Influence of uncertainties in discharge determination on the parameter estimation and performance of a HBV model in Meuse sub basins

Enschede, April 2010

Sander P.M. van den Tillaart
Water Engineering & Management
University of Twente
The Netherlands

UNIVERSITY OF TWENTE.

Influence of uncertainties in discharge determination on the parameter estimation and performance of a HBV model in Meuse sub basins

Enschede, April 2010

Master Thesis

Sander P.M. van den Tillaart
s.p.m.vandentillaart@alumnus.utwente.nl

University of Twente
Department Water Engineering & Management
The Netherlands

UNIVERSITY OF TWENTE.

Supervisors:

Dr. M.S. Krol
Water Engineering & Management
University of Twente

Dr. Ir. M.J. Booij
Water Engineering & Management
University of Twente

Summary

Water institutes from all over the world have an important task of predicting future short-term and long-term discharges and water levels in river basins. These predictions are of importance for example to estimate the influence of climate change on future discharges and water levels. With adequate predictions possible threats of floods and droughts in the future can be estimated.

Before a model is applicable to a certain river basin, the model has to be calibrated and validated. In the calibration process a set of parameters is searched which approximates the measured discharge best, given sets of measured input data series. The HBV model (Bergström, 1976) is an example of a model that is used for hydrologic modeling. A lumped version of this rainfall runoff model is used in this research. It uses precipitation, temperature and potential evapotranspiration as input and the simulated discharge as output. The model contains equations and eight parameters which together describe a hydrological system.

Measurement errors of input and output series may result in errors in estimated parameters and hence errors in simulated discharge. In particular, the effect of sampling errors in precipitation on the estimated parameters and simulated discharge has frequently been studied. In hydrological modeling often the assumption is made that the effect of errors in discharge is negligible. In this research the effect of discharge errors on model performance and model parameters is investigated, by applying the HBV model to two sub basins of the Meuse River, namely the Ourthe and Chiers basins.

First of all a calibration is performed using the original data. The calibration procedure is a global parameter optimization method named SCEM-UA (Vrugt et al., 2003a) in which a combined objective function is used which emphasizes both the water balance and the shape of the hydrograph. The calibration period is 1984 – 1998 and the five most sensitive parameters of the model are calibrated. The calibration resulted in a higher value for the objective function in the Ourthe compared to the Chiers basin. This was also the case in the validation, which was performed over a period of 16 years (1986 – 1983).

Four different sources of errors in discharge determination are considered. Two error sources concern errors in discharge measurement. This can be (1) a combination of systematic and random errors without autocorrelation or (2) measurement errors which are random and auto correlated. The other two error sources are a consequence of the use of the discharge-water level (Q-h) relation. Firstly, (3) the Q-h relation does not take some processes in the hydrograph into account, such as hysteresis or the properties of a high water event, or (4) the effects of an outdating of the Q-h relation. The original discharge data are adapted in a way that the series are disturbed with each of the above errors. For every error source several different discharge data series are constructed with different errors. The quality of each data series is characterized by using two quality functions, named *QOD* and *BALANCE*.

The HBV model is calibrated for each of the discharge data series and corresponding quality functions and model performance are determined. It turned out that the random errors without autocorrelation do not have any significant influence on model performance and that the systematic errors have a considerable influence, even if the error is relatively small. One remarkable fact is that in both basins the model performance increases with respect to the original situation if a small

positive systematic error is present. Random errors with autocorrelation have some influence, depending on the autocorrelation coefficient. The error source which emphasizes the properties of a high water event does not have any significant influence on model performance. The effect of an outdating of the Q-h relation has similar effects compared to the systematic errors. This is because this error source contains a kind of systematic error. The effects of the errors do not vary much between the two basins.

If a significant influence on model performance is present, the parameters are influenced as well. If the influence of the error sources on model performance is small, the influence on model parameters is also small. Within the five used parameters two types can be distinguished: three parameters are mainly influenced by changes in the water balance and therefore by systematic errors, and two parameters are more related to the shape of the hydrograph and therefore influenced by random errors. The water balance related parameters show logical patterns regarding their physical representation if errors are present, while for the other two parameters no logical patterns can be distinguished.

The effects of the error sources on model performance together with the expectation of what is real are the basis of the choice for a realistic scenario of errors. The assumption is made that in both basins discharge determination is done by using the Q-h relation. In the realistic scenario it is assumed that this Q-h relation loses its validity after some time subsequent to a revision and that measurement errors occur in the water level determination. The realistic scenario consists of a set of possible discharge series and calibrations, because of the randomness character of the scenario.

The influence of the different discharge series on model performance and parameters in the realistic scenario is mainly caused by the systematic error due to the expiration of the Q-h relation. In general, unfavorable values for the discharge quality functions lead to a worse model performance. The highest value for the objective function is found if *BALANCE* has a small positive value, so if a small systematic error is present in the discharge data.

The HBV model has a better model performance in the Ourthe basin than in the Chiers basin. This might be caused by the presumption that the quality of the data in the Ourthe basin is better than in the Chiers basin. Another possibility is that the HBV model can perform better in basins which have a discharge regime with low base flow and high peaks like the Ourthe basin, compared to basins with a higher base flow and less high peaks, like the Chiers basin.

Error sources which contain a systematic error, such as the combination of systematic and random errors without autocorrelation or an outdated Q-h relation and the developed realistic scenario have effects on the water balance related parameters. Therefore the uncertainty due to the used discharge data is quite large, because these parameters are quite sensitive to systematic errors. These parameters have a small uncertainty due to the calibration method. For these parameters no big differences between the Ourthe and Chiers basins are found.

The uncertainty of the other two parameters due to the measurement errors is large in both basins, because values within the entire parameter ranges are found and no patterns are visible. The uncertainty in parameter value due to the calibration method for these parameters is large in the

Chiers basin. In the Ourthe basin the uncertainty is small if the value of the objective function is high, but the uncertainty increases if the value of the objective function decreases.

In general it can be concluded that the quality functions *QOD* and *BALANCE* give a good picture of the effects of the different errors on model performance and parameter estimation. Some patterns recur, particularly if model performance is expressed against *BALANCE*. Also regarding well-identified parameters, *BALANCE* has a logical influence on the parameter values.

Errors in discharge series that have a systematic character have much influence on model performance in both basins, while random errors and errors that are a result of processes in the hydrograph do not show much influence. The water balance related model parameters are mainly influenced by systematic errors, while the other parameters do not show any logical patterns. A recommendation is done to perform more research about the presence and magnitude of systematic errors, for example if a Q-h relation is used, so more knowledge about the influence of systematic errors in discharge on model calibration can be acquired.

Preface

About one year ago I started this research, not knowing what this year would bring. Now, one year later, I can say that I have learned a lot. During the WEM master courses I was introduced into the field of hydrology, which immediately fascinated me. When I started looking for a graduation research, I quickly came in contact with my present daily supervisor, Martijn Booij. Together with him I developed a proposal for a research, which seemed interesting to me. This was the beginning of a very instructive period that I have enjoyed very much.

All over the world, water institutes have an important task of predicting future short-term and long-term discharges and water levels in rivers. The field of hydrology is essential in this context, because it is of importance to estimate the influence of for example changes in climate on future discharges and water levels. With adequate predictions possible threats of floods and droughts in the future can be estimated. I think my research is a step towards a useful addition to the previous studies about hydrological modeling. Hopefully this research gives some more insight about the importance of adequate and accurate methods for discharge determination, because a hydrological model turned out to be quite sensitive for uncertainties in discharge measurements.

I would like to express thanks to some people that helped me during this research. First of all, I would like to say thanks to my supervisors. Martijn, you always advised me if I had problems with the HBV model or the Matlab and Fortran programs and gave helpful tips for relevant literature. Maarten, you were the person that proposed critical questions during our meetings. You often approached the problem from a different angle than the hydrological view and tried to provoke me with the important questions that kept me having the big picture in mind. I experienced the meetings with you both as very helpful and interesting, even during the period that I had some difficulties in motivating myself.

Apart from my supervisors, there are some people I would like to mention. First of all, I would like to thank Jasper Vrugt from the University of Amsterdam (UvA) and UT-alumnus and former classmate Han Vermue for providing me the SCEM-UA algorithm and for supplying me the relevant literature that helped me setting up the model calibrations.

Furthermore I would like to thank my present and former roommates of the graduation room and employees of the WEM-department. I really enjoyed the atmosphere at the UT with you, as well as during lunch times, 'borrels', 'daghaps' and barbecue evenings. I also would like to give a word of thanks to my friends from SHOT for their numerous cups of coffee and tea and social amusement during the long working days and Thursday nights. All you guys made that my graduation period became a successful completion of my student life!

My final word of thanks I would like to give to my parents, Jeanne and Piet van den Tillaart, and my sweet girlfriend Jessica. Without your support and motivation, but also your love and patience it would have been much more difficult to complete this research. Thank you all very much!

Sander van den Tillaart
Enschede, April 2010

Table of Contents

List of Figures

List of Tables

1 Introduction

In this chapter, an introduction of the research is presented. First, the background of the problem is explained in paragraph 1.1. Some similar previous studies are treated in paragraph 1.2. Together these elements lead to the problem description in paragraph 1.3. Subsequently the objective, research questions and used methodology are presented in that paragraph.

1.1 Background

Water institutes from all over the world have an important task of predicting future short-term and long-term discharges and water levels in river basins. An issue like climate change indicates the importance for adequate discharge and water level predictions. With adequate predictions possible threats can be estimated and the future risk of floods or droughts can be evaluated. For predicting the future discharges and water levels, hydrological models can be used. Hydrological models can be (semi-)distributed or lumped and can either be conceptual or physical.

A semi-distributed model is used if a basin can be separated into a number of sub basins and that each of these basins is distributed according to elevation and vegetation. A lumped model does not take into account the spatial variability of processes, input, boundary conditions and watershed geometric characteristics (Singh, 1995). If a model is conceptual, it means that the model parameters do not directly represent physical properties. That is why model parameters cannot be measured in the field. The model parameters which represent some basin characteristics are determined by calibration of the model. The advantage of a conceptual model is that it has a simple model structure. A disadvantage is that most parameters are empirical, which may reduce the validity of the model.

Before the model is applicable to predicting of future discharges in a certain river basin, the model has to be calibrated and validated. In the calibration process a set of parameters is estimated which results in the best simulation of the observed discharge, given sets of measured input data series. For this, discharge measurements are needed. These measurements are used as a reference.

