marquardt hits the boundaries again.

6.3.2 The obtained parameter values

The c_G seems to increase when the influence of the regularisation term drops. The less influence the regularisation has, the more the c_G deviates from the value found in the research of Brauer et al. (2014b). The parameter c_V eventually also hits a boundary. The lowering of the lower-boundary did not seem to help on its own to avoid that the parameter hit a boundary. The results obtained with the calibration process with the least influence of the regularisation term, show that at least for c_V , c_G and c_Q regularisation is needed.

6.3.3 Validation

The NS of the best result that did not hit any boundaries, is lower than the NS found with the initial calibration. The NS found for $\alpha = 2e-7$ is 0.68 while the NS found in Chapter 4 equals 0.72. The big difference between these two is that when the parameter-set is obtained with a regulated calibration process, the values are closer to the ones found in the research of Brauer et al. (2014b). The c_G and the c_V are considerably lower than the unrealistic values that are found with the initial calibration.

The validation run done for the fitted parameters of the initial calibration yielded NS's of respectively 0.83 and 0.44 (see Chapter 4). When a validation run is performed for the fitted parameters that are obtained with regularisation ($\alpha=2e-7$), NS's of 0.81 and 0.77 are obtained for the periods in respectively 2011 and 2013. Figure 11.6 and Figure 11.7 in the appendices show the plots that the WALRUS R-script created of these runs. Like mentioned in Chapter 4, the first months of 2011 are dry. The period in 2013 is considerably wetter than 2011, but also has a short dry period in the beginning of spring. The used calibration period (in 2012) is relatively wet. There are thus clear differences between the periods. The results for the validation runs of WALRUS with regulated parameters are in general better than the results for the fitted parameter-set in Chapter 4. This is because the NS's stay high throughout the different periods.

6.4 Conclusion

Better results can be obtained for the Hupsel catchment when the calibration approach from Chapter 4 is expanded with regularisation. It is important that good a priori information on what the possible optimal parameter values are, is available. The computation speed also increased, possibly because boundaries are not hit during the calibration process.

7 | Calibration and Validation

In this chapter the method and results of the calibration process for the four catchments will be discussed. See Chapter 3 for a more elaborate explanation of the properties of the catchments and Chapter 4 and 6 for a more elaborate explanation of the calibration method.

Throughout this research it is shown that for the used calibration approach, information about the possible optimal parameter values is needed. This information is necessary for determining good starting values and boundary values, but is also needed for regularisation.

The goal of this chapter is to validate the calibration approach on the other catchments. First the calibration approach will be expanded with a multi-start method. This method is needed to detect possible local minima in the other catchments. After this, WALRUS will be applied on two different categories of catchments. The first category represents a catchment with good information on the possible parameter values, meaning that there is good information for the implementation of lower- and upper-boundary values, and for the centre-parameters. The second category represents the catchments which have low quality a priori information. In Table 7.11 the best obtained parameter-sets are shown for both categories. To investigate the seventh research question, WALRUS simulations are coupled in this Chapter.

7.1 Methodology

As mentioned above, three types of calibration runs are discussed in this chapter. 1) The calibration of the Stadtlohn catchment. This calibration runs is performed to see how WALRUS performs with good a priori information of the possible optimal parameter values. The a priori information for the Stadtlohn catchment is available via the Hupsel catchment data, which originates from Brauer et al. (2014b). 2) For the Diesdonk, Lieveren and Tun-gelroyse beek catchments, good a priori information about the possible optimal parameter values is lacking. The obtained parameter-sets from the Stadtlohn catchment will be used as a priori information. The Cabauw parameter-set, originating from the research of Brauer et al. (2014b), will also be used as a priori information. 3) Lastly, the sub-catchments of the Berkel will be calibrated from the most upstream sub-catchment (Lutum) to the most downstream sub-catchment (Lochem) and will be coupled with each other, without introducing a hydraulic component.

First an addition to the developed calibration approach will be discussed.

7.1.1 Multi-start method

It is mentioned in Section 4.1.3 that a multi-start method for Levenberg-Marquardt is useful to avoid

possible local minima. Local minima can be caused by the available data and/or by the non-linear structural characteristics that are typical for conceptual rainfall-run off models (Duan et al., 1992). Of the earlier used Hupsel-data, it was already known that when this data is used in WALRUS, there are no real local minima (Brauer et al., 2014b). Thus for this part of the research it was not necessary to adopt a multi-start approach. WALRUS will now be applied to different catchments for which it is not known if local minima can occur. For the calibration of the four catchment a multi-start approach will thus be adopted.

With a multi-start approach a sampler is used to create different starting values for one catchment. For all these starting values calibration runs will be performed. When there are two or more clearly discernible groups in the results, this might indicate that there are more minima. When there is clearly one group, and maybe a few - not with each other related - other results, the possibility that there is only one (global) minimum significantly increases.

For this research the evolutionary optimisation algorithm hydroPSO will be used as sampler. hydroPSO uses particles to find an optimal value. The algorithm is applied in largely the same way as the Levenberg-Marquardt algorithm. hydroPSO needs starting values and lower- and upper-boundary values defined before-hand. The starting and boundary values will be based on the information obtained with the 'rough-runs'. The parameters will also be scaled, so that the convergence process will be sped up. The lower- and upper-boundary values obtained with the 'rough-runs' will be used as indicators for the amount of scaling. The algorithm will be limited to the use of 10 particles and will stop converging after 10 iterations. This has as benefit that there is a more random-like method of finding starting values, reducing the effect of possibly manually choosing wrong starting values. The generally slow hydroPSO algorithm will not be run for too long due to the fact that it is limited to 10 iterations, it will stop before it has converged. After these 10 iterations the best found parameter values per particle (the so-called *pbest*) will then be used as starting values for the regulated Levenberg-Marquardt runs. This means that Levenberg-Marquardt will be run 10 times, with different starting values.

7.1.2 Calibrating with good a priori information: Stadtlohn

Due to the fact that there is not a lot of information yet available on the parameter values for WALRUS, the catchment closest to the Hupsel will be calibrated first. The idea behind this is that such a catchment might share the most (geo-)hydrological characteristics with the Hupsel-catchment.

The Stadtlohn catchment is selected because it is the second most distant sub-catchment of the Berkel, when compared to the Hupsel. This way,

the catchments will possibly not be too identical in characteristics. There still can be some (geo-)hydrological differences between the catchments, thus allowing the calibration algorithms the possibility to find more optimal values. The Stadtlohn catchment is selected over the Lutum sub-catchment due to the small size of the Lutum catchment (38 km^2). Because the data of the calibrated Stadtlohn catchment (193 km^2) will be used for the Diesdonk (approximately 260 km^2), Lieveren (approximately 106 km^2) and the Tungelroyse beek catchments (152 km^2), it is at least practical to pick a catchment that in size can be placed between these catchments. This is important because some parameters might also be depended on the catchment size. A parameter like for instance the quick-flow parameter, might also be influenced by the size of the catchment. The Hupsel parameter-set will thus be used as a priori information for the possible parameter values: as starting values for the 'rough-runs' and as centre-parameters for the regularisation.

The parameters

In Chapter 3 it is mentioned that the Berkel catchment is mainly dominated by silty sands. The WALRUS soil type that comes closest to this type is "sand". For the channel depth (c_D) and for the surface water area fraction (a_S) the same values as the ones used for the simulation of the Hupsel-catchment will be used.

The calibration and validation periods

Two years are selected from the dataset for calibration and validation. When compared to the other available data, which spans over a period of six years, these years look quite average. The year of 2008 will be used for calibration. The starting period will be the first of May 2008. Officially the hydrological year starts at the beginning of April. However, at the first of April the system still seems to be moving out of the winter period. The beginning of May is thus selected because around that time the system seems to be in a sort of equilibrium. The calibration period ends the next year at the beginning of May. The validation period is from the first of May 2009 till the end of the fourth of May 2010. Because the data contains a small gap around the first of May, the fourth of May 2010 is selected as end date. WALRUS uses interpolation to fill gaps in the discharge data, and thus does not permit the start date or end date to start in a gap.

7.1.3 Calibrating with low quality a priori information: Diesdonk, Lieveren, Tungelroyse beek

For the Diesdonk, Lieveren and Tungelroyse beek catchment the obtained data from the Stadtlohn catchment will be used as starting values for the 'rough-runs' and as centre-parameters for the regularisation. For the centre-parameters, the mean of the 10 obtained parameter-sets will be used.

Similar runs will be performed with the Cabauw parameter-set as a priori information. This way the catchments in question will be calibrated with data originating from a freely draining lowland catchment (Stadtlohn) and from a completely regulated polder area (Cabauw). The Stadtlohn catchment and the Cabauw catchment thus vary from each other, which should lead to differences in results for the same catchment. When a catchment performs well with the Cabauw parameter-set as centre-parameters, the catchment in question might share certain hydrological characteristics with that catchment. This can tell something about how the optimal parameter-set for that catchment at least should look like. Using different a priori information for one catchment can also shed some light on the dependency of WALRUS on a priori information.

The parameters

The Diesdonk catchment mainly contains fine sands. Because of this, the WALRUS input field *st* will be assigned the WALRUS soil type "sand". In the Lieveren catchment mainly loamy fine sands can be found. This means that for the *st* input field the WALRUS soil type "loamy sand" will be used. The Tungelroyse beek catchment is mainly dominated by sand soils. There is however some loam and clay present. Because of this the "loamy sand" WALRUS soil type will be assigned to the *st* input field. The fixed values (c_D and a_S) that are used for the Stadtlohn catchment, will also be used for the other three catchments.

The calibration and validation periods

The period of available data for the Diesdonk is two years. The first year, from the first of april (start of the hydrological year) 2010 till the first of april 2011 will be used for calibration.

The period of the available data for the Lieveren catchment spans only over one year (2011). The data set contains quite a big hole at the beginning of May. The calibration period spans from May the thirteenth 2011 (just after the hole in the data), till November the thirtieth 2011.

For the Tungelroyse beek two years of data are available. Like the Diesdonk data, the first year will be used for calibration and the second for validation. The calibration starts at the first of May 2011 till the first of May 2012.

7.1.4 Coupling of WALRUS models: Berkel sub-catchments

A catchment like the Berkel catchment is quite large (849 km^2). It might be the case that it difficult to simulate when only using WALRUS and no extra hydraulic component. The heterogeneity of the catchment might have to much influence on the results, which could lead to parameters that only represent heavily averaged processes. When more sub-catchments are used, WALRUS can find more optimal values for smaller areas. The quality of the found parameter-sets might be higher then when WALRUS would have tried to find parameter-sets

for the complete Berkel catchment. Besides the increase of quality of the simulated results, it also gives the opportunity to investigate the behaviour of the hydrological system in more detail. Speaking in water management terms, this could lead to more effective, local solutions. For the coupling of the catchments the (f_{XS}), will be used as explained in Section 2.1. The multi-start method will not be applied for these calibration runs.

The parameters

For the starting values of the Levenberg-Marquardt algorithm and for the centre-parameters for the Lutum sub-catchment the earlier found values of the Stadtlohn catchment calibration will be used. For the centre-parameters for the Stadtlohn sub-catchment, the values found for the Lutum sub-catchment will be used. The same will be done for the Rekken sub-catchment, only then the found values of Stadtlohn sub-catchment will be used as centre-parameters. For the Haarlo and Lochem sub-catchment the Hupsel parameter-set will be used as centre-parameters. This is because the Hupsel catchment is situated close to these two sub-catchments. The quality of the a priori information might thus be possibly better.

The calibration period

The same calibration period as used for the Stadtlohn catchment will be used for these calibration runs.

7.1.5 Regularisation

For the regularisation an α of $1e-6$ will be used. For the Hupsel catchment an α value of $2e-7$ proved to be the best one. Because this value was clearly a border-case, i.e. with the higher value of $5e-7$ a boundary was hit, the value of $1e-6$ is selected to be on the safe side. This means that the regularisation term might have more influence on the objective function than is the case with the best regulated results of the Hupsel as described in Chapter 6. When too many boundaries are hit with one calibration run (meaning 10 Levenberg-Marquardt runs) the α value might be increased, so that the regularisation term has more influence on the objective function.

7.2 Results

7.2.1 Calibrating with good a priori information: Stadtlohn

Table 7.1 show the 10 regulated Levenberg-Marquardt runs, with their respective fitted parameter-sets. Table 11.1, that can be found in the appendices, shows the validation run done for all the 10 parameter-sets.

As one can see in Table 7.1 WALRUS is able to perform reasonably well with the obtained parameter-sets. The highest obtained NS equals 0.62. A total of six parameter-sets have NS val-

ues above the 0.6. The parameter values of these parameter-sets show clear similarities. This indicates that for the regulated calibration algorithm there is a clear global minimum. The other obtained parameter-sets, the ones that have a NS lower than 0.6, do not show similarities between their respective parameter-sets. There is thus no clear indication for a second local minimum. The fact that relatively high NS's are obtained, can indicate that the a priori information is of good quality.

The NS's that are obtained with the validation runs are shown in Table 11.1. When looking at the values it is clear that all the obtained parameter-sets perform well. All the NS's are above the 0.7.

Table 7.2 shows the water balance of the regulated Levenberg-Marquardt run with the highest NS (0.62). Figure 7.1 shows the plot that WALRUS created for this parameter-set. The flux from the groundwater to the surface water is negative. This is clearly visible in the water balance and in the bottom window of Figure 7.1. This means that the direction of the flux is actually from the surface water reservoir to the ground water reservoir. This is the case for all the 10 regulated Levenberg-Marquardt runs. This negative flux is unrealistic. Although the height difference in the German area is larger than the height differences in the Dutch area — thus some differences are to be expected — both areas share a lot of characteristics. The Hupsel area, which is situated close to the Berkel catchment, also does not show this trend. An explanation for this is that both the channel depth (c_D) and the initial groundwater depth have the same starting values (1500 mm) for the calibration process. During the calibration, the d_{G0} apparently gets a higher value than the c_D . This is clearly visible in Table 7.1. The d_{G0} for the final WALRUS run is thus lower than the c_D . This means that the model starts with the groundwater depth situated below the river bottom. WALRUS is not able to get the d_G above this river bottom, thus producing f_{GS} that is negative for the complete year.

The influence of the d_{G0} parameter value

A new run is performed with the same parameter-set, only now the c_D is changed to 2500 mm. This means that this value is higher than the fitted value of the d_{G0} (1807 mm). Figure 7.2 shows the results of this run. The flux from the groundwater reservoir to the surface water reservoir is now clearly not always negative. This is underlined by the water balance of this new run (Table 7.3), the f_{GS} is not negative. The NS also significantly increased, from 0.62 to 0.74.

To test the effect of different c_D values, a few extra runs are performed. One run gets a c_D that is closer (but still higher) to the initial groundwater depth, 2000 mm. Two other runs get a higher value than the c_D of the second run for the Stadtlohn catchment, respectively 3000 mm and 3500 mm. The NS's for these runs are shown in Table 11.2, which can be found in the appendices. You can deduct from this table that there is a optimum value for the c_D . You cannot say however, that this optimal

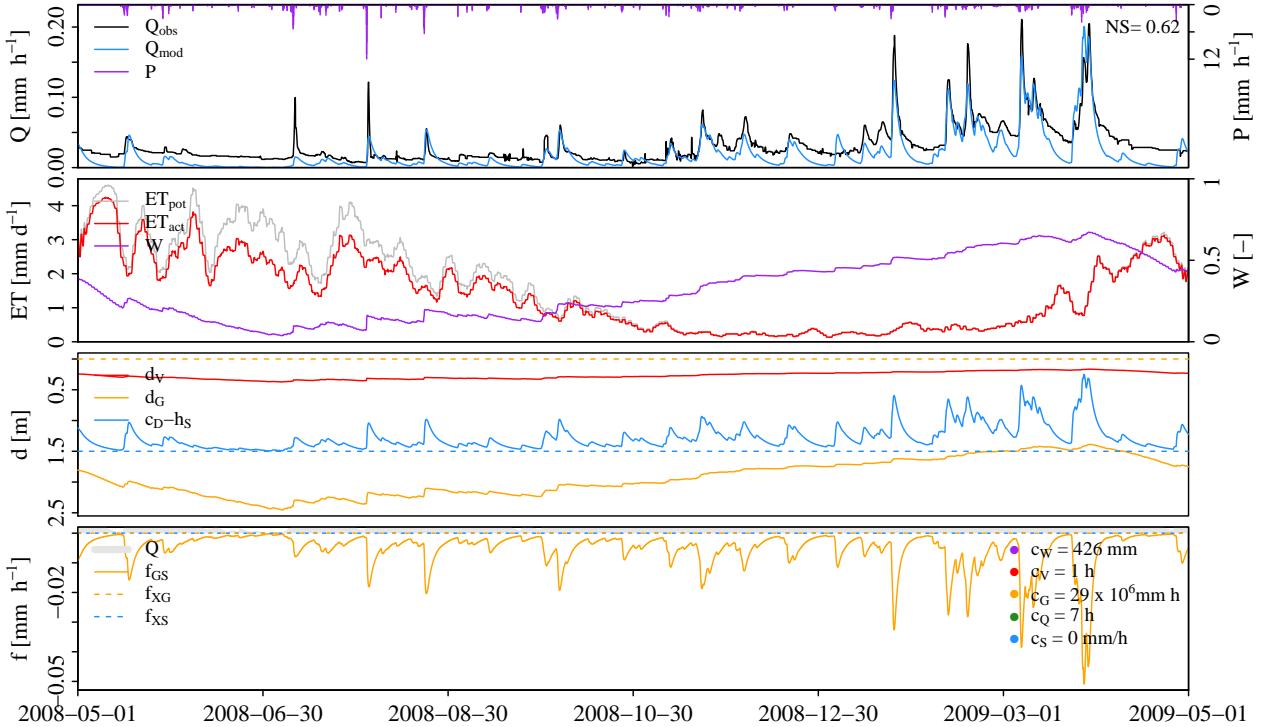


Figure 7.1: The model run for the Stadtlohn catchment with the fitted parameters that produced the highest NS (0.62) without hitting any boundaries. The simulated discharge seems to follow the observed discharge quite well, judging from the top window. However, the f_{GS} in the bottom window is negative for the whole year. Physically this is not realistic for this area. The negative f_{GS} is caused by the fact that the calibrated d_{G0} has a higher value than the c_D , i.e. the groundwater level starts beneath the bottom of the surface water reservoir. Possibly because of this initial condition the groundwater depth will not rise above the bottom of the surface water reservoir. Figure 7.2 shows the results of a run were the initial groundwater depth value is smaller than the channel depth value.

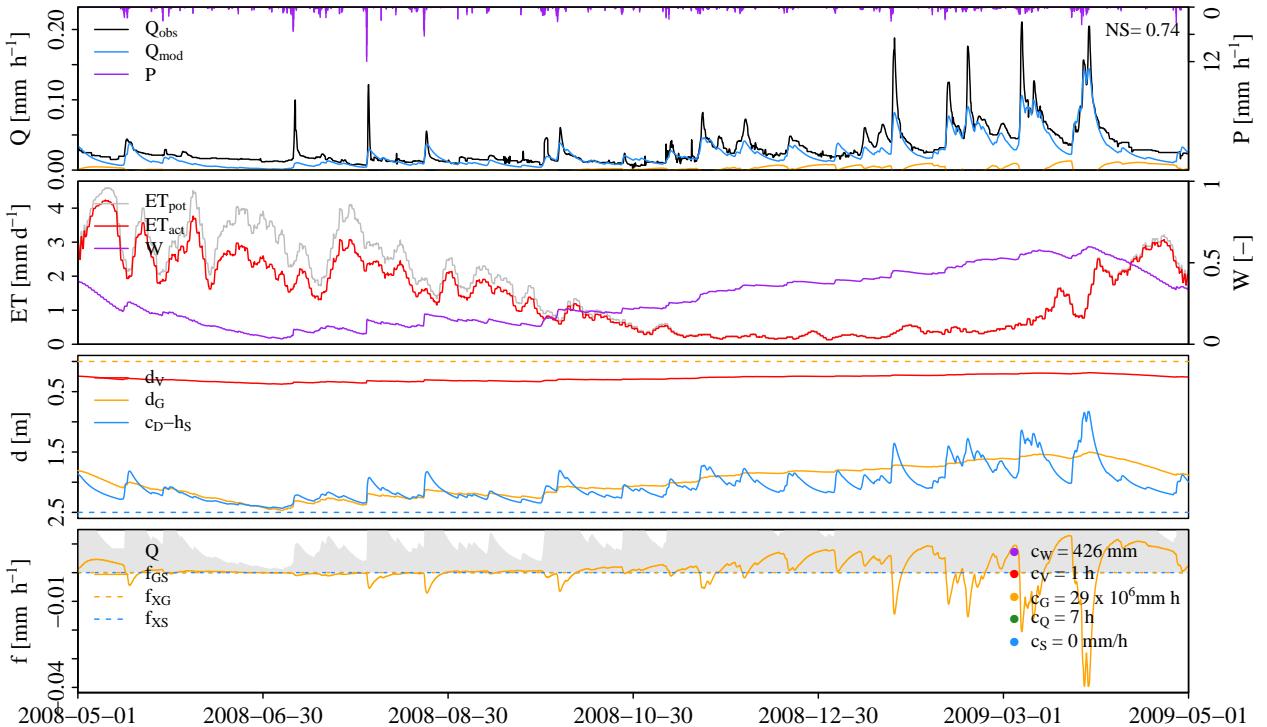


Figure 7.2: The model run for the Stadtlohn catchment with the fitted-parameters and adjusted d_{G0} . For this WALRUS run the d_{G0} has a smaller value than the c_D . The f_{GS} in the bottom window is not negative for the whole year any more. Physically this is far more realistic for this area. The NS also significantly increased, from 0.62 to 0.74.

Table 7.1: The obtained parameter sets of 10 regulated Levenberg-Marquardt runs ($\alpha=1e-6$) for the Stadtlohn catchment. The used start values are the best particle values of the hydroPSO sampler. For the regularisation, the Hupsel parameters are used as centre-values. It seems that there is a clear global minimum, because their is only one clear discernible group with similar results. When looking at more than three decimals, run number eight has the highest NS. There are no parameters that equal their boundary values.

Run	cW	cG	cV	cQ	dG0	cS	NS
1	415.7	650649	0.07333	11.261	1744	0.3715	0.52
2	426.0	28567226	0.65846	7.068	1811	0.264	0.62
3	405.1	2175968	0.10165	14.051	1699	0.5243	0.47
4	427.5	28456234	0.76295	7.139	1819	0.2653	0.62
5	426.9	27298198	0.65484	7.110	1814	0.2656	0.62
6	422.0	18433463	0.07623	7.897	1772	0.2803	0.59
7	424.7	17704151	0.52707	7.754	1785	0.2786	0.59
8	425.7	28580918	0.66354	7.037	1807	0.2641	0.62
9	426.9	27327347	0.54126	7.120	1816	0.2658	0.61
10	419.2	7739965	0.06925	10.243	1763	0.3462	0.52

Table 7.2: The water balance for the Stadtlohn catchment. This water balance is from the regulated Levenberg-Marquardt run with the highest NS (0.62). The external fluxes are zero because no external fluxes were given before hand. The flux from the groundwater reservoir to the surface water reservoir (f_{GS}) is negative. Physically this is not realistic.

P	ET_{pot}	Q_{obs}	f_{XG}	f_{XS}	ET_{act}	Q	f_{GS}	f_{QS}	d_V	d_G	h_S	h_Q	n_{days}	check
692	595	269	0	0	512	170.0	-46.1	214	10.7	56	-1.1	-0.2	365	0

Table 7.3: The water balance for the Stadtlohn catchment with a higher c_D . When the c_D is higher than the d_{G0} it clearly influences the results. Contrary to the first run, the f_{GS} is now positive.

P	ET_{pot}	Q_{obs}	f_{XG}	f_{XS}	ET_{act}	Q	f_{GS}	f_{QS}	d_V	d_G	h_S	h_Q	n_{days}	check
692	595	269	0	0	506	202	10.7	189	-14.5	-74.9	-1.3	-0.2	365	0

value lies exactly between 2500 mm and 3000 mm. Or in other words, when the c_D value would be plotted against the NS, one cannot assume a parabolic shape.

Excluding the d_{G0} parameter from the calibration process

The whole calibration process for the Stadtlohn catchment was repeated in order to investigate the effect of the exclusion of the d_{G0} parameter from the calibration process. For this run however, the d_{G0} is excluded from the calibration process. This means that this parameter value will be calculated by the default method in WALRUS, as described in Chapter 2. The c_D gets a value of 2500 mm, this is partly based on the results of Table 11.2. In Table 7.4 the results are shown. The highest obtained objective values are considerably higher than those of the first set of calibration runs. When compared to the first runs, the NS's from the second runs seem to vary a bit more. This can be explained by the starting values hydroPSO created in combination with the possibility that the Levenberg-Marquardt algorithm might have selected a wrong starting direction. The calculation process of hydroPSO is stochastic in nature. This means that every time one would run the algorithm anew, one would not get exactly the same results (which would be the case with a deterministic algorithm). This means that due to the used multi-start method, it is difficult to com-

pare the first set of runs with the second set. Due to the form of the parameter-landscape, Levenberg-Marquardt might have selected a direction that over a small change seems to lead to the best improvement of the objective value, but over a series of steps does not lead to the global minimum. The convergence direction Levenberg-Marquardt chooses depends on the starting values. Table 7.5 shows the water balance of the run with the highest NS (0.8205). When the simulated f_{GS} and the f_{QS} values are compared with those of Table 7.3, the values of this water balance seem to indicate that the groundwater reservoir plays a larger role.