Errors in input and output series may result in errors in estimated parameters and hence errors in simulated discharge. In particular, the effect of sampling errors in precipitation on the estimated parameters and simulated discharge has frequently been studied. The effect of discharge determination errors is less often investigated. More information about the influence of discharge determination errors on model performance and parameter estimation of a hydrological model can direct future discharge determination methods and research and may improve short- and long-term discharge predictions.

The HBV model (Bergström, 1976) is an example of a hydrological model which is used in this research. HBV is a conceptual, rainfall-runoff model and can be used as a semi-distributed or lumped model (Liden and Harlin, 2000; Lindström et al., 1997). Because there has not been much research in the past that is aimed at uncertainties in hydrological modeling due to measurement errors, the choice for a conceptual model with a simple structure is made. The HBV model uses precipitation, potential evapotranspiration and temperature as input variables. The simulated discharge is the output of the model.

1.2 Previous research

As indicated in paragraph 1.1, in the past there has not been much research about the influence of errors in discharge determination on model performance and parameter estimation of hydrological models. In hydrological modeling, often an assumption is made there are no uncertainties in discharge data series or that the presence of uncertainties would not influence the behavior of the model. In this research, the fairness of this assumption is examined.

To investigate the uncertainties in discharge data and the influence of these uncertainties on the calibration of a hydrological model, it is important to learn from previous studies. One important aspect is that information has to be collected about uncertainties in discharge measurement. This part is mainly treated in chapter 3. Furthermore it can be useful to look at studies which focus on the uncertainties in input variables (incorrect or missing data) and their influence on model performance and parameter estimation. Other uncertainties in input and output of a hydrological model can be caused by applying a wrong spatial and/or temporal resolution. Studies that treat these kinds of uncertainties can contain useful elements for this research.

In studies regarding incorrect input data the focus is often on the quality of the precipitation data. An example of this is the research of Andréassian et al. (2001). They presented a method in which the quality of the precipitation data is assessed using quality functions. The *GORE* and *BALANCE* indices assess the quality of precipitation time distribution and the total depth respectively. The used hydrological models were GR3J, TOPMODEL and IHACRES, applied to three river basins, differing in surface area. The overall conclusion of this research was that with improving the quality of input data, the model performance increases.

Several previous studies are aimed at assessing the influence of varying spatial resolution of the rainfall input on model performance. Five of these researches are those from Bárdossy and Das (2008), Dong et al. (2005), Brath et al. (2004), Booij (2002b) and Bormann (2006). The first three were aimed at the distribution of rain gauges in a certain river basin. Bárdossy and Das (2008) investigated the influence of varying the distribution of the rain gauge network on model calibration using the HBV model. The outcome of these researches showed that if the rain gauge network changes, a new calibration of the HBV model parameters has to be performed. Specifically, the calibrated model with dense precipitation input fails when run with sparse precipitation information. On the other hand it turned out that a calibrated model with sparse rainfall information can perform well when run with dense precipitation information. Dong et al. (2005) and Brath et al. (2004) tried to find the optimal number of rain gauges in a catchment. Although different sizes of catchments were used (17 000 km^2 and 1050 km^2) the outcome of both researches was that the optimal number of gauges was five.

The research of Booij (2002b) was aimed at assessing the effects of coupled spatial and temporal basin model resolution and spatial and temporal rainfall input resolution on the response of a large river basin, namely the Meuse River basin (21 000 km^2). The used model was a simple stochastic rainfall model and a river basin model with uniform parameters. The results of the research showed that the effect of the spatial model resolution on extreme river discharge is of major importance as compared with the effect of the input resolution. The highest spatial model resolution seemed to be rather accurate in determining extreme discharge.

Bormann (2006) investigated the effect of spatial input data resolution on the simulated water balances and flow components using a multi-scale hydrological model, named TOPLATS. The conclusion of this research was that using a larger spatial resolution, the model performance decreases.

The studies mentioned above have their focus on uncertainties in input or spatial resolution in hydrological modeling and their influence on model performance. The problem is that no research is aimed at the influence of uncertainties in discharge data on model calibration. Some important elements from the previous studies that can be useful for this research are:

- In order to draw decent conclusions it is useful to focus on multiple watersheds, with differing properties;
- It is useful to use a 'simple' conceptual and/or lumped model, because in that case the influence of uncertainties can be evaluated relatively easy and the calculation time is limited. Furthermore, this research is one of the first studies that focus on uncertainties in hydrological modeling due to discharge measurement uncertainty. That is why it is logical to use a model that has a relatively simple structure;
- In the previous studies several objective functions are used, which assess the quality of a calibration. There are different kinds of objective functions, each with a certain focus on the hydrograph. For this research one or two objective functions have to be used, or they can be combined into one objective function;
- It is important to express the relationship between the magnitude of the uncertainties and the influence on model performance and/or the estimation of parameters.

1.3 Problem statement, objective, research questions and research model

The findings in previous studies lead to a problem that is stated below. This problem can be translated into a general research objective and three research questions.

1.3.1 Problem statement

The problem that is derived from the previous studies is that often an assumption is made that discharge uncertainties do not have any significant influence on model performance and model parameters after calibration. In this research it is examined whether this assumption can be justified.

1.3.2 Objective

The objective of this study is to investigate the influence of uncertainties in discharge determination on the estimation of the parameters and the performance of a lumped version the HBV model for two sub basins in the Meuse River, by applying an automatic global searching calibration method and using adapted observed discharge time series.

1.3.3 Research questions

The objective stated before leads to the following three research questions:

1. Which version and schematisation of the HBV model, which sub basins of the Meuse River and which calibration procedure are most adequate for calibration?

2. What kind of uncertainties in discharge determination can be present and how can these errors be brought into existing discharge time series?

3. What is the effect of uncertainties in discharge determination on model performance and parameter estimation of the HBV model applied to different sub basins of the Meuse River?
    a. What is the effect of uncertainties in discharge determination on model performance of the HBV model, applied to different sub basins of the Meuse River?
    b. What is the effect of uncertainties in discharge determination on the estimation of parameter sets of the HBV model, applied to these sub basins?

1.3.4 Research method

In Figure 1 a simple research model is given. The first two research questions form a foundation in order to be able to answer the third question. The third research question will directly contribute to the objective of the study.

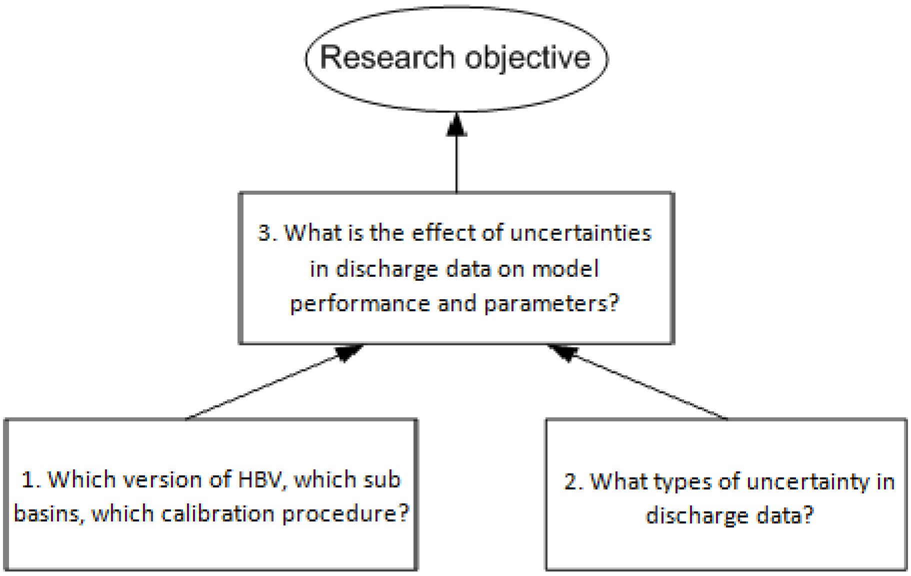


Figure 1: Research model

Step 1 in this research is that data need to be collected and the hydrological model, sub basins and calibration procedure need to be chosen. Also a calibration is performed with the original discharge data. This ‘base case’ serves as a reference for future calibrations.

Step 2 in this research is to make an investigation about all possible uncertainties in discharge data. Subsequently for every error source a method is chosen about how the uncertainty can be integrated into an existing data set. This is done to simulate different kinds of discharge determination uncertainties. In this research an assumption is made that the original discharge data do not contain uncertainties or errors.

Step 3 of the research consists of answering the third and most important research question. The adapted discharge series which are a result of step 2 are used to perform calibrations. After that, the model performance and the behavior of the model parameters are examined. The original ‘base case' calibration is used as a reference to assess the influence of the adapted discharge data on model performance and parameter sets of the HBV model.

1.4 Outline of the report

In Chapter 2 an answer is provided for research question 1. In this chapter the data collection and the reference HBV model is presented. At first the type and sources of data are explained. Also a choice for two sub basins of the Meuse is made and explained for these sub basins. After that, a description of the used rainfall runoff model, the HBV model, is given. Subsequently the calibration method that is used in the research is introduced. At the end of Chapter 2, the calibration and validation of the HBV model in the two sub basins is performed. This calibration and validation form the 'base case', i.e. a reference for all following calibrations.

In Chapter 3 research question 2 is treated. This chapter contains some theory behind uncertainties in discharge determination. First, the origins of errors in discharge time series are explained. After that, a distinction is made between different types of errors. These different types of errors can be combined to several error sources. These four error sources are explained further and also a method is presented to introduce these errors into the original discharge data. These artificially constructed discharge time series are used subsequently to perform calibrations. The quality of the adapted discharge data series is quantified by two quality functions, which are defined at the end of Chapter 3. A part from the four error sources, a realistic scenario is developed.

In Chapter 4 the third and main research question is answered. In this chapter the results are presented and discussed. For each error source the influence of the errors on model performance and model parameters is shown. Also the results from the realistic scenario are analyzed and discussed.

The model results lead to a discussion in Chapter 5. After that, conclusions are drawn and recommendations for future research are proposed in Chapter 6.

2 Data collection and reference HBV model

In this chapter, the collection of the used data and the set up of the reference HBV model are discussed. In paragraph 2.1 the schematization of the chosen sub basins and the used data can be found. Paragraph 2.2 gives a description of the used hydrological model, the HBV-15 model. In paragraph 2.3 the used calibration procedure is explained. Paragraph 2.4 contains the calibration and validation of the base case, i.e. the situation with the original data, which results in a reference HBV model.