Coincidentally, the eight parameter-sets of both sets of calibration runs are the ones that lead to simulations with the highest objective value. The c_G , c_V and c_S clearly have different values. Table 11.3, which can be found in the appendices, shows the validation runs for the second set of calibration runs of the Stadtlohn catchment. Just like the NS's of the calibration runs, the NS's of the validation runs vary a bit more than those in Table 11.1. The parameter-sets which have the best objective value (the ones with NS's above 0.8) also have good results for the validation (NS's above the 0.7).

Table 7.4: The obtained parameter-sets of 10 regulated Levenberg-Marquardt runs ($\alpha=1e-6$) for the Stadtlohn catchment, with the d_0 excluded from the calibration process. The used start values are the best particle values of the hydroPSO sampler. The parameters that equal a boundary value are printed in bold font. For the regularisation, the Hupsel parameters are used as centre-values. For all the 10 runs a boundary is hit (c_V and c_S). The NS's are generally better than those of the calibration process which included the d_{G0} .

Run	cW	cG	cV	cQ	cS	NS
1	363.7	9117594	72.2000	4.2424	0.4806	0.7377
2	338.6	4363963	2.2445	0.2056	9.2000	0.5193
3	335.9	4697230	3.1179	0.2010	9.2000	0.5171
4	319.3	162345371	1.1553	26.5535	9.2000	0.3845
5	352.7	4754644	0.4536	6.9620	9.2000	0.6260
6	351.9	5104243	0.4680	7.2149	9.2000	0.6348
7	354.6	4046198	0.4041	6.5711	9.2000	0.6062
8	413.2	11225742	0.1000	6.6081	0.7081	0.8205
9	414.4	11149412	0.1000	6.3532	0.7057	0.8195
10	407.2	11580807	0.1000	7.3575	0.9199	0.8164

Table 7.5: The water balance for the Stadtlohn catchment for the run with the highest NS (see Table 7.4). When the f_{GS} and the f_{QS} values are compared with those of Table 7.3, the values of this water balance seem to indicate that the groundwater reservoir plays a larger role.

P	ET _{pot}	Q _{obs}	f _{XG}	f _{XS}	ET _{act}	Q	f _{GS}	f _{QS}	d _V	d _G	h _S	h _Q	n _{days}	check
692	595	269	0	0	5091	250	92.6	155	-66.1	-341	-0.8	0	365	0

7.2.2 Calibrating with low quality a priori information: Diesdonk, Lieveren, Tungelroyse beek

The calculations for the Diesdonk and Lieveren catchment were done before the conclusion was drawn that the negative flux was produced by the inclusion of d_{G0} in the calibration process. Thus for these two catchments the a priori information originating from the Stadtlohn catchment will be based on Table 7.1.

Diesdonk

Table 7.6 shows the results of the calibration of the Diesdonk catchment, with the data of the calibrated Stadtlohn catchment as a priori information. Table 7.7 shows the results of the calibration of the Diesdonk catchment, with the data of the Cabauw catchment as a priori information. As mentioned above, the uncertainty concerning the quality of the a priori information is much larger when compared to the uncertainty of the quality regarding the a priori information of the Berkel catchment. This is directly visible in the calibration results. The use of the Hupsel data as a priori information for the Berkel catchment is sufficient, but the use of the Berkel parameter values as a priori information for the Diesdonk catchment does not seem to be sufficient at all. The highest NS that is obtained is 0.37. However, the associated parameter-set shows that during the calibration process boundaries are hit (c_G and c_V). All the other parameter-sets also show parameter that equal lower- or upper-boundary values.

The Cabauw parameter-set as a priori information for the possible parameter values, leads interestingly enough to far better results for the Diesdonk catchment. Because the Diesdonk is

freely draining and the Cabauw catchment is not, this is somewhat unexpected. The highest NS that is obtained is 0.43. This NS is obtained for six of the 10 regulated Levenberg-Marquardt runs. The other four runs do not show clear similarities among each other. This indicates that there is a clear global minimum. For the runs that have the high NS no boundaries are hit. In this research, Diesdonk is the largest catchment that is calibrated on its own. When one looks at the parameter values that are obtained in Table 7.7, they seem to fit the idea of a large catchment. The high values for the c_V and c_Q both might indicate the reaction of the Diesdonk is slower due to its size. The c_S is also quite high, which is probably related to relatively large amount of discharge this catchment produces, when compared to the other catchments in this research. The c_G is quite high for a catchment that, besides sand also contains loam. However, this could indicate that there is not much loam present in the catchment.

Lieveren

When the first five Levenberg-Marquardt runs are performed the results of the runs are all the same. The runs lead to exactly the same parameter-sets with the same low objective value, NS equals -1.03. The c_V and the c_Q both hit the upper-boundary. The regularisation parameter α is changed a few times from values ranging from $\alpha=1e-8$ to $\alpha=30$. This did not lead to better results, the boundaries are still hit when the α is changed. As test, the weights of the two parameters in question are increased. As a consequence, the algorithm does not let the c_V and the c_Q hit the boundary any more. However, the other four parameters now do. Because of the bad results, and because of time restrictions, it is

Table 7.6: The obtained parameter sets of 10 regulated Levenberg-Marquardt runs ($\alpha=1e-6$) for the Diesdonk catchment. The start values are the best particle values of the hydroPSO sampler and the centre values are the means of the 10 obtained parameter values (Table 7.1) of the Stadtlohn catchment. The parameters that equal a boundary value are printed in bold font. It is clear that calibrating the Diesdonk catchment with the Stadtlohn data as the a priori information of the possible parameter values, does not lead to good results. In all the 10 runs boundaries are hit. The highest obtained NS value is 0.38. The associated parameter-set hit the upper-boundaries for the c_G and c_V parameters.

Run	cW	cG	cV	cQ	dG0	cS	NS
1	582.0	276000000	0.1000	7.31184	1945.0	0.5203	-0.10
2	500.3	14894023	0.1146	25.38055	120.9	9.2000	-1.44
3	514.0	276000000	1.2965	0.01000	1886.7	0.1139	0.09
4	530.8	26426486	0.1000	13.76616	2058.8	9.2000	0.04
5	585.1	34984630	0.1000	18.93816	2157.0	1.3279	0.28
6	587.8	42133302	72.2000	8.38059	1947.8	0.6042	0.25
7	531.1	24216504	0.1000	13.79992	2058.5	9.2000	0.04
8	518.3	276000000	0.4170	12.24832	2022.2	9.2000	-0.02
9	530.7	24132861	0.1000	13.73511	2058.3	9.2000	0.03
10	617.6	276000000	72.2000	0.04309	1791.9	0.1003	0.38

Table 7.7: The obtained parameter sets of 10 regulated Levenberg-Marquardt runs ($\alpha=1e-6$) for the Diesdonk catchment. The start values are the best particle values of the hydroPSO sampler and the centre values are the parameter values Brauer et al. (2014b) obtained for the Cabauw catchment. The parameters that equal a boundary value are printed in bold font. The Cabauw catchment data as part of the a priori data leads to better results than the Stadtlohn catchment data. The highest obtained NS equals 0.43. The associated parameter-sets do not hit any boundaries. The parameter-sets for which the NS does not equal 0.43, do not show clear similarities among each other, indicating that there is a clear global minimum.

Run	cW	cG	cV	cQ	dG0	cS	NS
1	633.0	117698425	13.96	75.23	1875	2.0723	0.43
2	632.1	119430378	14.01	75.22	1871	2.1218	0.43
3	635.7	31687808	13.96	74.64	1857	2.0676	0.43
4	632.6	119573255	13.62	75.59	1860	2.0860	0.43
5	630.9	119633902	14.01	75.16	1871	2.1514	0.43
6	400.0	74744363	14.03	135.20	1817	0.7126	-0.78
7	572.1	100877699	13.97	75.55	1819	9.2000	0.33
8	571.3	60108847	15.04	75.49	1818	9.2000	0.32
9	619.0	276000000	13.98	75.38	1842	1.7873	0.41
10	630.7	119711018	13.99	75.04	1859	2.1432	0.43

decided to not investigate the Lieveren catchment in combination with Stadtlohn a priori information, any further. Table 7.8 shows the results of the calibration with the Cabauw parameter-set as part of the a priori information. The important difference with the first calibration run with Lieveren is that the calibration algorithm is able to find different results. However, the NS's of all the runs are negative, meaning that the quality of the simulations are very low. For nine of the 10 runs the c_G hits the lower- or upper-boundary. Just like with the Stadtlohn parameter-set, it is clear that the Cabauw parameter-set can also not be used as part of the a priori information.

Tungelroyse beek

The new Stadtlohn calibration runs (Table 7.4) are used as a priori information for the Tungelroyse beek catchment. This also means that the d_{G0} will not be part of the parameters that will be calibrated. In Table 7.9 the results are shown. As one can see the quality of the calibrations is very low. All the objective values are below zero, meaning that they simulate discharge with WALRUS that is worse than

the average of the observed discharge. For eight of the 10 parameter-sets a boundary is hit. Although the d_{G0} is excluded from the calibration process and the better results of the Stadtlohn calibration are used as a priori information, this does not lead to good results for the Tungelroyse beek.

Due to time restrictions, the Tungelroyse beek is not calibrated with the Cabauw data as information for the possible parameter values.

7.2.3 Coupling of WALRUS models: Berkel sub-catchments

For the coupling of the sub-catchments the d_{G0} will not be included in the calibration. This is due to the results found in Section 7.2.1. Based on Table 11.2, a c_D of 2500 mm will be used. For the a priori information on the possible optimal parameter values of the Lutum sub-catchment, the second Stadtlohn catchment results (Table 7.4) are used.

Table 7.10 shows the obtained parameter-sets for each sub-catchment.

For the German sub-catchments (Lutum, Stadtlohn and Rekken) WALRUS seems to perform relatively well. However, for the Dutch sub-catchments

Table 7.8: The obtained parameter sets of 10 regulated Levenberg-Marquardt runs ($\alpha=1e-6$) for the Lieveren catchment, with as start values the best particle values of the hydroPSO sampler and the Cabauw parameters as centre-values. The parameters that equal a boundary value are printed in bold font. The use of the Cabauw parameter-set as part of the a priori data leads to more varied results as is the case with the Stadtlohn parameter-set. However, the results are still of very low quality, all the NS's are negative. The c_G parameter hits nine of the 10 times a boundary. This can indicate that with the used starting values, this parameter is quite sensitive.

Run	cW	cG	cV	cQ	dG0	cS	NS
1	288.1	3000000	14.67	75.99	1077	0.2742	-0.5963
2	324.3	3000000	14.00	75.67	1039	0.1973	-0.3196
3	379.0	276000000	13.98	76.15	1650	0.2077	-1.9351
4	549.2	3000000	14.44	76.23	1213	0.2243	-0.8469
5	412.5	3000000	14.20	75.98	1182	0.2441	-0.7449
6	33.9	276000000	72.20	135.20	1793	0.3851	-13.1640
7	722.0	276000000	72.20	75.93	1950	0.1000	-3.6064
8	358.5	38582396	72.20	75.90	1677	0.3806	-3.0493
9	284.5	3000000	14.58	76.08	1088	0.3338	-0.7564
10	353.9	276000000	72.20	75.90	1691	0.3909	-3.0879

Table 7.9: The results of the Tungelroyse beek calibration runs. The a priori information (the starting values and the centre-parameters) originate from the second calibration runs of the Stadtlohn catchment (Table 7.4). The d_{G0} was excluded from the calibration process. The parameters that equal a boundary value are printed in bold font. The quality of the parameter-sets is low. All the NS's are below zero, meaning that they simulate discharge with WALRUS that is worse than the average of the observed discharge. In eight of the 10 parameter-sets a boundary is hit.

Run	cW	cG	cV	cQ	cS	NS
1	166.5	68639062	3.199	14.679	9.200	-2.107
2	166.5	68486358	3.169	14.652	9.200	-2.112
3	135.6	276000000	6.752	29.068	5.550	-1.115
4	166.6	68477559	3.100	14.736	9.200	-2.114
5	141.5	276000000	6.787	34.246	9.200	-1.315
6	166.4	68592914	3.246	14.616	9.200	-2.108
7	249.1	276000000	7.915	1.115	0.100	-3.370
8	162.9	71284402	2.449	13.138	5.756	-1.904
9	166.4	68498370	3.105	14.509	9.200	-2.109
10	163.0	71311348	2.443	13.451	5.751	-1.906

Table 7.10: The obtained parameter-sets per Berkel sub-catchment and their respective objective value. The parameters that equal a boundary value are printed in bold font. For three of the five parameter-sets (Lutum, Stadtlohn and Lochem), there are parameter values that equal boundary values. The parameter-sets that are obtained for the German catchments (Lutum, Stadtlohn and Rekken) seem to lead to good objective values. The parameter-set obtained for the Dutch catchment Haarlo does not lead to a high NS. The simulation results for Lochem are of low quality.