2.1 Data collection and schematizations sub basins

In Figure 2 the Meuse River Basin is shown (Riou vzw, 2010). The Meuse Basin is located in France, Luxemburg, Belgium, Germany and the Netherlands.

Figure 2: Meuse River Basin

In Figure 3 a schematization for the Meuse River Basin, upstream of Borgharen is given (Booij, 2005). The Meuse River Basin upstream of Borgharen can be de divided into several sub basins. The following 15 sub basins can be distinguished.

The used climate data in this research are from Météo France (French sub basins) and the Belgian Meteorological Institute (Belgian sub basins). The Meuse Basin data are from RIZA and the discharge data come from Rijkswaterstaat Limburg.

1: Meuse Lorraine sud
2: Chiers
3: Meuse Lorraine nord
4: Bar-Vence-Sormonne
5: Semoins
6: Viroin
7: Meuse midi
8: Lesse
9: Sambre
10: Ourthe
11: Ambleve
12: Vesdre
13: Mehaigne
14: Meuse nord
15: Jeker

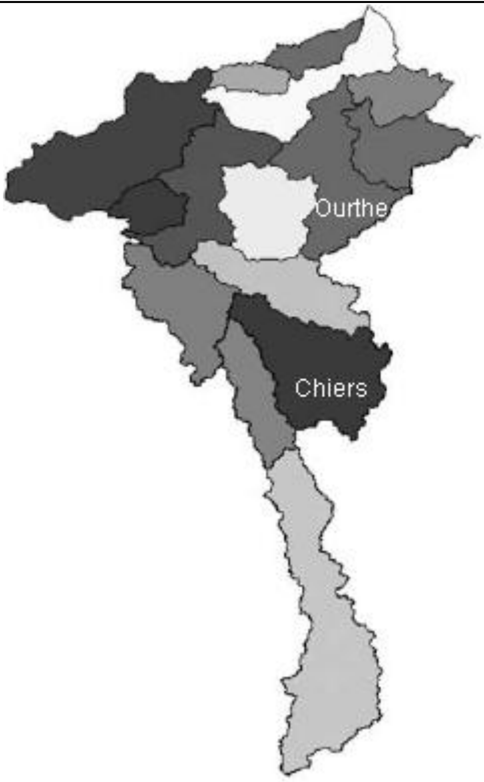


Figure 3: Entire Meuse River basin upstream of Borgharen, containing Ourthe and Chiers (Booij, 2005)

A longitudinal profile of the Meuse River and its main tributaries is shown in Figure 4. In this figure the length and slopes of the tributaries can be found.

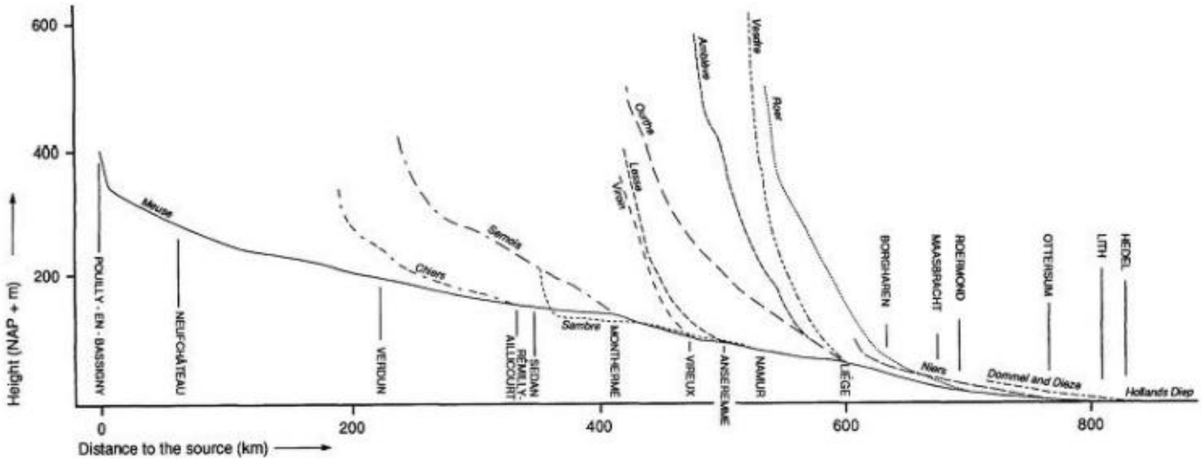
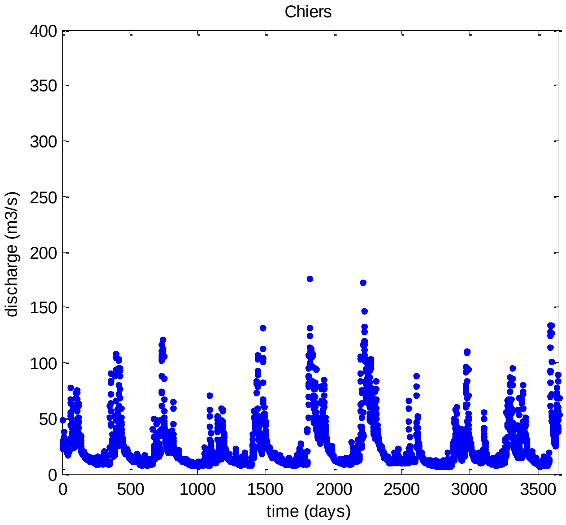


Figure 4: Longitudinal profile of the Meuse River and its main tributaries (Berger, 1992)

In the research the focus will be on two of these sub basins, namely Ourthe and Chiers. The choice for these two basins has several reasons. There are some similarities and some differences between the sub basins. Firstly, both basins have no inflowing runoff from upstream basins. This means that the only inflow of water into the system is from precipitation. The inflow from groundwater flow is neglected in this model. Furthermore, both sub basins have a surface area with a size that is of comparable order (Ourthe 1597 km^2 and Chiers 2207 km^2).

A difference between the basins is the average slope of the rivers, as shown in Figure 4. The slope of the Ourthe River is significantly steeper than the slope of the Ourthe River. Another difference, which is probably related to the slope of the rivers, is the shape of the discharge time series. While the Ourthe has a relatively low base flow and high peaks, the Chiers River has less high peaks and a higher base flow. This can be seen in Figure 5. In this figure daily measurement values of the discharge are shown for 10 years (1989-1998) for the Chiers and Ourthe rivers. The steeper slope of the Ourthe basin may be the reason for the extremer discharge regime of the Ourthe.



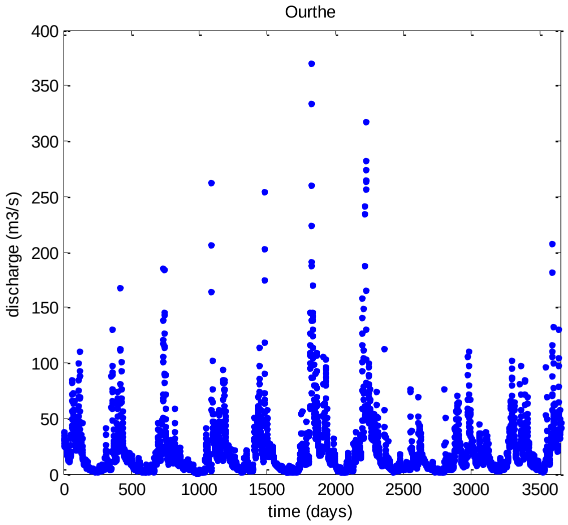


Figure 5: Discharge graphs Chiers and Ourthe Rivers

The average discharge in the basins is 23,0 and 25,5 m^3/s respectively for the Ourthe and Chiers river, while the standard deviations of the discharges are respectively 29,8 and 23,4 m^3/s. These values indicate that the average discharge is higher in the Chiers River, but that the Ourthe River shows a more extreme behavior.

The climate data, such as daily average precipitation, temperature and potential evapotranspiration, are of comparable magnitude between the basins. This is shown in Table 1. The mean and standard deviation from the available daily data are calculated and it turns out that there is not much difference in input variables between the basins.

Table 1: Properties of climate data in Ourthe and Chiers rivers

Climate data		***Ourthe***	***Chiers***
Precipitation [mm]	Mean	2.7	2.6
	Standard deviation	4.7	4.7
Temperature [^oC]	Mean	8.6	9.2
	Standard deviation	6.5	6.6
PET [mm]	Mean	1.5	1.6
	Standard deviation	1.5	1.3

2.2 HBV model

The used hydrological model in this research is the HBV model. This model is developed in 1972 by Bergström. Initially, the model was developed for the forecasting in hydropower developed rivers of

Scandinavia (Bergström, 1976). Since then, the model has been applied in more than 50 countries and regions in the world (Bergström, 1995).

In this research, a lumped version of the HBV model is used, the so called HBV-15 model, developed by Booij (2002a). In the research of Booij (2002a) 15 sub basins of the Meuse River upstream of Borgharen were used. In the HBV-15 model, each of the 15 sub basins is schematized without spatial variability. So, for every sub basin a lumped model is used and these 15 lumped models are linked and together form the HBV-15 model. In this research just two of these sub basins are considered, so two of these lumped models are used. In the following sections a description of this version of the lumped HBV model is given. For this research, the HBV model is programmed into MATLAB.

2.2.1 Description HBV model

In general, the HBV model is a conceptual, rainfall-runoff model and can be used as a semi-distributed or lumped model. It was developed at the Swedish Meteorological and Hydrological Institute (SMHI) in the early 70s (Bergström, 1976; SMHI, 2003). A lumped version of the HBV model is chosen for this research because of its conceptual, simple model structure. Because there has not been much research in the past that considers uncertainties due to discharge measurement errors, a choice is made for a simple model which is easy to use. A semi-distributed model is used if a basin can be separated into a number of sub basins and that each one of these is distributed according to elevation and vegetation (Singh, 1995). In this research the choice for a lumped version of the HBV model for the Ourthe and the Chiers rivers is made and it does not take into account the spatial variability of processes, input, boundary conditions and watershed geometric characteristics.

The used HBV is called conceptual because the model parameters do not directly represent physical properties. That is why model parameters cannot be measured in the field. The model parameters, which indirectly represent the basin characteristics, are determined by calibration of the model. The advantage of a conceptual model is that it has a relatively simple model structure. A disadvantage is that most parameters are empirical, which may reduce the validity of the model.

The HBV model uses precipitation, potential evapotranspiration and temperature as input variables. The simulated discharge is the output of the model. The used time step is one day, because the discharge and climate data contain daily values. There are eight model parameters which are used for the calibration. The model contains three routines which describe the most important runoff processes. In Figure 6 a schematization of the model including the three routines is shown. Also the location of the input (green), parameters (red) and output (black) can be found in this schematisation. In the following section these routines as well as the used parameters are discussed.

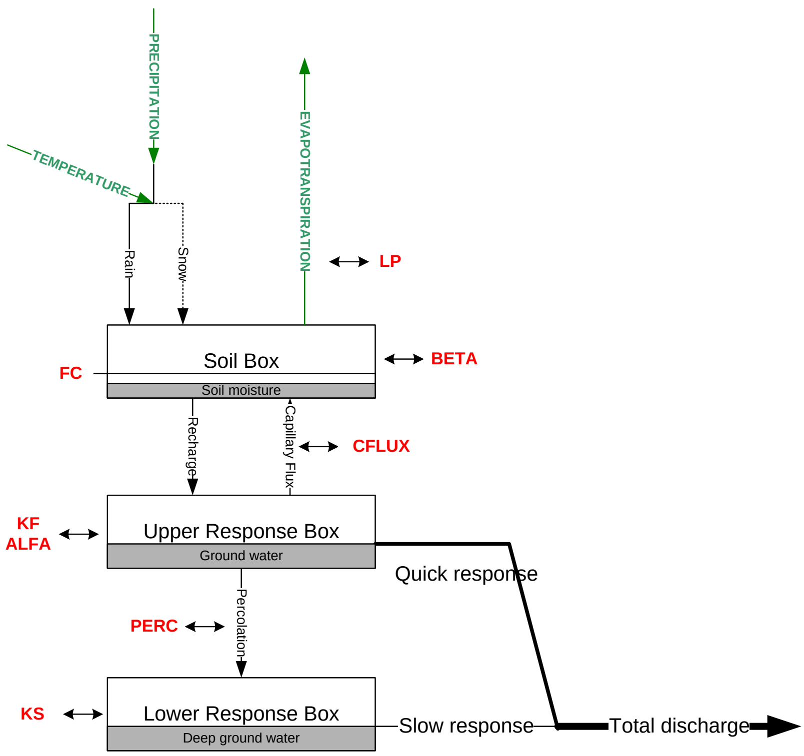


Figure 6: Schematization of the used HBV model

2.2.2 Description of the lumped HBV model

Figure 6 shows the schematization of the model. The lumped version of the HBV model consists of three routines: the precipitation routine, the soil moisture routine and the runoff generating routine, which can be divided into quick and slow runoff. These routines are discussed below (Bergström, 1976).

Precipitation Routine

The precipitation, which is the initial input of the model, is divided into rainfall and snowfall. If the temperature is above a certain threshold, rainfall will be present. Below this threshold, the precipitation consists of snowfall. Also the melting and refreezing processes are taken into account in this routine.

Soil Moisture Routine

This routine controls which part of precipitation is evaporated or stored in the soil. The ratio of actual soil moisture (*SM*) and the maximum water storage capacity of the soil (parameter *FC [mm]*), and the soil routine parameter (*BETA [-]*) together assess the runoff coefficient. With this runoff coefficient, the part of the precipitation *P* which forms the recharge *R* to the upper response box can be calculated, by using equation (1). If the soil is saturated (*SM=FC*), then the recharge is equal (if *BETA*=1) or larger (if *BETA*>1) than the precipitation, dependent on the value of *BETA*.

$$R(t) = \left(\frac{SM}{FC}\right)^{BETA} * P(t) \tag{1}$$

LP [-] describes the limit of water storage for potential evapotranspiration. Above this limit, potential and actual evapotranspiration will be equal to the potential evapotranspiration. Data for potential evapotranspiration are input for the HBV model. The parameter *CFLUX [mm day^{-1}]* represents the maximum capillary flux from the runoff routine into the soil.

Runoff Generation Routine (quick and slow runoff)

The runoff generation routine is the response function that transforms excess water from the soil routine to runoff. The runoff generation routine consists of two reservoirs. The first one, the upper response box, is a non-linear reservoir which represents the quick runoff. *KF* [day^{-1}] is a recession parameter in the upper or fast response box. *ALFA* [-] is a measure for non-linearity of the quick runoff.

The second reservoir is the linear lower response box. This box represents the slow response (with recession coefficient *KS* [day^{-1}]), i.e. the base flow that is fed by groundwater. The fast (Q_f [mm/day], equation (2)) and slow (Q_s [mm/day], equation (3)) response can be characterized by the following equations, in which S_f [mm] and S_s [mm] represent the storage in respectively the fast and slow response box.

$$Q_f(t) = KF * S_f(t)^{(1+ALFA)} \quad \textbf{(2)}$$

$$Q_s(t) = KS * S_s(t) \quad \textbf{(3)}$$

Groundwater recharge is ruled by a maximum amount of water that is able to penetrate from soil to groundwater (parameter *PERC* [mm day^{-1}]) through the upper response box.

2.3 Calibration procedure

For the calibration procedure an optimization method and an objective function have to be chosen for the research. Choices for these elements of the calibration are explained in the following sections.

2.3.1 Optimization method: SCEM-UA

The used method for model optimization is the SCEM-UA algorithm. This method has been developed by Vrugt et al. (2003a). SCEM-UA is an automatic global searching method which is based on the SCE-UA algorithm (Singh, 1995). Instead of using the Downhill Simplex method that is used in the SCE-UA algorithm, an evolutionary Markov Chain Monte Carlo (MCMC) sampler is used. This means that a controlled random search is used to find the optimum set of parameter values in the parameter space. The choice for the SCEM-UA method is based on the fact that it is an automatic global searching method that converges quite fast to the optimal value. An advantage of this algorithm is that the chance of finding the global optimum is very high. In Appendix 1 more information about the SCEM-UA method can be found. The SCEM-UA method is also programmed into MATLAB and linked with the HBV model. This makes that all optimizations in this research take place in the MATLAB program.

First a calibration is performed with all eight parameters. The number of iterations of the SCEM-UA algorithm in this first calibration is 4000. After that, a sensitivity analysis is performed to determine which parameters have a large influence on the objective function in these basins. To reduce calibration time, only the most influencing parameters are used in the calibrations further on in this research. The parameters which do not have much influence on model performance get a fixed value. This results in a smaller calibration time, as the SCEM-UA method needs less iterations to find the optimum. When an optimum is found after calibration, this model with parameter set and objective function is used as a reference or 'base case' in this research.

2.3.2 Objective function

There are different kinds of objective functions which can determine the model performance given a certain parameter set. In calibrations single or combined objective functions can be used. A single objective function is an objective function which is aimed at a specific property of the hydrograph. Some objective functions for example assess the quality of the shape of the hydrograph or a correct water balance over the entire calibration period. Other functions evaluate the quality of specific parts of the hydrograph, such as peak flows or low flows. In this research it is important to have a good representation of the entire hydrograph. It is also useful to have a correct water balance. This is why a combined objective function will be used. This combined objective function combines the single objective functions *NS* and *RVE* (Nash and Sutcliffe, 1970). The functions *NS* and *RVE* are shown in equation (4).

$$NS = 1 - \frac{\sum_{i=1}^{N} \left(Q_{sim}(t_i) - Q_{obs}(t_i) \right)^2}{\sum_{i=1}^{N} \left(Q_{obs}(t_i) - \overline{Q_{obs}} \right)^2} \quad \text{and} \quad RVE = \frac{\sum_{i=1}^{N} \left(Q_{sim}(t_i) - Q_{obs}(t_i) \right)}{\sum_{i=1}^{N} \left(Q_{obs}(t_i) \right)} \qquad \textbf{(4)}$$

Where $Q_{obs}(t_i)$ and $Q_{sim}(t_i)$ are observed and simulated daily discharge at time step t_i respectively and $\overline{Q_{obs}}$ is mean observed daily discharge and N is the total number of time steps. *NS* assesses the quality of the shape of the hydrograph and has a value of 1 if a perfect match in hydrograph is present, while *RVE* is aimed at the relative volume difference between the observed and simulated discharge and has an optimal value of 0.

The combined objective function used in this research is called *y* (Akhtar et al., 2009) and defined as follows:

$$y = \frac{NS}{1 + |RVE|} \qquad \textbf{(5)}$$

In case of an optimum, *NS* has a value of 1 and *RVE* has a value of 0. This makes that the optimal situation has a value of 1 of the objective function *y*.

2.3.3 Calibration time period

There are climate and discharge data available for these basins over a 31-year period, from 1968 to 1998. It is important to have a large period available for an adequate calibration. The calibration is performed over a period of 15 years, from 1984 to 1998. The other data, in the 16-year period from 1968 to 1983, is used for the validation. It is possible that there are not data available for the entire 16-year validation period, for example in the Chiers basin. In that case, less than 16 years of data are used for validation.

2.4 Calibration and validation results, reference HBV model

In this paragraph the reference HBV model is set up. The calibrations of the model with the original data of the two sub basins form the 'base cases' or reference models for this research. Therefore at first the calibration with all eight parameters takes place and after performing a sensitivity analysis this calibration is further optimized and subsequently validated for both basins.

2.4.1 Calibration with 8 parameters

In the first calibration, eight model parameters are used to optimize the model. Some parameters are more sensitive than other parameters. In other words, some parameters have a larger influence on the objective function than other parameters. That is why these parameters need more attention in a calibration. To select the most influential parameters, a univariate sensitivity analysis is performed. This means that the variables are varied once at a time while the other parameters keep a constant value. In Table 2 an overview of the parameter values and ranges is given. The ranges of the Ourthe basin are based on the research of Booij and Krol (2010). Initially, the ranges for the Chiers basin were also coming from this research, but it seemed that this did not deliver the maximum value for the objective function, because for some parameters the optimal parameter value was situated at the border of the parameter range. This might be an indication that the real optimal parameter value is lying outside the parameter range. To solve this problem, the parameter range is changed in a way that the optimal value is not at the boundary of the range any more. The parameter ranges after the modifications still contain realistic values and therefore are suitable for calibration.