Sub-catchment	cW	cG	cV	cQ	cS	NS
Lutum	438.3	13640383	4.3173	0.1000	0.2773	0.7326
Stadtlohn	415.1	13363008	0.1000	7.56	0.8844	0.8042
Rekken	403.8	3427321	0.5480	12.56	2.0242	0.8136
Haarlo	361.6	11008887	0.5426	11.80	1.8197	0.5234
Lochem	346.8	4616910	0.1000	0.1000	0.1319	-0.3598

the performance considerably drops. Haarlo has a NS of 0.52, which is not that good. The simulation for the Lochem sub-catchment even led to a negative NS. For the two Dutch sub-catchments, a priori information originating from the Hupsel catchment is used. The results from Section 7.2.1 show that it is possible to obtain good results with this information for the Stadtlohn catchment. This can indicate that the Hupsel catchment and the German sub-catchments share a lot of characteristics, and the Hupsel catchment and the other Dutch sub-catchments not. In Chapter 3 it is elaborated that there are differences between the German and Dutch side of the Berkel catchment.

As one can see, for three of the five parameter-sets of the sub-catchments (Lutum, Stadtlohn and Lochem), there is at least one parameter value that equals a boundary value. For the Lochem sub-catchment two boundaries are hit. The regularisation term apparently cannot make sure that all the parameters stay within the hard boundaries.

7.3 Conclusion

In Table 7.11 the best obtained parameter-sets are shown from the calibration runs performed for the two categories.

The parameter d_{G0}

One of the conclusions that can be drawn specifically for this chapter, is that the inclusion of the d_{G0} in the calibration process, can lead to unrealistic results. When the parameter is excluded from the calibration process, higher objective values can be obtained.

Stadtlohn

It can be concluded that the Hupsel data can be used as a priori information for the possible values of the parameters. With the used calibration approach objective values above 0.8 could be obtained. When one looks at the first and second set of calibration runs for the catchment, it can be concluded that high objective values (above the 0.7) can be obtained for the Stadtlohn catchment, with varying parameter-sets.

Diesdonk

With the used approach it is not possible to obtain good parameter-sets for the Diesdonk catchment. When the a priori information on possible optimal parameter values of the Stadtlohn catchment is used, the objective values are all below the 0.4. When the Cabauw data is used as a priori information, the results are slightly better. Most of the fitted parameter-sets lead to an objective value of 0.43. It is possible to conclude that this might indicate that the Cabauw and Diesdonk catchments share certain characteristics.

Lieveren

The Stadtlohn and Cabauw parameter-sets cannot be used for the Lieveren catchment as a priori

information for the starting values and centre-parameters. The objective values that are obtained are all below zero. For the set of calibration runs for which the Cabauw data is used as a priori information, the c_G often equals boundary values. This can indicate that with the used starting-parameters, this parameter value is very sensitive.

Tungelroyse beek

When the Stadtlohn parameter-sets are used as a priori information, low objective values are obtained for the Tungelroyse beek. All the objective values are below zero, meaning that the simulations are worse than the average of the observed discharge.

Coupling of the Berkel sub-catchments

The seventh research question can be answered in this chapter. The question is as follows:

is it possible to find optimal parameter-sets for coupled WALRUS simulations without introducing a hydraulic component?

To a certain extent it is possible to find optimal parameter-sets for coupled WALRUS simulations, without introducing a hydraulic component. The objective values that are obtained for the German sub-catchments are high. For the Dutch sub-catchments the objective value is considerably lower. It might thus be possible to conclude that at a small scale, WALRUS can be coupled without introducing a hydraulic component. At a larger scale the results significantly drop and the addition of a hydraulic component might improve the results.

Table 7.11: The best obtained parameter-sets for the two categories of catchments that are calibrated in this chapter. For the Stadtlohn catchment only the the parameter-set of the second calibration run is shown. This parameter-set has thus only five parameters, because the d_{G0} was excluded from the calibration run. For the Diesdonk catchment, the results of both calibration runs are shown. Respectively the calibration run for which a priori information from the Stadtlohn catchment was used and the calibration run for which the a priori information originating from the Cabauw catchment was used. For the Lieveren catchment only the results of the calibration run with a priori information originating from the Cabauw catchment are shown. This is because the other calibration process was stopped after five Levenberg-Marquardt calibration runs, due to the low quality of the results. Due to time restrictions the Tungelroyse beek was only calibrated with information originating from the Stadtlohn catchment. For the Tungelroyse beek, the results from the second Stadtlohn calibrations were used as a priori information. Thus the d_{G0} was excluded from the calibration run. The parameters that equal a boundary value are printed in bold font.

Catchment	cW	cG	cV	cQ	dG0	cS	NS
Stadtlohn	413.2	11225742	0.1000	6.6081	-	0.7081	0.8205
Diesdonk (Stadtlohn)	617.6	276000000	72.2000	0.04309	1791.9	0.1003	0.38
Diesdonk (Cabauw)	630.7	119711018	13.99	75.04	1859	2.1432	0.43
Lieveren (Cabauw)	288.1	3000000	14.67	75.99	1077	0.2742	-0.5963
Tungelroyse beek (Stadtlohn)	135.6	276000000	6.752	29.068	-	5.550	-1.115

8 | Uncertainty Analyses

In this Chapter a uncertainty analysis method will be validated. The uncertainty analysis will actually be performed through sensitivity analyses. See Chapter 5 for an explanation on how the sensitivity analyses are done in this research. Saltelli and Annoni (2010) state that with a sensitivity analysis used as uncertainty analysis, one can answer the question "How uncertain is this inference?". Or in other words, how large is the uncertainty regarding the conclusion that is reached on basis of the evidence.

In hydrological models — and in environmental models in general — uncertainty can be attributed to five different sources (Beck, 1991):

1. The initial conditions;
2. The used data (for forcing and for calibration);
3. The used calibration algorithms;
4. The values of the parameters;
5. The model structure.

The initial conditions (1) in this research form an important aspect of the uncertainty. Throughout this research the a priori information regarding possible parameter values, has proven to be important. The a priori information mainly plays an important role for the starting-values and boundary-values of the so-called 'rough runs' (Chapter 4) and for the centre-parameters of the regularisation (Chapter 6). The used data (2) comes from different sources and thus there is an uncertainty regarding the quality. This quality of the data might be part of the cause of some of the low objective values that are obtained in this research. The Levenberg-Marquardt algorithm (3) can also increase the uncertainty. Due to equifinality one cannot be sure if the optimal parameter set might have been obtained with the algorithm. The fixed parameter values (4) like the c_D and the a_S that are used for all the catchments are all based on the Hupsel catchment. Because parameters are thus only approximations for the other catchments, there is a degree of uncertainty caused by these fixed values. The model (5) is set-up to simulate specific behaviour of the environmental system. This means that certain aspects of this environmental system are left out of the model, or that they are simplified. There often is a degree of uncertainty regarding these simplifications.

Due to the multitude of sources of uncertainty, the analysis in this chapter can mainly say something about the amount of uncertainty, and not so much about where it is coming from.

The uncertainty analysis method is validated on the results of the second calibration runs of the Stadtlohn catchment. Due to time restrictions the other catchments are left out of this analysis.

8.1 Methodology

The exact same method as explained in Chapter 5 is used. This means that an 'One-At-a-Time' (OAT) sen-

sitivity analysis method is used. The sensitivity analysis scripts are run for the same period that is also used for calibrating the Stadtlohn catchment. A discharge peak that occurs in this period will be used for the analysis. The subsequent recession period is also included in this period. This gives the chance to investigate what happens with the parameter sensitivity in two different situations. The discharge peak, and its subsequent recession period occur between the 24th of March and the 14th of April, 2009.

8.2 Results

Figures 8.1, 8.2 and 8.3 show the results of the sensitivity analysis. The parameters will now be discussed one by one.

The c_W shows not a lot of sensitivity when compared to the other parameters (except when compared to c_G). The contribution of the parameter to the total sensitivity mainly seems to increase during the recession period, until the second smaller rainfall events start. This increase in contribution to the total sensitivity could be linked to the decrease of influence of the quickflow reservoir during the recession period. In such a period the storage deficit (d_V) will slowly increase due to the decrease of excess water (see Equation 2.1). The groundwater reservoir thus starts to play a more important role during such a period and the quickflow reservoir a less dominant role. The relative un-sensitivity of c_W is also visible in Chapter 7. You can see in most of the Tables that this parameter does not change a lot and almost never hits a boundary. This is behaviour that significantly deviates from the c_G , c_V , c_Q and c_S parameters. These parameters do have a tendency to hit boundaries and have (sometimes) wildly varying values. The fact that the c_W does not change a lot, might indicate that the uncertainty regarding the correctness of the value is low.

The c_G is not visible in Figure 8.2. When one looks at the related d_G in Figure 8.1 (third window from above), one can see that the groundwater depth (d_G) is relatively constant. This could indicate that the groundwater reservoir has a significantly different time-scale than the other parameters. The shorter time-scale of the parameters that react faster on a rainfall event (mainly c_V , c_Q and c_S) might lead to the difference in sensitivity. Another possible explanation for the low sensitivity of the groundwater reservoir parameter is that the scale value, which indicates the amount of variation the parameter value can have in a specific catchment, might be low when compared to the other values. This scale value than can give an indication on how large the uncertainty is regarding the c_G value. The smaller this value, the smaller the variation in c_G values that were found for the catchment in question. Because the c_G does not change a lot, it might be possible to conclude that the uncertainty regarding the correctness of the value is low.

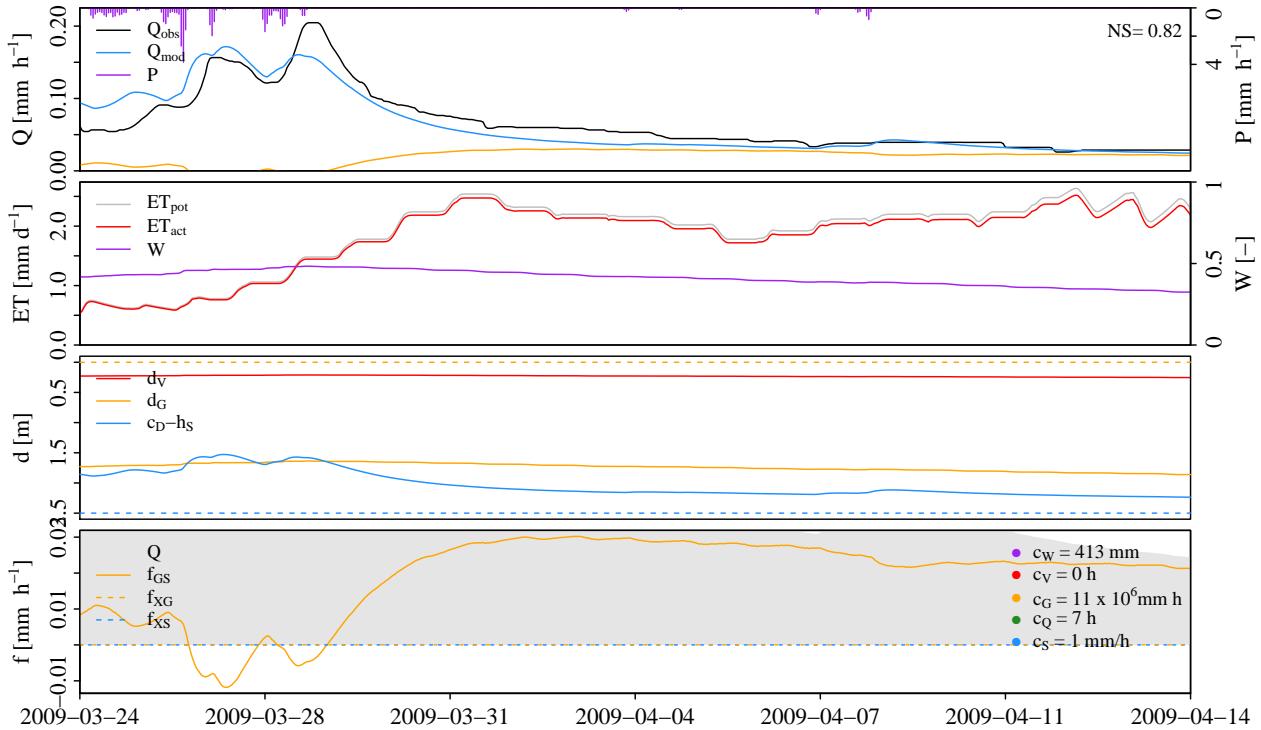


Figure 8.1: The rainfall event that occurred in the Stadtlohn catchment, which was used for uncertainty analysis. It shows a rainfall event and its subsequent recession period. The rainfall event occurred between the 24th of March 2009 and the 14th of April 2009. In the beginning of the period there is clear rainfall peak. In the recession period a few small rainfall peaks occur. These small rainfall peaks do not seem to influence the system that much.

The c_V shows a similar sort of sensitivity as the c_W parameter, only than a certain factor larger. Just like with the wetness index parameter, the vadose zone relaxation time parameter possibly becomes more sensitive during the recession period because the groundwater reservoir becomes more important.

The quickflow reservoir parameter is mainly sensitive directly after the rainfall event. During the recession period the influence of the c_Q decreases. This could indicate that the excess water from the rainfall event has been processed by the system, and that thus the influence of the quickflow reservoir drops.

The c_S seems to react heavily on the rainfall peaks. However, the sensitivity of this parameter is clearly not only related to rainfall peaks. In Figure 8.2 three peaks in sensitivity are visible for c_S , while in Figure 8.1 only rainfall peaks are visible for the first two peaks in c_S sensitivity. This indicates that the bankfull discharge parameter does not only correlate with rainfall peaks but also with something else. When one looks at Figure 8.3 one can also clearly see the peaks in sensitivity of the bankfull discharge parameter. For the heavy rainfall peak in the beginning of the period the c_S clearly shows the most sensitivity. For the smaller rainfall event at roughly two thirds of the period one can see that the c_S shows less sensitivity. The third peak in sensitivity of the bankfull discharge is clearly the smallest of the three. When one compares Fig-

ure 8.2 and Figure 8.3, it is again underlined (as in Chapter 5) that one cannot deduct the amount of sensitivity from the sensitivity as represented in Figure 8.2.

8.3 Conclusion

The ‘One-At-a-Time’ method can be used to perform uncertainty analysis. For the c_W and c_G parameters one could say something about the amount of uncertainty due to the fact that both parameters showed a relative constant un-sensitivity. For the other parameters it is however difficult to conclude something about the uncertainty. It mainly depends on what the user of these type of analyses wants to investigate. For instance, a user could be more interested in the rainfall peak. The conclusions he or she would draw then, are mainly based on the sensitivities the parameters show around that peak. However, these conclusion would probably be different when the user wants to research the end of the recession peak.

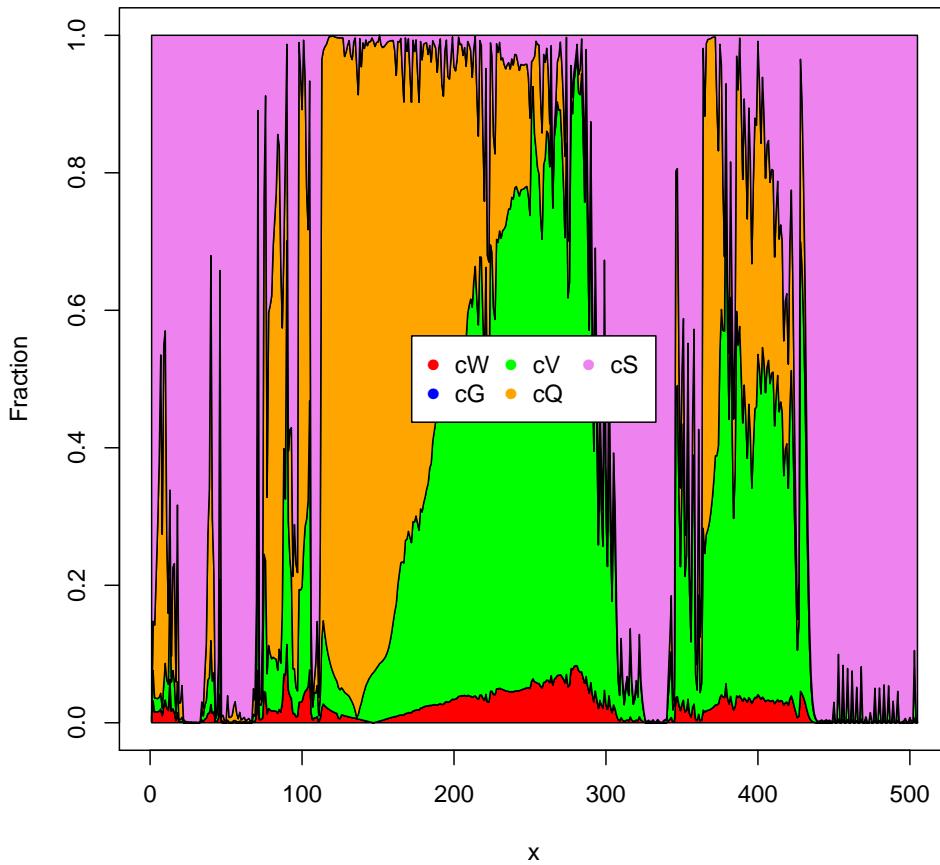


Figure 8.2: The total fractional sensitivity of the Stadtlohn parameters, used for uncertainty analysis. This Figure shows the fraction of the sensitivity of the parameters on the y-axis. The x on the x-axis stands for the time steps. The period represented in this picture is the same one which is shown in Figure 8.1. This Figure seems similar to Figure 5.3. However, this Figure shows the sensitivity of a different catchment and the sensitivity is used for a different goal, namely uncertainty analysis. The first two c_s sensitivity peaks can be related to the rainfall events. However, the third peak does not let itself explain when only looking at the rainfall peaks. When one looks at Figure 8.3 one can see that the amount of sensitivity of c_s decreases per peak. This underlines that the above shown figure, which expresses the sensitivity in fractions, only shows the amount of sensitivity a parameter contributes to the total amount of sensitivity at a specific moment.

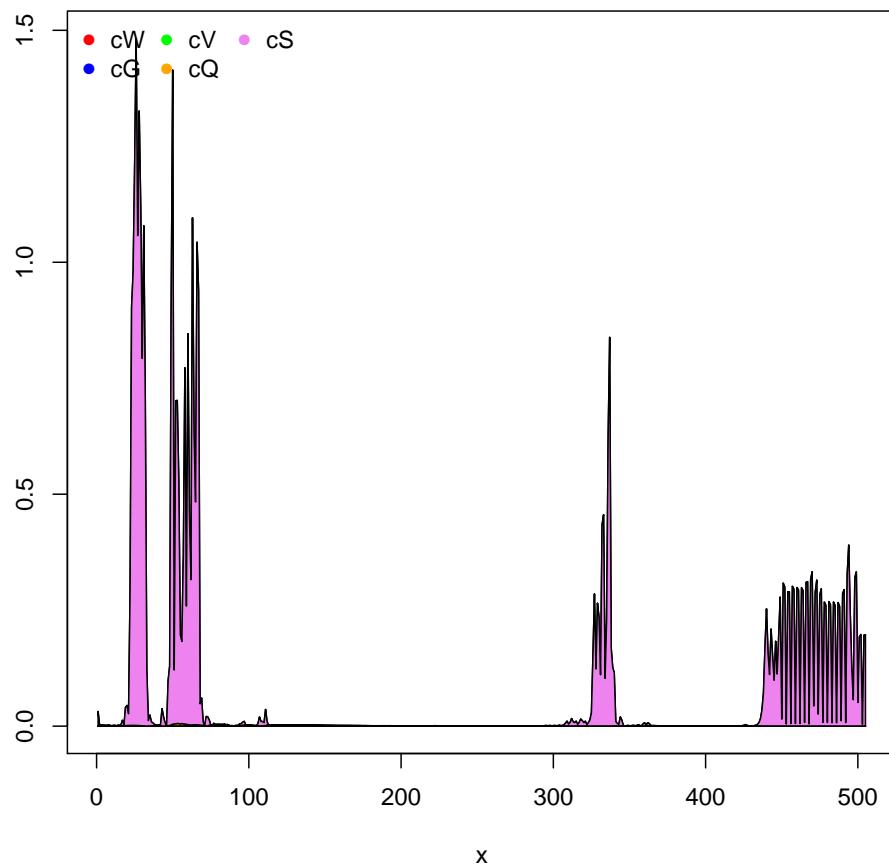


Figure 8.3: The total cumulative sensitivity of the Stadtlohn parameters, used for uncertainty analysis. This plot shows the total cumulative sensitivity of the parameters on the y-axis. The x on the x-axis stands for the time steps. The period represented in this picture is the same one as is shown in Figure 8.1. This Figure seems similar to Figure 5.2. However, this Figure shows the sensitivity of a different catchment and the sensitivity is used for a different goal, namely uncertainty analysis. You can see that the c_s contributes the most to the total sensitivity. When one looks at the decreasing height of the peaks of the sensitivity of the bankfull discharge parameter, it is clear that one cannot deduct from Figure 8.2 the amount of sensitivity. Only the amount of contribution of sensitivity of a parameter at a specific moment.

9 | Discussion

In this chapter the main points of the research will be discussed. Like mentioned in Chapter 1, WALRUS depends on inverse modelling. This is in itself not a method without limitations. There is a large dependency on the quality of the data, insight into the model behaviour, the knowledge on parameter behaviour during calibration and on the used calibration methods. When the model is not yet fully understood, it is difficult to say something about its behaviour. It can be the case that a supposed indicator of a certain process is not only correlated with that specific process, but also with something else. In this research, inverse modelling methods are identified and validated for finding optimal parameter sets for WALRUS. Because WALRUS is not yet fully understood, there are thus side remarks that can be made about this research. The points of attention will be the data, the sensitivity analysis, the calibration method and the calibration of the catchments.

9.1 Data

The data that is used throughout this research and discussed in Chapter 3, is obtained from different sources. There is quite some uncertainty regarding the quality of the data. The original source is not always known. Besides that, the data has been put in different formats by the previous users, and again in this research. Due to this, errors, or shifts in different columns might have occurred. This can have significant influence on the results. For example, the observed discharge might have been shifted, but the precipitation and the evapotranspiration not. This will lead to different calibration results. The model will be adjusted in such a way that it might simulate a catchment that reacts faster or slower than it in reality does.

9.2 Model

WALRUS has a few unique features which make the model stand out, when compared to other rainfall-runoff models. An example of such a feature is the coupling between the groundwater and the unsaturated zone (i.e. vadose zone). These two zones are divided by a dynamic groundwater table. Another example of such a feature are the wetness depended flowroutes. These are controlled by the dynamic vadose zone, which in turn influences the wetness depended divider through the storage deficit (d_V in Figure 2.1). Another important aspect of the model is the groundwater to surface water feedback. This feedback is important, because drainage plays a prominent role in lowland countries like The Netherlands. The model can also incorporate seepage and surface water supply (Brauer et al., 2014a).

In the research of Brauer et al. (2014b) WALRUS was calibrated for two catchments. For these two well-monitored catchments four parameters (c_W , c_G ,

c_V and c_Q) were calibrated. In this research, besides the four above mentioned parameters, one, but mostly two extra parameters are calibrated (d_{G0} and c_S). This was done because the catchments in question are less well-monitored than the ones used by Brauer et al. (2014b). However, this led to an increase in the amount of used parameters, thus increasing the risk of over-parametrization. Simply put, with over-parametrization the calibration algorithm gets more knobs which it can adjust to be able to obtain a good objective value, regardless of physical realism (see Chapter 4 for an example). The research of Jakeman and Hornberger (1993) suggests that because of the nature of the data that is used for rainfall-runoff models, four is the maximum amount of parameters that should be used. The amount of calibrated parameters in this research thus exceeds that number. More extensive data collection in the used catchments could lower the amount of parameters that should be calibrated. For instance, when a catchment specific Q,h-relationship is created, the parameter c_S does not have to be included in the calibration any more. Detailed research into the soil type of a catchment could lead to a catchment specific soil type. However, the danger in an increase of dependency on more specific data, is that it might go beyond the idea of designing relatively simple models like WALRUS.

9.3 Sensitivity Analyses

In this research One-At-a-Time (OAT) sensitivity analyses have been used for answering the questions "Where is the uncertainty coming from?" in Chapter 5 and "How uncertain is this inference?" in Chapter 8 (Saltelli and Annoni, 2010). According to Saltelli and Annoni (2010) these OAT sensitivity analyses can only be used for simple, linear models. WALRUS is however a non-linear model. Saltelli and Annoni (2010) state that OAT's are not able to detect interactions among factors. For instance, one cannot conclude from an OAT, that when the bankfull discharge parameter seems to be very sensitive to a small change, this might be strongly correlated with maybe the rainfall. This is clearly visible in Chapter 8. Two of the three sensitivity peaks of c_S might be related to rainfall, however, at the time of the third sensitivity peak there is no rainfall event. More extensive sensitivity analyses might be able to tackle problems like these.