Table 2: Parameter values

Parameter [unit]	*Physical interpretation*	*Range Ourthe Basin*	*Range Chiers Basin*	*Optimal value Ourthe*	*Optimal value Chiers*
FC [mm]	Maximum soil moisture storage	150 – 500	400 – 700	224	485
BETA [-]	Shape parameter of runoff generation	1 – 3	0.9 – 1.5	1.907	1.157
LP [-]	Fraction of *FC* above which potential evapotranspiration occurs and below which evapotranspiration linearly reduces	0.2 – 1	0.1 – 1	0.388	0.259
ALFA [-]	Measure of non-linearity for fast flow	0.1 – 1.5	0.05 – 0.5	0.505	0.202
KF [day^{-1}]	Recession coefficient for fast flow reservoir	0.005 – 1	0.005 – 1	0.0219	0.0328
KS [day^{-1}]	Recession coefficient for slow flow reservoir	0.005 – 1	0.005 – 1	0.0069	0.005
PERC [mm day^{-1}]	Drainage from the fast flow reservoir to the slow flow reservoir when sufficient water is available	0.1 – 2.5	0.1 – 2.5	0.569	0.326
CFLUX [mm day^{-1}]	Maximum value for capillary flow	0.1 – 2.5	0 – 2.5	1.363	0.454
		Value objective function		0.933	0.759

2.4.2 Sensitivity analysis

After the calibration with eight model parameters, a univariate sensitivity analysis has been performed for both basins. This analysis is done to investigate which parameters are most sensitive, i.e. which parameters have the largest influence on the model performance if the parameters would change. The parameters which are less sensitive on model performance will get a fixed value in further calibrations in this research, because this leads to a decrease in calibration time. In the sensitivity analysis each parameter is varied one at a time, while the other seven are kept constant. The variations of the parameters influence the model in a way which results in a change in objective function. In Appendix 2 the outcome of the sensitivity analyses of both sub basins is shown.

In both the Ourthe and Chiers basins the most sensitive parameters are *ALFA*, *FC*, *LP*, *BETA* and *KF*. These parameters are chosen to optimize the calibration. *CFLUX*, *PERC* and *KS* are not sensitive in the Ourthe and Chiers basins with the present climate and discharge data and will get the values as shown in Table 2 as a fixed value in the calibrations in this research.

In a study of Booij and Krol (2010) the three parameters *ALFA*, *FC* and *LP* are considered the most identifiable for the Ourthe and the Chiers basins. This means that these parameters are most sensitive to a certain combined objective function in which the single objective functions *NS*, *RVE*, NS_L and NS_H (Nash Sutcliffe coefficients for relatively low and high flows) (Nash and Suttcliffe, 1970) were included.

In this research, the next most sensitive parameters in the two basins are *BETA* and *KF*. In this sensitivity analysis, *ALFA*, *FC*, *LP*, *BETA* and *KF* also turned out to be the most sensitive parameters.

2.4.3 Optimization by calibration of the five most sensitive parameters

As seen in section 2.4.2 the parameters *FC*, *BETA*, *LP*, *ALFA* en *KF* are used in the calibration of both the Chiers and Ourthe basins. In the sensitivity analysis, the used ranges were rather large. To perform a more accurate calibration, the ranges are made narrower. The minimum and maximum values of the parameters are based on the parameter values from the first calibration (shown in Table 2). The minimum value is 80% of the original value, and the maximum value is 120% of this value. In Table 3 the ranges of the five calibration parameters that are used for the final calibration are shown. The other three parameters get a fixed value which is also shown in the table. The maximum number of iterations in the SCEM-UA algorithm in this calibration is 1500.

Table 3: Parameter values and ranges after calibration

Parameter [unit]	*Range/value Ourthe Basin*	*Range/value Chiers Basin*		*Optimal value Ourthe*	*Optimal value Chiers*
FC [mm]	179 – 268	388 – 583		211	520
BETA [-]	1.52 – 2.29	0.90 – 1.40		1.71	0.94
LP [-]	0.31 – 0.47	0.20 – 0.32		0.310	0.202
ALFA [-]	0.40 – 0.61	0.16 – 0.25		0.559	0.175
KF [day^{-1}]	0.017 – 0.027	0.026 – 0.040		0.0175	0.0388
KS [day^{-1}]	0.0069	0.005		0.0069	0.005
PERC [mm day^{-1}]	0.569	0.326		0.569	0.326
CFLUX [mm day^{-1}]	1.363	0.454		1.363	0.454
	Values objective functions		***y***	**0.934**	**0.767**
			NS	0.934	0.767
			RVE	0.002 %	0.000 %

Table 3 shows the values of the parameters after optimization, as well as the values for the combined objective function *y* and the functions *NS* and *RVE*. The Ourthe basin has a significantly higher value for the objective function compared to the Chiers basin. This means that the HBV model has more difficulties in simulating the discharge of the Chiers basin compared to the Ourthe, given the climate data. The reason for that can be that the quality of the data in the Ourthe basin is better, or that the HBV model is more suitable for rivers with a relatively low base flow and high peaks like the Ourthe, i.e. more extremes and variability.

The value for *RVE* is very small in both basins. The use of this particular combined objective function *y* implies that it is important that the water balance over the entire calibration period corresponds with the real situation.

The parameters are of the same magnitude as in the research of van Deursen (2004). In the study of van Deursen several sub basins of the Meuse River are calibrated by using the HBV model. There are some small differences in the exact parameter values. These differences occur because of a different calibration method (Monte Carlo method), a different objective function (Nash Sutcliffe coefficient) and by using a semi-distributed model in that research.

In Figure 7 and Figure 8 show the outcome of the calibrations of the two basins. The blue dots represent daily discharge measurements and the red graph represents the simulated data.

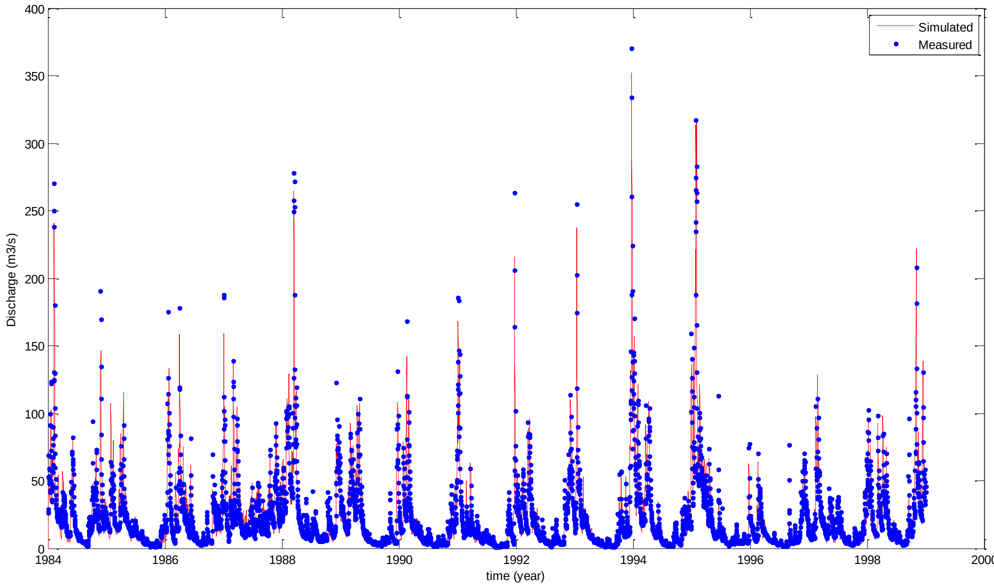


Figure 7: Calibration Ourthe basin period 1984 – 1998

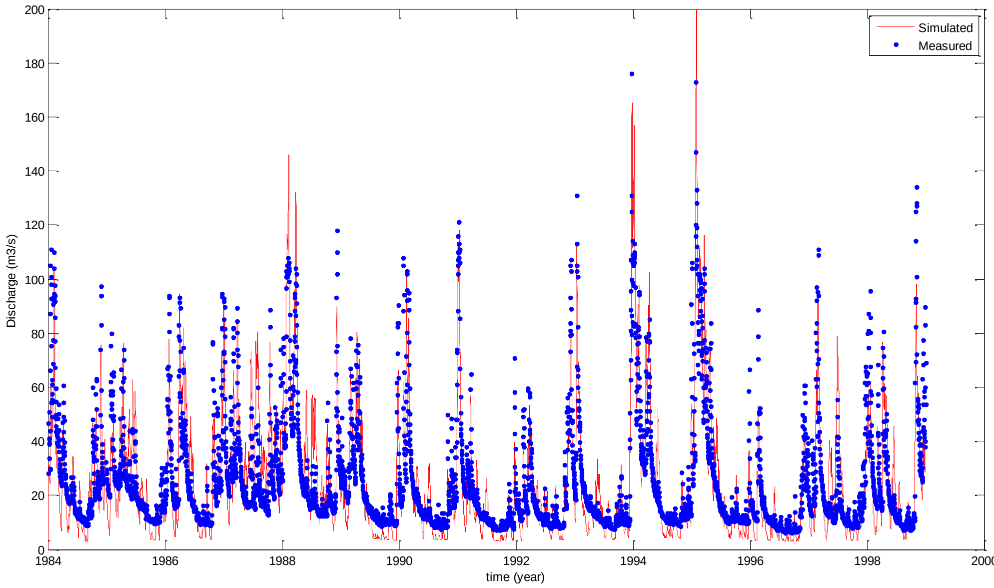


Figure 8: Calibration Chiers basin period 1984 – 1998

These figures show that the discharge data in the Ourthe basin are simulated better than in the Chiers basin. This is consistent with their values of the objective functions after calibration. In both basins the observation can be made that the model finds some difficulties in the peaks and the low flows. The peaks tend to be underestimated by the model, while the low flows are often overestimated.

All parameters in the HBV model have a physical background. Below, for each parameter the comparison of the differences in parameter value between the basins is made. In paragraph 2.1 the differences between the hydrographs of the two basins are explained. Figure 5 showed that the Ourthe has a relatively low base flow and high peaks, the Chiers river has less high peaks and a higher base flow.

FC

The Chiers basin has a higher value for *FC* than the Ourthe basin. This means that maximum soil moisture storage is higher in the Chiers basin. This higher value means that there is a possibility of buffering if there is a lot of precipitation. This will lead to less high and steep peaks in the Chiers river compared to the Ourthe. The discharge data of the two rivers confirms this expectation.

BETA

BETA is a parameter which describes the relative contribution of precipitation from the soil moisture box to the upper response box (recharge) if the soil moisture box is not saturated. If *BETA* has a high value, the recharge is smaller in case of a moisture deficit in the soil box than with a low value of *BETA*. The value for *BETA* is highest in the Ourthe basin, which means that given a certain soil moisture deficit and precipitation, less recharge will occur in this basin compared to the Chiers basin. As a result, the upper response box in the Chiers will be filled faster than in the Ourthe basin. It is difficult to clarify what is the influence of this on the discharge regime.