The scale-values for the sensitivity analysis are based on the variations parameters show when they are calibrated throughout a certain period of time with help of a "moving-time-window" (see Chapter 5). The values that are obtained with the "moving-time-window" calibration runs, can be depended on the calibration algorithm that is used. It is possible that the hydroPSO algorithm obtains different values than the Levenberg-Marquardt algo-

rithm. Because of this, it might be possible that the scale-values used for the sensitivity analysis — but also those that are used for regularisation — are of questionable quality.

9.4 Calibration and validation

Throughout this research a systematic approach is identified and validated for the calibration process with WALRUS. This approach is then applied on the four catchments that are described in Chapter 3. This systematic approach however has a tendency to conceal the high dependency on a priori information regarding the possible parameter values. For instance the ranges and starting values of the 'rough-runs' are selected in such a way that they are thought to be broad enough to capture all the possible results. Due to the for now lacking knowledge on the possible parameter-sets, it is difficult to support the choices for these ranges. The starting values for the Levenberg-Marquardt and hydroPSO runs in turn are often based on the plots of the 'rough-runs'. These 'rough-runs' did give some insight into the behaviour. However, they can be compared with the earlier mentioned OAT's. All the values are fixed, except for one, which is then changed ten times. There might be correlation between the parameters. However, when the parameters are changed one by one, this will not be apparent in the plots.

A drawback of the used regularisation method is that it increases the dependency on a priori information. When one wants it to work well, one needs good information so that correct centre-parameters are selected.

To handle the problem with lacking a priori information on the possible parameter values, the different catchments in Chapter 7 are calibrated in order of approximation. First the one that is the closest to the Hupsel catchment is calibrated, and then the catchment closest to that catchment, and so on. For the Stadtlohn catchment this seems to work well. Good objective values are obtained. However, this does not mean that this is a general rule that can be applied to all the different cases. The idea behind the "calibrating in order of approximation" is that the catchment might share the most (geo-)hydrological characteristics with the previously calibrated catchment. Due to the heterogeneity of the environment — a much debated and written about topic in the hydrological sciences (e.g. McDonnell et al., 2007; Sivapalan, 2003) — there is a lot of ground to question this method. The environment can change quite drastically over a distance of only a few metres. Using the data from the closest catchment that is available, can lead to bad results because of the possibly completely different characteristics the two catchments can have.

In this research the catchments are only calibrated with help of the discharge. However, better results might have been obtained when multi-objective calibration was used. In this case, the model would then not only be calibrated with help of the discharge, but also with help of the ground-

water depth. The groundwater depth data can be obtained via DinoLoket (TNO Geological Service Netherlands). WALRUS uses the concept of groundwater volume. The measurement pipes which collect the data however, are focussed on monitoring the fluctuation in a specific area. It could prove to be difficult to deduct groundwater volume from the DinoLoket data. However, the obtained groundwater depth data can be used in a different way for the calibration of WALRUS. When the model is calibrated with help of the discharge, the daily average simulated groundwater depth can be compared with the strongest correlating measurement location. This means that the model can be further calibrated via signal analysis.

In the opinion paper of Kirchner (2006) it is stated that the traditional way of validating, as done in this research, might not be that revealing about the models behaviour and functioning. Validating the model on an area for which it is calibrated, the so-called split sample tests, does not say a lot about the model performance because it is already calibrated for that area. The two time series both represent similar conditions and are thus likely to show similar behaviour. Kirchner (2006) advocates the use of differential split sample tests. The model is calibrated for one area and the parameters are applied on an other area. This is not done with the validation runs in this research. It might revealing to apply this method to the validation runs in this research. It should be noted that the method Kirchner (2006) advocates, is however partially done in an other way. The parameter-sets of the Cabauw catchment and/or Stadtlohn catchment are used as centre-values of the other catchments. For the Diesdonk catchment this led to the suspicion that the catchment might share some characteristics with the Cabauw catchment.

10 | Conclusion

The objective of this Master thesis research is:

to identify and validate techniques for obtaining optimal parameter-sets for WALRUS in different catchments in The Netherlands.

A calibration approach is identified and validated in this research. This calibration approach is elaborated in the Chapters 4, 6 and 7. In short, the term 'calibration approach' encloses the 'rough-runs', the multi-start method and the Levenberg-Marquardt algorithm with scaled parameters, hard boundaries and added regularisation term.

In Chapter 1, research questions were proposed as guidance for the investigation. Questions one to six will now be answered. Question seven is already answered in Chapter 7.

1. *Can a different calibration approach than the one used in Brauer et al. (2014b), obtain good objective values?*

Good objective values can be obtained with the used calibration approach for the Hupsel and the Berkel catchments. This is not the case for the Diesdonk, Lieveren and Tungelroyse beek catchments. This is because the calibration approach depends on information regarding possible optimal parameter values of the catchment in question. For the Hupsel and Berkel catchments this information was available. For the Diesdonk, Lieveren and Tungelroyse beek catchment it was not.

2. *Is the used calibration approach efficient?*

When the approach is used without the multi-start method, it is a relatively quick and efficient method. The multi-start method however makes the used calibration approach considerably slower. This method is however only necessary when local minima are suspected to be present. When this is not the case the calibration approach can be applied without the multi-start method.

3. *To what extent do boundaries on the parameters have to be set for the calibration algorithm, to be able to identify optimal parameter-sets?*

Hard boundaries have to be implemented in the form of lower- and upper-boundary values per parameter. Regularisation has to be implemented to overcome some disadvantages of using hard boundaries. Without regularisation the chance is significantly higher that the Levenberg-Marquardt algorithm will obtain parameter-sets with fitted parameter values that equal boundary values. This can give unrealistic simulations (see Chapters 4 and 7) and can lead to numerical problems when the obtained parameter-sets are used for subsequent calibration runs that have the same hard boundary values (see Chapter 5).

3.1 *Is there a way to identify boundary values for the parameters?*

The 'rough-runs', as elaborated in Chapter 4, give to some extent the possibility to identify boundary values. However, to apply this method successfully, there has to be an a priori idea on what the optimal parameter values might be.

3.2 *Is it possible to avoid that fitted parameter values equal boundary values?*

Regularisation can partly avoid that fitted parameter values equal boundary values. Most of the calibration results that led to good objective values have non/almost no parameter values that equal boundary values. When the objective values are low, the respective parameter-sets often have more parameter values that equal boundary values.

4. *Can WALRUS be calibrated with a limited amount or lacking knowledge on possible optimal parameter values?*

When there is almost no a priori information on what the parameter values might be, it is difficult to obtain optimal parameter-sets with the used calibration approach. For the used 'rough-run' method and for the regularisation, information has to be available on what the possible optimal parameter values might be. This research shows that when there is less information available on these values, the objective value can become quite low.

5. *Does equifinality occur with WALRUS?*

There are indications that equifinality occurs when calibrating WALRUS. High NS values were obtained with varying parameter-sets for the Hupsel and the Berkel catchment.

6. *Do the same parameters as mentioned in Brauer et al. (2014b) prove to be the most sensitive?*

c_G , c_V and c_Q but also c_S , show behaviour that can be related to sensitivity. The sensitivity analysis shows that at certain moments c_V , c_Q and c_S significantly contribute to the total sensitivity. The sets of calibration runs, show that the four above mentioned parameters can have a tendency to hit the lower- and upper-boundaries in one catchment. This behaviour can be caused by their sensitivity.

11 | Recommendations

There are a lot of ways to calibrate WALRUS. This research only covers a small part of all the options that are available. Based on the findings of this research the following recommendations are made.

To improve validation, tests with split-sample validation runs as proposed by Kirchner (2006) are recommended. These types of validation runs can be revealing with respect to the behaviour of WALRUS.

Although the multi-start method works in the sense that it can detect possible local minima, it is relatively slow. The calibration approach without the multi-start method is on the other hand quite efficient. Research into a more efficient multi-start method can improve the applicability of the calibration approach that is identified and validated in this research.

As mentioned in the conclusion, equifinality does seem to occur when calibrating WALRUS. It might be interesting to investigate the effect of multi-objective calibration, as elaborated in the discussion. A method like this can narrow down the possible optimal parameter-sets.

It is recommended to research what makes the Diesdonk catchment perform well with the Cabauw parameter values as a priori information. Because of the differences between these two lowland catchments, it can lead to interesting insights into why certain parameter values are obtained.

It is advised to use the Cabauw data as a priori information on the possible parameter values for the calibration of the Tungelroyse beek. Just like with the Diesdonk catchment, the calibration process might perform better with the Cabauw data as a priori information. When this is the case, this might lead to insights into why certain parameter values are obtained.

It is recommended to test different methods for determining the scale values and weights that are used for, respectively, the sensitivity/uncertainty analysis and for the regularisation. These scale values and weights have quite some influence on the process. A more thorough insight into the influence of the used method might improve the insight into the behaviour WALRUS shows and it might improve the calibration approach.

Because the OAT method used for sensitivity analysis in this research is not that good at finding correlations, it is recommended to investigate more extensive sensitivity analysis methods. This way sensitivities can be better explained.

Bibliography

- Bauer, F., Lukas, M. A., 2011. Comparing parameter choice methods for regularization of ill-posed problems. *Mathematics and Computers in Simulation* 81 (9), 1795–1841.
- Beck, M., 1991. Principles of modelling. *Water Science & Technology* 24 (6), 1–8.
- Beven, K., Young, P., 2003. Comment on ‘bayesian recursive parameter estimation for hydrologic models’ by m. thiemann, m. trossen, h. gupta, and s. sorooshian. *Water Resources Research* 39 (5).
- Brauer, C., Teuling, A., Torfs, P., Uijlenhoet, R., 2013. Investigating storage-discharge relations in a lowland catchment using hydrograph fitting, recession analysis, and soil moisture data. *Water Resources Research* 49 (7), 4257–4264.
- Brauer, C., Teuling, A., Torfs, P., Uijlenhoet, R., 2014a. The wageningen lowland runoff simulator (walrus): a lumped rainfall-runoff model for catchments with shallow groundwater. *Geoscientific Model Development Discussions* 7 (1), 1357–1411.
- Brauer, C., Torfs, P., Teuling, A., Uijlenhoet, R., 2014b. The wageningen lowland runoff simulator (walrus): application to the hupsel brook catchment and cabauw polder. *Hydrology and Earth System Sciences Discussions* 11 (2), 2091–2148.
- Byrd, R. H., Lu, P., Nocedal, J., Zhu, C., 1995. A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific Computing* 16 (5), 1190–1208.
- Duan, Q., Sorooshian, S., Gupta, V., 1992. Effective and efficient global optimization for conceptual rainfall-runoff models. *Water resources research* 28 (4), 1015–1031.
- Eberhart, R., Kennedy, J., 1995. A new optimizer using particle swarm theory. In: Micro Machine and Human Science, 1995. MHS'95., Proceedings of the Sixth International Symposium on. IEEE, pp. 39–43.
- Fletcher, R., Powell, M. J., 1963. A rapidly convergent descent method for minimization. *The Computer Journal* 6 (2), 163–168.
- Gupta, H., Perrin, C., Kumar, R., Blöschl, G., Clark, M., Montanari, A., Andréassian, V., 2014. Large-sample hydrology: a need to balance depth with breadth. *Hydrology & Earth System Sciences* 10 (7).
- Jakeman, A., Hornberger, G., 1993. How much complexity is warranted in a rainfall-runoff model? *Water Resources Research* 29 (8), 2637–2649.
- Jansen, F., 2012. Relation between catchment characteristics and peak discharges. Bachelor's thesis, Wageningen UR.
- Kavetski, D., Kuczera, G., Franks, S. W., 2006. Calibration of conceptual hydrological models revisited: 1. overcoming numerical artefacts. *Journal of Hydrology* 320 (1), 173–186.
- Kirchner, J. W., 2006. Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology. *Water Resources Research* 42 (3).
- Marquardt, D. W., 1963. An algorithm for least-squares estimation of nonlinear parameters. *Journal of the Society for Industrial & Applied Mathematics* 11 (2), 431–441.
- McDonnell, J., Sivapalan, M., Vaché, K., Dunn, S., Grant, G., Haggerty, R., Hinz, C., Hooper, R., Kirchner, J., Roderick, M., et al., 2007. Moving beyond heterogeneity and process complexity: A new vision for watershed hydrology. *Water Resources Research* 43 (7).
- Melsen, L. A., December 2012. Comparison of recent lowland floods in the netherlands. Master's thesis, Wageningen UR.
- Nash, J., Sutcliffe, J., 1970. River flow forecasting through conceptual models part i—a discussion of principles. *Journal of hydrology* 10 (3), 282–290.
- Piotrowski, A. P., Napiorkowski, J. J., 2011. Optimizing neural networks for river flow forecasting—evolutionary computation methods versus the levenberg–marquardt approach. *Journal of hydrology* 407 (1), 12–27.
- Saltelli, A., Annoni, P., 2010. How to avoid a perfunctory sensitivity analysis. *Environmental Modelling & Software* 25 (12), 1508–1517.
- Shin, M.-J., Guillaume, J. H., Croke, B. F., Jakeman, A. J., 2013. Addressing ten questions about conceptual rainfall-runoff models with global sensitivity analyses in r. *Journal of Hydrology* 503, 135–152.
- Sivapalan, M., 2003. Prediction in ungauged basins: a grand challenge for theoretical hydrology. *Hydrological Processes* 17 (15), 3163–3170.
- Sorooshian, S., Gupta, V. K., 1985. The analysis of structural identifiability: Theory and application to conceptual rainfall-runoff models. *Water Resources Research* 21 (4), 487–495.
- Spendley, W., Hext, G. R., Hinsworth, F., 1962. Sequential application of simplex designs in optimisation and evolutionary operation. *Technometrics* 4 (4), 441–461.
- Wagener, T., McIntyre, N., Lees, M. J., Wheater, H. S., Gupta, H. V., 2003. Towards reduced uncertainty in conceptual rainfall-runoff modelling: dynamic identifiability analysis. *Hydrological Processes* 17 (2), 455–476.
URL <http://dx.doi.org/10.1002/hyp.1135>

Wagener, T., Wheater, H. S., 2004. Rainfall-runoff modelling in gauged and ungauged catchments. Imperial College Press.

Wiese, B., Nützmann, G., 2011. Calibration of spatial aquitard distribution using hydraulic head changes and regularisation. *Journal of Hydrology* 408 (1), 54–66.

Wright, S., Nocedal, J., 1999. Numerical optimization. Vol. 2. Springer New York.

Zambrano-Bigiarini, M., Rojas, R., 2013. A model-independent particle swarm optimisation software for model calibration. *Environmental Modelling & Software*.

Appendices

11.1 Chapter 4

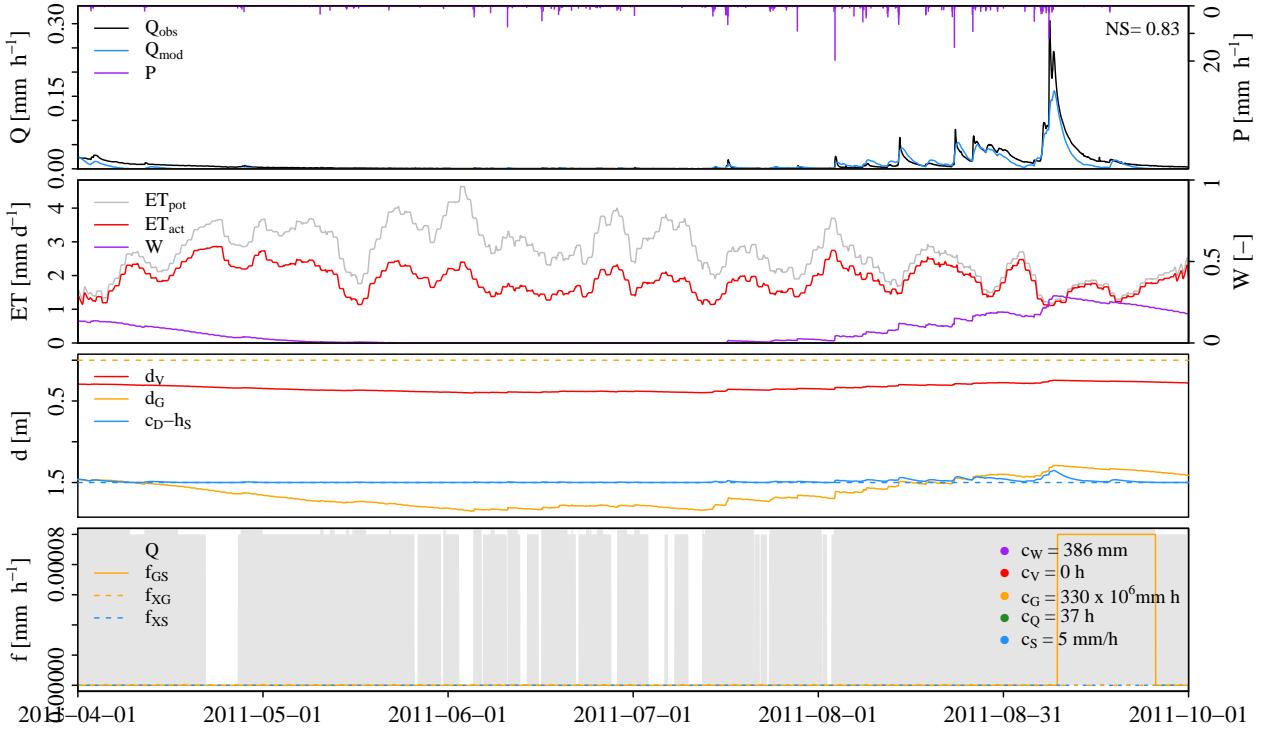


Figure 11.1: The first validation run that is performed with the fitted parameters obtained with the initial calibration. The validation period is from the beginning of April 2011 till the beginning of October 2011. It is clear that WALRUS performs even better with the parameters for this period. The NS equals 0.83. However, the c_G and c_Q are physically unrealistic.

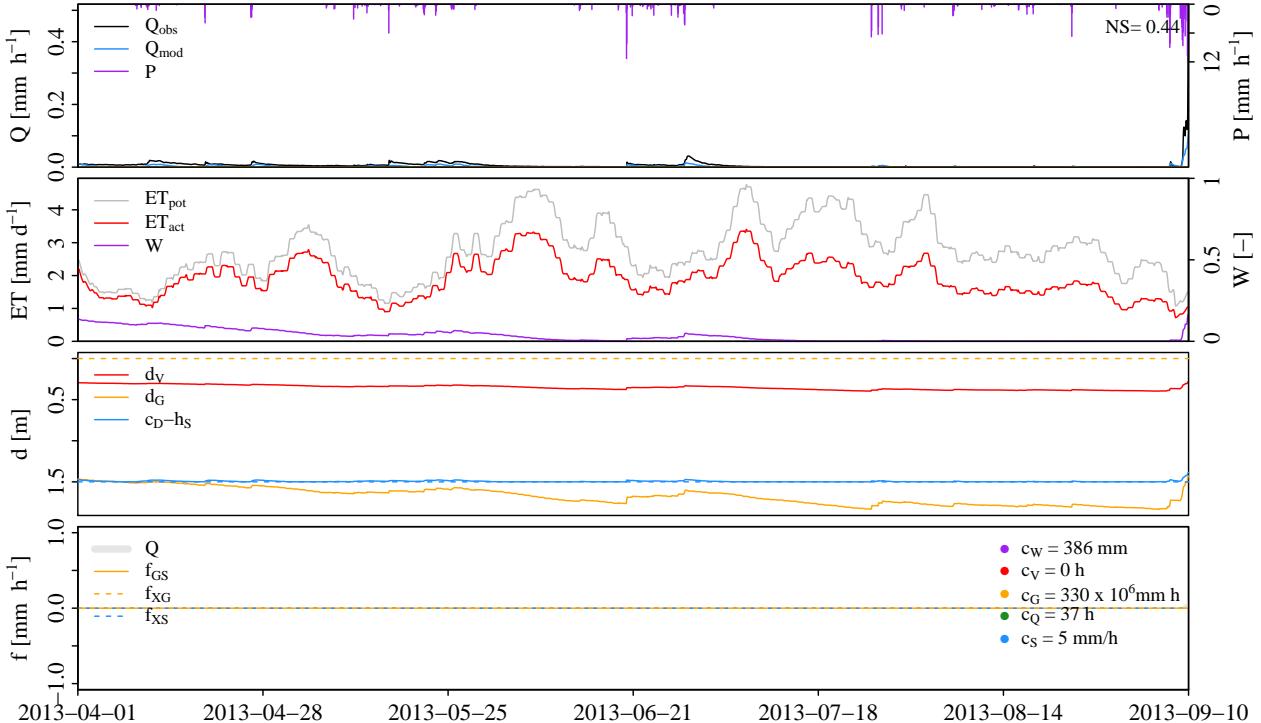


Figure 11.2: The second validation run that is performed with the fitted parameters obtained with the initial calibration. The validation period is from the beginning of April 2013 till half of September 2013. WALRUS considerably less well for this period when compared to the initial calibration itself and to the first validation run. The NS for this second validation run equals 0.44. This shows that not all the types periods can be modelled well with the obtained (physically unrealistic) fitted parameter set.

11.2 Chapter 5

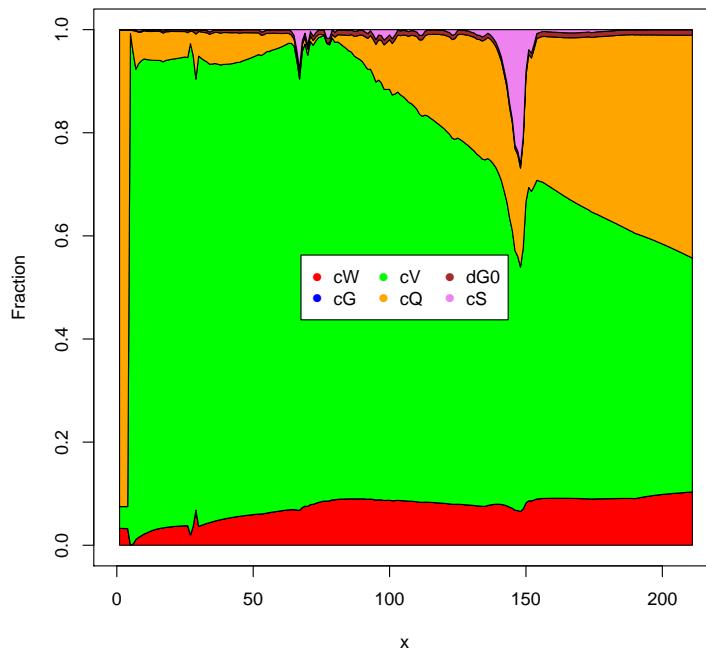


Figure 11.3: The fraction of sensitivity per parameter per time step obtained for the Hupsel catchment, in the period of approximately the 27th of July till the 15th of August in 2012. On the y-axis the fraction of sensitivity is given. On the x-axis the time steps are shown. The scale values are determined with help of Formula 5.3, which uses the difference between the maximum and minimum values found in the 19 calibration runs. This plot is smoother than when the values are calculated with Formula 5.4, as in Figure 11.4.

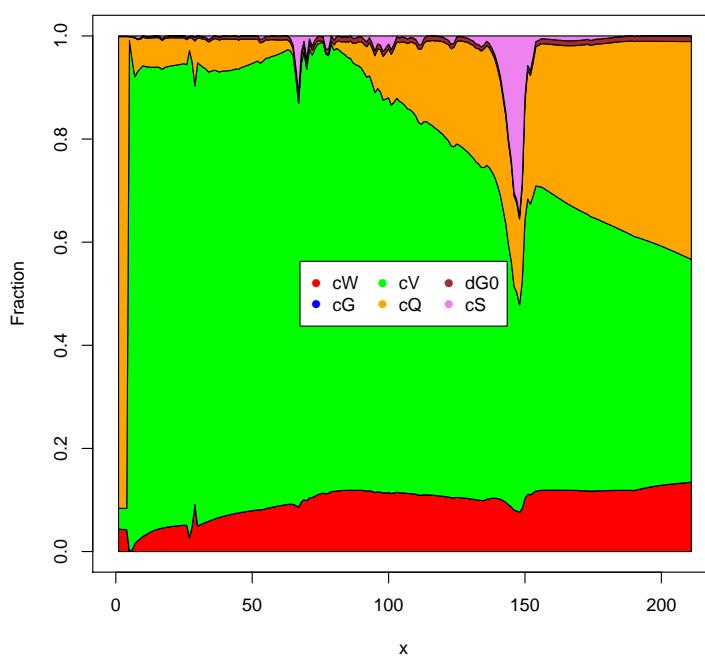


Figure 11.4: The fraction of sensitivity per parameter per time step obtained for the Hupsel catchment, in the period of approximately the 27th of July till the 15th of August in 2012. On the y-axis the fraction of sensitivity is given. On the x-axis the time steps are shown. The scale values are determined with help of Formula 5.4, which uses the standard deviations of the values found in the 19 calibration runs. This plot is more pronounced than when the values are calculated with Formula 5.3, as in Figure 11.3.