LP

LP represents the fraction of the maximum soil capacity (*FC*) above which potential evapotranspiration occurs and below which evapotranspiration linearly reduces. So if *LP* is high, the chance that the actual evapotranspiration is equal to the potential evapotranspiration is smaller than in cases *LP* has a low value. So in the basins with a high *LP*, most of the time the actual evapotranspiration is less than the potential evapotranspiration. Because *LP* is a fraction of *FC*, the effect of a certain value of *LP* also depends on the value of *FC*. The Ourthe has a bigger value for *LP* than the Chiers basin, but also has a smaller value for *FC*, which indicates that the Ourthe basin has a smaller field capacity. *LP* has a slightly higher value in the Ourthe basin compared to the Chiers, but since the value of *FC* is smaller in the Ourthe it is difficult to determine in which basin the most evapotranspiration takes place and what is the influence of this on the hydrograph.

ALFA

The value of *ALFA* is higher in the Ourthe basin. This means that at a certain change in storage in the upper response box the quick response has a more non-linear behavior in the Ourthe basin compared to the Chiers basin. This means that if the storage increases, so if precipitation is present, the quick discharge faster increases. In these basins the Ourthe has the highest and steepest peaks and the parameter *ALFA* confirms that.

KF

The value of *KF* is higher in the Chiers basin. This means that if the storage in the upper response box is small, the quick runoff in the Chiers basin is higher than at the same storage in the Ourthe basin. Due to the value of *ALFA* this relation is not present if the storage in the upper response box is high.

Considering the hydrographs of the rivers, it makes sense that *KF* has a higher value in the Chiers basin, because it has larger discharges at low flows compared to the Ourthe basin.

KS

For the parameter *KS* a same conclusion can be drawn. In contrast with *KF*, this parameter has a linear relation with the slow flows from the lower response box. The values do not vary much between the two basins. The Ourthe basin has a higher value. The effect of this value on the discharge regime is difficult to determine.

PERC

The value of *PERC* is higher in the Ourthe basin. This means in the Ourthe basin there is more percolation from the upper response box into the lower response box. It is difficult to estimate the influence of the difference between the basins on the hydrograph.

CFLUX

The value of *CFLUX* is considerably higher in the Ourthe basin, so in this basin there is more capillary flux from the upper response box into the soil moisture box. So if the soil moisture is smaller than the field capacity, the vertical flux from the upper response box into the soil is bigger in the Ourthe basin.

The influences of the values of the last three parameters on the hydrograph are difficult to clarify. However these parameters are less important in this research, as they are not of importance in the rest of this research, because their insensitivity in these basins.

In Figure 9 and Figure 10 the development of the parameters is shown for respectively the Ourthe and Chiers basin. The vertical axis represents the parameter value and on the horizontal axis the number of iterations is shown, which is set to a maximum number of 1500.

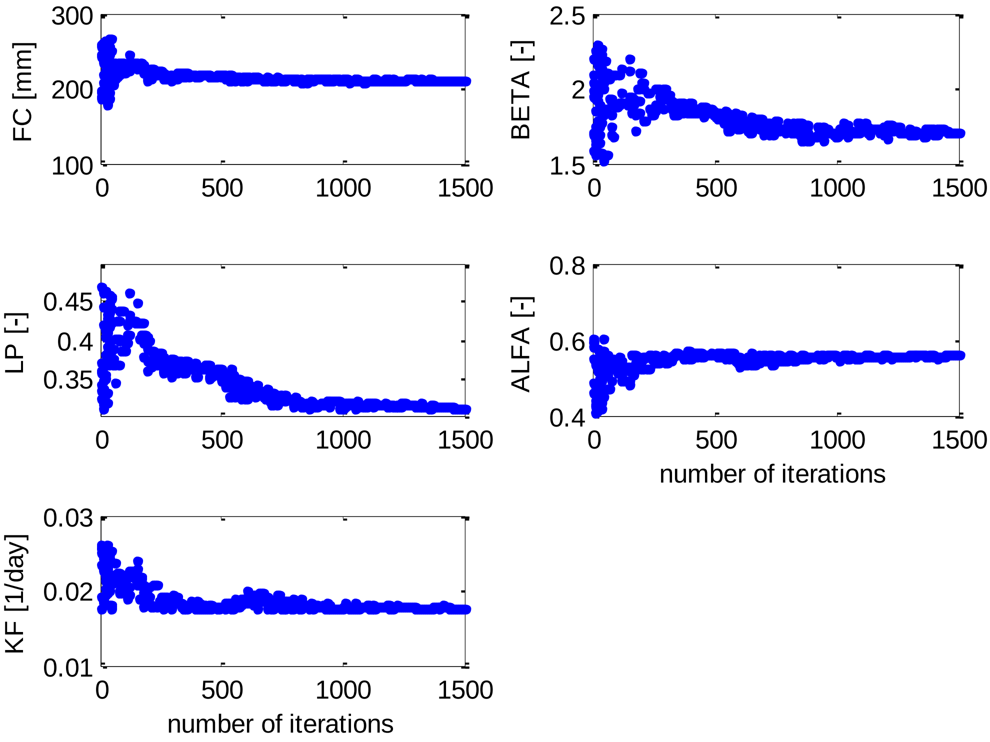


Figure 9: Development of the parameters in the calibration of the Ourthe basin

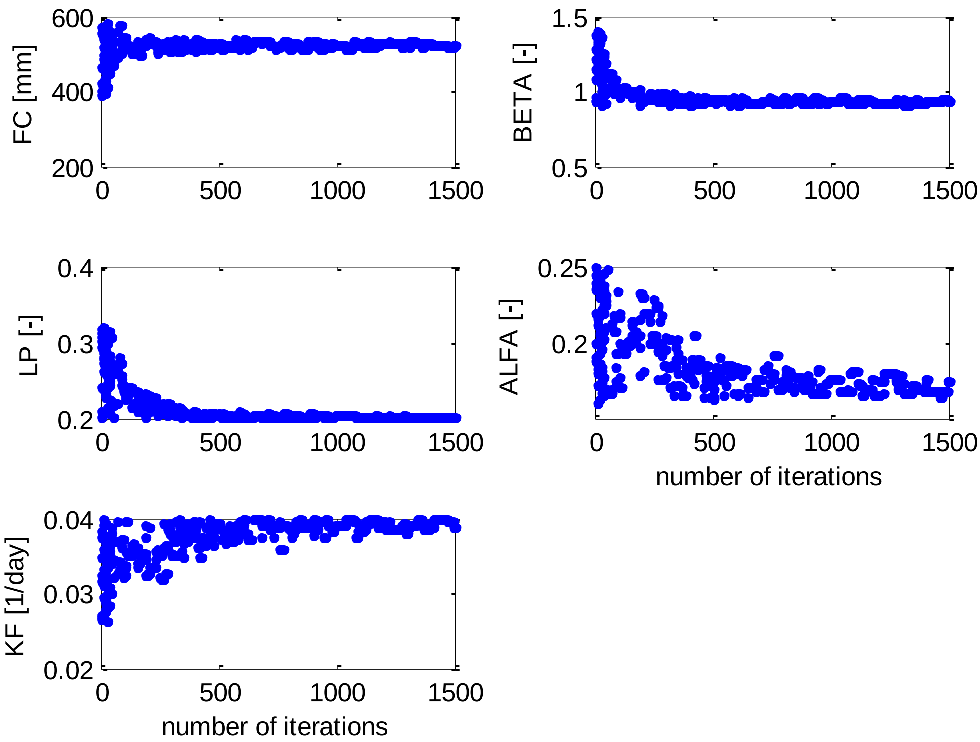


Figure 10: Development of the parameters in the calibration of the Chiers basin

The figures show that in the Ourthe basin the parameters *FC*, *ALFA* and *KF* have a strong convergence in the beginning of the calibration, while the optimizing algorithm seems to have more difficulties in determining *BETA* and *LP*. *FC*, *ALFA* and *KF* are therefore well identifiable.

In the Chiers basin *FC*, *BETA* and *LP* show a strong convergence and thus are well identifiable, while *ALFA* and *KF* show a weaker convergence and are less identifiable.

2.4.4 Validation

In the validation, the parameter values of the calibration are used. The validation period is from 1968 to 1983. In this time span there are a couple of periods in the Chiers basin in which no measurement data are present.

In Figure 11 and Figure 12 the outcome of the validation of respectively the Ourthe and the Chiers basin is shown. The blue dots are the daily determined discharge values and the red graph is the simulated discharge. If discharge data are unavailable, no blue dots are present. The figures also show the values of the objective functions *NS*, *RVE* and *y*. The objective functions are only calculated over the periods in which measurement data are present.

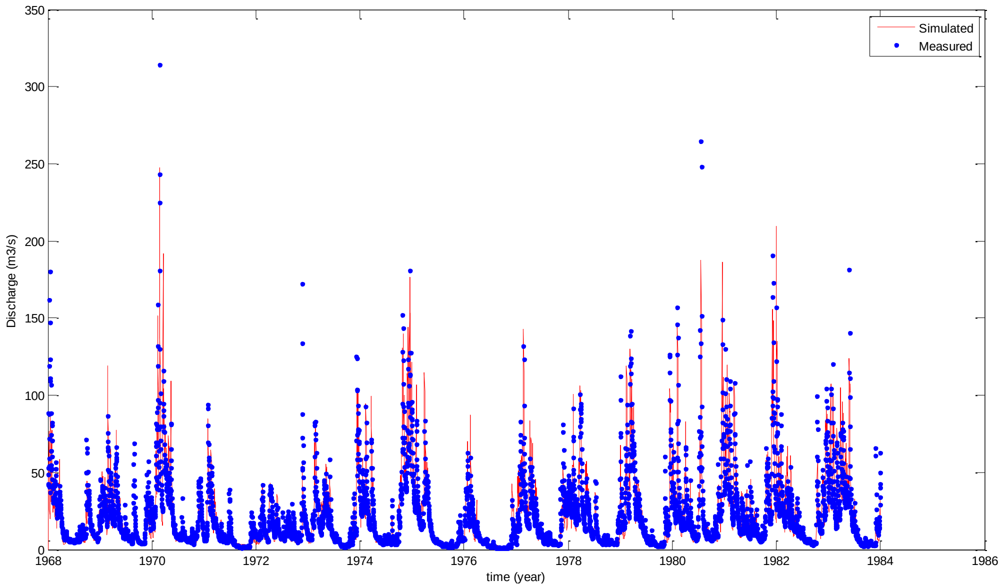


Figure 11: Validation Ourthe basin period 1968 – 1983

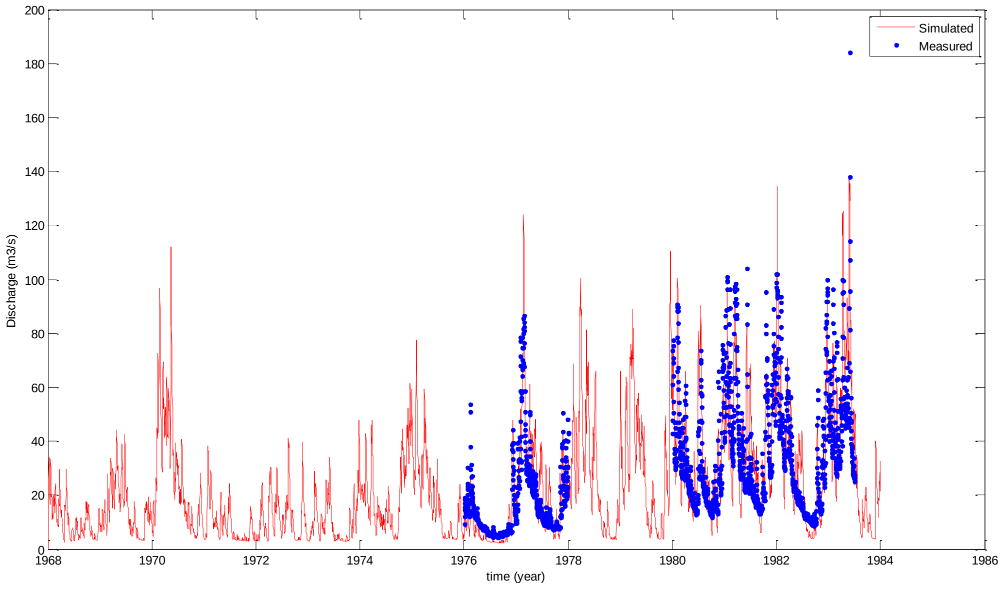


Figure 12: Validation Chiers basin period 1968 – 1983

In Table 4 the values of the combined objective function, as well as the functions *NS* and *RVE* are shown. The values of the objective functions in the calibration of the base case are shown between brackets.

Table 4: Values of the objective functions after validation (values calibration between brackets)

	Ourthe		*Chiers*	
y	**0.842**	(0.934)	**0.752**	(0.767)
NS	0.849	(0.934)	0.785	(0.767)
RVE	0.90 %	(0.002%)	4.44 %	(0.000%)

The validation over the period 1968 to 1983 yields an objective function *y* of 0.842 for the Ourthe basin, while the objective function in the calibration had a value of 0.934. However, 0.842 is still a good value for the validation. In the Chiers basin, the objective function y has a value of 0.752. This value is comparable with the objective function in the calibration, which was 0.767.

In the period 1968 – 1983 not much data are present in the Chiers basin. However, in section 2.4.1 it is explained that the data in the period 1984 – 1998 are used for calibration because it is important to have a large amount of data for an adequate calibration. The total amount of days in the validation of the Chiers basin is 2011 days, or 5.5 years, divided over two periods. It is assumed that these periods contain enough data to perform a decent validation. In general, the conclusion can be drawn that the validation results for both the Ourthe and Chiers basin are rather good. The validation of the Ourthe is good, because the objective function of 0.842 is quite high. The validation of the Chiers is satisfactory because the objective function value does not differ much between calibration and validation.

3 Methodology: uncertainties in discharge determination

Uncertainties in discharge time series are present due to errors in discharge determination. In paragraph 3.1 a number of possible types of errors are discussed. In paragraph 3.2 these errors are classified into different sources of errors. These error sources contain one or a couple of the possible errors. This classification is set up to construct discharge time series including errors. These artificial discharge series are used in the research to analyze the influence of the error sources on model performance and model parameter estimation. A part from the classification of the errors, paragraph 3.2 also contains descriptions of integrating the types of errors into the discharge series. An assumption is made that the original discharge data do not contain errors. The different error sources lead to a “realistic scenario”, which is expected to approach the real situation best. This realistic scenario is treated in paragraph 3.3. In paragraph 3.4 a method for quality assessment of these discharge series is presented. These classification functions give an estimation of the quality of the artificial time series compared to the original time series.

3.1 Errors in discharge time series

There are different kinds of errors which can cause deviations form the real discharge time series. In general, there are two types of errors. The first type of error occurs at the execution of the measurement. This is explained in section 3.1.1. Furthermore, the use of the Q-h relation can bring errors into discharge data, which is clarified in section 3.1.2.

3.1.1 Measurement errors

Errors which originate during the execution of the measurement are referred to as measurement errors. Measurement errors can have different causes. The estimate of the discharge at a certain moment is determined by measurements of the water level, cross section and/or velocity. Uncertainties that can occur during the measurements of these elements are (Jansen, 2007):

- Uncertainty in measured data (empirical quality);

The difference between measured discharge and actual discharge is caused by random and systematic errors. The random error can be indicated as the spread in the measurement. The random errors can occur with or without a certain autocorrelation.

- Uncertainties regarding the executing of the measurement (methodological quality) ;

There are different reasons for a bad execution of the measurements. In extreme conditions like high water levels, standard procedures cannot be followed and errors can occur. Other reasons for a bad execution can originate if a certain flow is completely different from uniform flow, while an assumption is made that the flow is uniform.

- Uncertainties regarding the performance of the measuring equipment;

As a consequence of breaking down or failing of measurement equipment, an instrument can malfunction.

Other uncertainties can originate from less common sources (Singh, 1995) like methods for transferring the data to computer-ready media, approximations in formulae used to convert field observations to volume estimates, negligence and fraud. In this research, the focus lies on measurement errors due to empirical and methodological quality.

Measurement errors can be systematic or random, but in reality a combination of random and systematic errors might be present. A random error can be completely random or can contain autocorrelation. This means that the error at a certain time step has influence on the error of the next time step. In the following sections, the systematic and random errors (with and without autocorrelation) are further explained.

Systematic error

Systematic errors are errors which are caused by incorrect use of the measuring instruments or by a defect in the instruments. A systematic error can either have a fixed absolute or relative deviation from the real value.

Random errors, with or without autocorrelation

A random error is caused by unknown unpredictable changes in the measurement. This can occur in the measurement equipment or can originate from external conditions. An assumption is made that these random errors are normally distributed with a certain mean and standard deviation (Jansen, 2007). In discharge measurement, random errors can have an autocorrelation in time. This means that the error depends on the error of the previous day. Random errors without autocorrelation have a mean of 0.

Table 5 shows the possible random errors in discharge measurement. The values of the spread of the random errors are found in the research of Jansen (2007). This research is based on measurement methods in de Meuse River at St Pieter, Borgharen and Maastricht. It is not clear what kind of measurement techniques are used in the Ourthe and Chiers rivers. That is why in this research several values for the spread in discharge for random errors will be used. The spread is defined as the maximum deviation in discharge of the measurements from the real values in 95% of the points, assuming a normal distribution of the errors.

Table 5: Overview possible random errors in discharge measurement Meuse River (Jansen, 2007)

Measurement error	***Present in period***	***Location of error***	***Spread in discharge of random error (percentage of points in 95% confidence interval)***
Level scale	1917-1956	Water level	10%
Level scale	1956-1975	Water level	5%
Registering level writer	1975-2000	Water level	5%
Digital level gauge	2000-2006	Water level	5%
ADM	2000-2006	Discharge	2.5%
Level stick and angustifolia posts	1917-1956	Cross section	2%
Measurement vessel and GPS	2000-2006	Cross section	1%

3.1.2 Errors from using Q-h relation

The Q-h relation can bring several errors in discharge determination. First of all, the discharge that corresponds with a certain water level is dependent on the properties of a high water event, like the shape of a certain wave and its gradient. Furthermore, the hysteresis phenomenon or an outdated Q-h relation can have influence on the discharge estimation. In the following sections these errors are discussed briefly.

Properties of a high water event

High water waves which last for a long period will have a different discharge compared to shorter waves with an equal water level. The differences are mainly noticeable at the peak of the wave. Steep and short high water waves will result in an underestimation of the discharge when applying the Q-h relation, while long and gradually increasing waves result in an overestimation of the discharge. This is due to the steepness of the wave. The steepness of the front of the wave has a relation with the celerity of the wave and therefore with the flow velocity. The changes in velocity that are a result of this, cause significant differences in water level in case of high water.

Hysteresis

The hysteresis effect describes the phenomenon in which the discharge in a hydrograph at a certain water level is higher if the water level is increasing, than in case the water level is decreasing, compared to the equilibrium situation (Boiten, 1986). This can be explained as follows: If a high water wave passes, diffusion occurs because a certain gradient in water level arises. This means that the peak of the wave, i.e. the flood maximum (the location with the highest discharge), gets a little smoother. The streaming velocity increases somewhat in front of the wave and thus the water depth decreases a little on that location. So at a certain fixed point along the river, the maximum discharge occurs at a certain moment, but the maximum water level occurs somewhat later (Jansen et al., 1979). The hysteresis effect is graphically shown in Figure 13. This figure shows that if a high water wave passes at a certain fixed point along the river, the maximum value for the discharge *Q* occurs earlier in time than the maximum value for the water level *h*. This means that at a certain water level *h* different values for the discharge *Q* can occur, depending on the presence of a rise or a fall in water level.