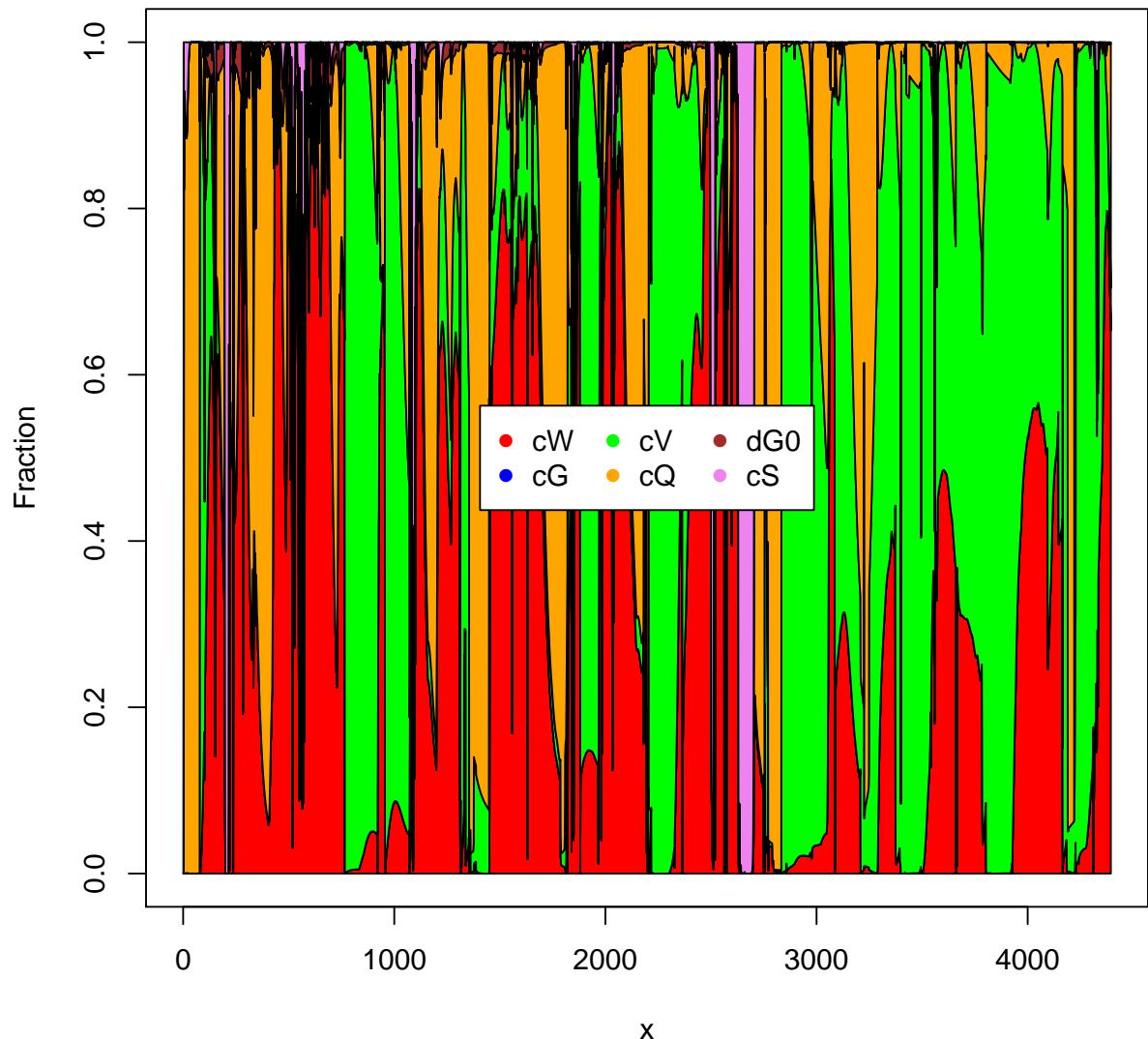


Figure 11.5: The fraction of sensitivity of the fitted Hupsel parameters per time step for the whole time period. On the y-axis the fraction of sensitivity is given. On the x-axis the time steps are shown. The scale values are determined with help of Formula 5.4, which uses the standard deviation of the values found in the 19 calibration runs.

11.3 Chapter 6

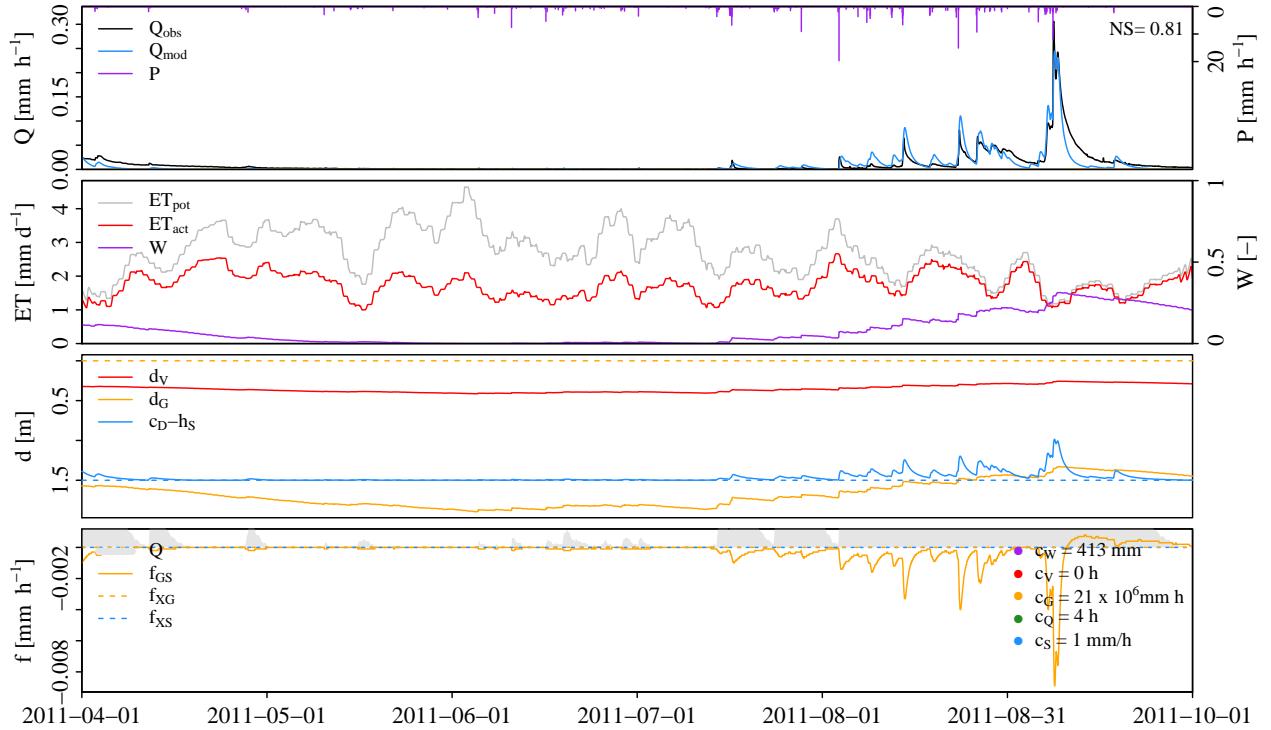


Figure 11.6: The first validation run that is performed with the fitted parameters obtained with the regulated calibration run for the Hupsel catchment. The validation period is from the beginning of April 2011 till the beginning of October 2011. WALRUS performs better with the parameters for this period. It performs only a bit less well when compared to the first validation run done for the initial calibration. The NS in this case equals 0.81. However, the parameter set obtained with the regulated calibration run is physically far more realistic than the one obtained with the initial calibration. The c_G and c_Q are considerably lower, which more realistic because the Hupsel is a fast-reacting catchment.

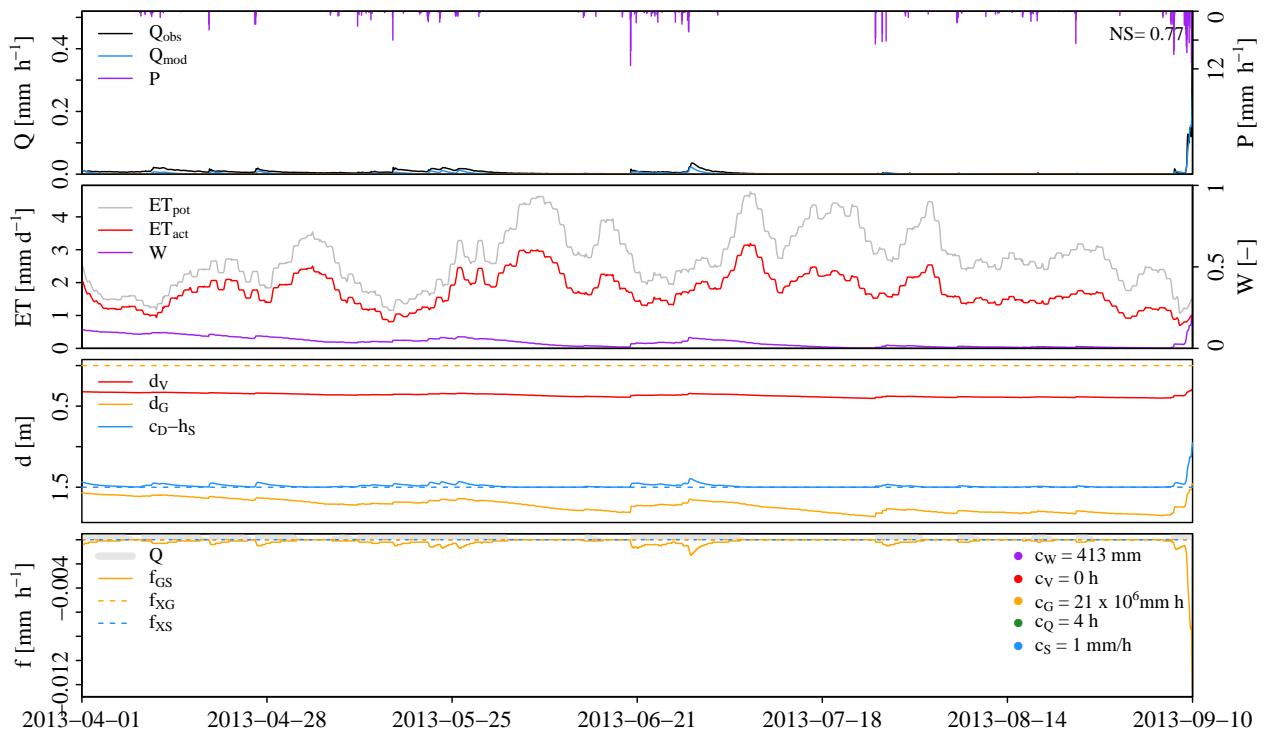


Figure 11.7: The second validation run that is performed with the fitted parameters obtained with the regulated calibration run for the Hupsel catchment. The validation period is from the beginning of April 2013 till half of September 2013. WALRUS performs well for this second validation run. The NS equals 0.77. When this is compared with the second validation run of the initial calibration, it is clear that the parameter-set obtained with the regulated calibration run is able to produce better results in more diverse situations.

11.4 Chapter 7

Table 11.1: The NS's of the validation runs performed with the ten fitted parameter-sets that are obtained with the calibration of the Stadtlohn catchment (see Table 7.1). The obtained parameter-sets all perform well for the validation period.

Run	1	2	3	4	5	6	7	8	9	10
NS	0.7033	0.7192	0.7153	0.7184	0.7181	0.7141	0.7135	0.7188	0.7183	0.7022

Table 11.2: The NS's of the WALRUS runs performed with different values for c_D for the Stadtlohn catchment. The chosen values are all values that produce positive f_{GS} 's in the water balance. You can clearly see that a c_D between 2500 mm and 3000 mm produces the best result. When the c_D becomes lower or higher than these values, the NS drops.

c_D	2000	2500	3000	3500
NS	0.69	0.74	0.74	0.61

Table 11.3: The NS's of the validation runs performed with the ten fitted parameter-sets that are obtained with the second set of calibration runs of the Stadtlohn catchment with the d_{G0} excluded from the calibration process (see Table 7.4). Contrary to the validation runs performed for the first calibration sets, the obtained parameter-sets do not all perform that well for the validation period. For the parameter-sets that got a NS value above 0.7 for the calibration period, the validation runs performed well. For the parameter-sets that got a NS value below 0.7 for the calibration period, the validation runs were considerably worse. This difference might be produced by the stochastic nature of the hydroPSO sampler.

Run	1	2	3	4	5	6	7	8	9	10
NS	0.7524	0.2712	0.2735	-0.6969	0.4098	0.4226	0.3806	0.7579	0.7556	0.7363