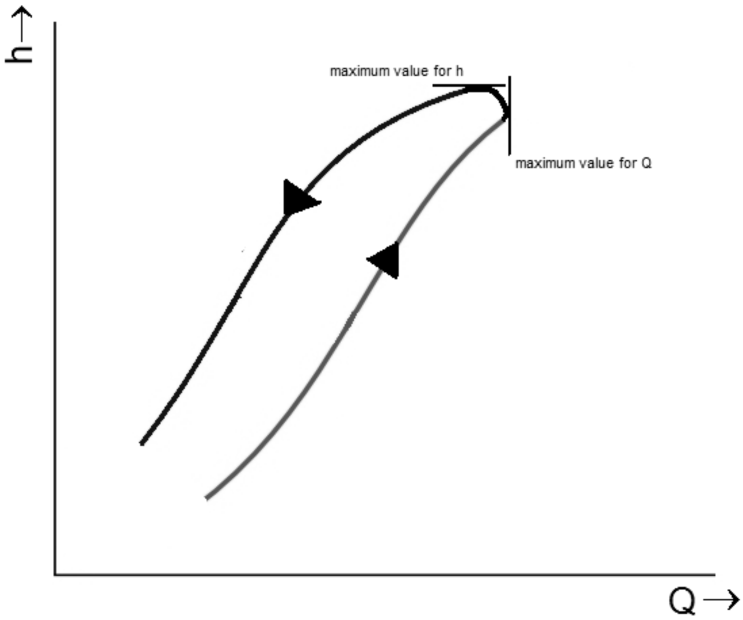


Figure 13: Graphical representation of the hysteresis effect

Outdated Q-h relation

If the Q-h relation loses its validity because it is outdated, different systematic errors can be present in the relation. The measured water level will not represent a correct value of the discharge. After some time a systematic error can be arisen. This systematic error can represent changes in the cross section of the river. For example, if sedimentation takes place at a certain location, the water level will be higher at a certain discharge compared to the water level before the sedimentation.

The different types of errors are summed up in Table 6.

Table 6: List of possible errors in discharge determination

Location of error	*Error*
Discharge measurement	(a) Systematic errors
	(b) Random errors without autocorrelation
	(c) Random errors with autocorrelation
Q-h relation	(d) Properties of a high water event
	(e) Hysteresis effect
	(f) Outdated Q-h relation

An addition to this overview, a combination of the two locations of the errors can be distinguished. When using the Q-h relation measurement errors in the water level *h* can be made. According to RIZA the spread of the random error in discharge determination when using the Q-h relation is between 5% and 10% (Jansen, 2007). This kind of error is not treated separately, because it will have a similar outcome as the random errors.

3.2 Sources of errors

In reality, combinations of the errors shown in Table 6 occur in discharge determination. To perform a realistic analysis, different errors are grouped by its source. These sources are developed because it is not clear what method is used for discharge determination in the Ourthe and Chiers basins. Error sources 1 and 2 are aimed at errors in discharge measurement itself, while sources 3 and 4 are cases in which the Q-h relation is used.

Measurement uncertainties:

- Error source 1: combination of a systematic error (a) and entirely random errors (b) in discharge, water level or cross section;
- Error source 2: Random errors with autocorrelation in time. These errors can be present in discharge, water level and cross section (c);

Q-h relation errors:

- Error source 3: use of Q-h relation with the uncertainties of the properties of a high water event (d) and the hysteresis effect (e);
- Error source 4: use of an outdated Q-h relation (f);

The original discharge time series are used as a reference. In reality it is not clear whether these data sets contain correct values, but since these are the only data that are available, the assumption is made that the present discharge data contain correct and realistic values. The errors on which each scenario is based result in a discharge data set that differs from the original discharge data. These adapted discharge time series are used in the calibration of the HBV model. The results of those calibrations can be found in chapter 4.

3.2.1 Error source 1: Combination of systematic and random errors

In this error source, a combination of systematic errors and random errors without autocorrelation is present. These errors can be a result of measurement errors in discharge, cross section or water level, but can also originate from the use of the Q-h relation, in which only the water level is measured.

Random errors without autocorrelation

The spread of the random errors is based on values in Table 5. The values in Table 5 are based on measurements in the Meuse River. Because it is not known what measurement techniques are used in the Ourthe and Chiers basin, several values for the spread are used. The random errors are integrated into the discharge data by using a random number generator. The distribution of the random chosen points is normal. The mean is the original discharge value and the standard deviation is half the spread of the error. For example, if the spread of the random error is 5%, i.e. 95% of the errors is smaller than 5%, then the standard deviation (σ) of the error is 2.5% of the original value. This is shown in Figure 14.

The used percentages for the spread are 5% and 10%.

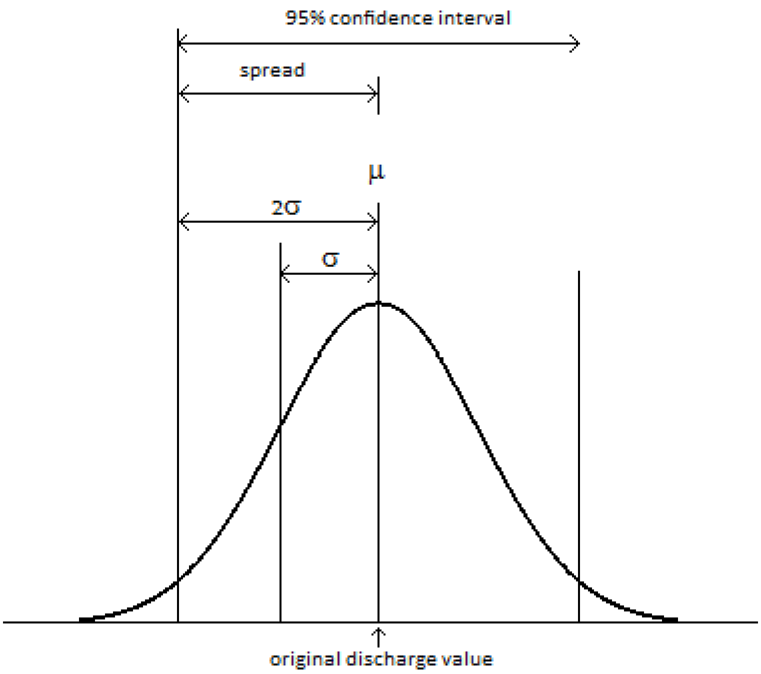


Figure 14: Random error

Systematic errors

Systematic errors are put into the model by adding a constant relative deviation into the discharge time series. This is done because an absolute deviation would have a large impact on the change in discharge if the discharge is small. In case a constant relative systematic error is used, the error will have a larger absolute deviation from the original value if the discharge is high. The used values for the systematic errors are -25%, -10%, -5%, 5%, 10%, and 25%.

The model results of this error source are shown in paragraph 4.1.

3.2.2 Error source 2: random errors with autocorrelation in time

Error source 2 contains random errors which have a certain autocorrelation in time. This means that the deviation from the real discharge value on a certain day has a positive influence on the deviation from the real discharge value on the next day.

A method for creating artificial discharge time series with random errors which have some autocorrelation is based on the method of De Kok and Booij (2009). They introduced a method for constructing artificial discharge time series with random errors including autocorrelation that is based on the following equations:

$$Q_g(t) = Q_0(t) + \varepsilon(t) \tag{6}$$

and

$$\varepsilon(t) = \delta(t)Q_0(t) + \alpha\varepsilon(t-1) \tag{7}$$

With Q_g = Discharge with deviation
Q_0 = Original discharge
ε = noise term
δ= randomly time-varying scaling factor
α= autocorrelation coefficient

The noise term $\varepsilon(t)$ forms the random error. $\delta(t)$ originally is a time-varying scaling factor uniformly distributed in the interval [-Δ,Δ]. Δ can have values between 0 and 1. In this research a fixed value for Δ is used: Δ = 0.05. This fixed value is chosen because it is expected that in this case the maximum spread of the random errors will be around 10%. Because the random error $\varepsilon(t)$ is auto correlated with the previous random error (equation (7)), the noise term will not be uniformly distributed around the mean, like δ. The values of ε near the mean will occur more frequent than larger deviations. The autocorrelation coefficient α is unknown for the used basins, so several values are assumed. The used values for α are 0.5, 0.7, 0.8, 0.9 and 0.95. A value of 0 for α would mean that there would be no autocorrelation.

3.2.3 Error source 3: wrong interpretation of the Q-h relation, take into account the properties of a high water event and hysteresis

In this case the Q-h relation is used. There are no uncertainties in the relation itself or the determination of the water level. However, the properties of a high water event and the hysteresis effect can bring uncertainties in the discharge determination.

Hysteresis

The hysteresis effect describes the phenomenon in which the discharge in a hydrograph at a certain water level is higher if the water level is increasing, than in case the water level is decreasing, compared to the equilibrium situation. In the equilibrium situation with uniform flow conditions, the following relation is valid:

$$Q_s = A_s C\sqrt{Ri_b} \tag{8}$$

In which Q_s= discharge during equilibrium situation [m^3/s]
C = Chézy coefficient [-]

R = hydraulic radius [m]
A_s = surface area of cross section [m^2]
i_b = bottom slope [-]

Jansen et al. (1979) presented an equation (9) which determines the difference of increasing and decreasing water levels, compared to the equilibrium situation in equation (8).

$$Q - Q_s = \frac{Q_s}{2 \cdot i_b \cdot c} \frac{dh}{dt} \tag{9}$$

In which Q = discharge during wave passage [m^3/s]
c = celerity of the wave propagation [m^3/s]

For Q_s the original discharge time series is used. An assumption has to be made regarding the parameter c. Jansen et al. (1979) stated that in case of a rectangular cross section, equation (10) can be used.

$$c = \frac{3}{2} u \tag{10}$$

In which: u = flow velocity
c = wave propagation velocity

The parameters u and c are unknown. For the flow velocity u, five different values are assumed. This results in different values of the wave propagation velocity c. These five values have five different discharge time series as a result, each with a different behavior of the hysteresis effect. The velocities of the flow and the wave propagation are shown in Table 7.

Table 7: Different values of parameters in hysteresis effect

Parameter	***Representation***	***Unit***	***Values***
u	Flow velocity	[m/s]	0.5, 0.75, 1.0, 1.25, 1.5
c	Wave propagation velocity	[m/s]	0.75, 1.125, 1.5, 1.875, 2.25

For i_b, the average bottom slope of the river will be used. In Table 8 the bottom slopes of the Chiers and Ourthe river are given (Booij and Krol, 2010).

Table 8: Bottom slopes of the Ourthe and Chiers river

River	***Bottom slope [-]***
Ourthe	$3.7 *10^{-3}$
Chiers	$1.0 *10^{-3}$

After that, $\frac{dh}{dt}$ has to be determined for every time step. Région Wallonne (2009) has published monthly values for water level and discharge in the period of 1988 to 2007 of the Ourthe river. These values are used to determine a Q-h relation of the Ourthe river. An assumption is made that this particular Q-h relation is used in the calibration period. The Q-h relation for the Ourthe river is shown in Figure 15. In this graph, a trend line is added. The equation that belongs to this trend line, is used to determine the water level with a certain discharge Q_s. Because five values of c are used, five different discharge series are constructed and used in this case.

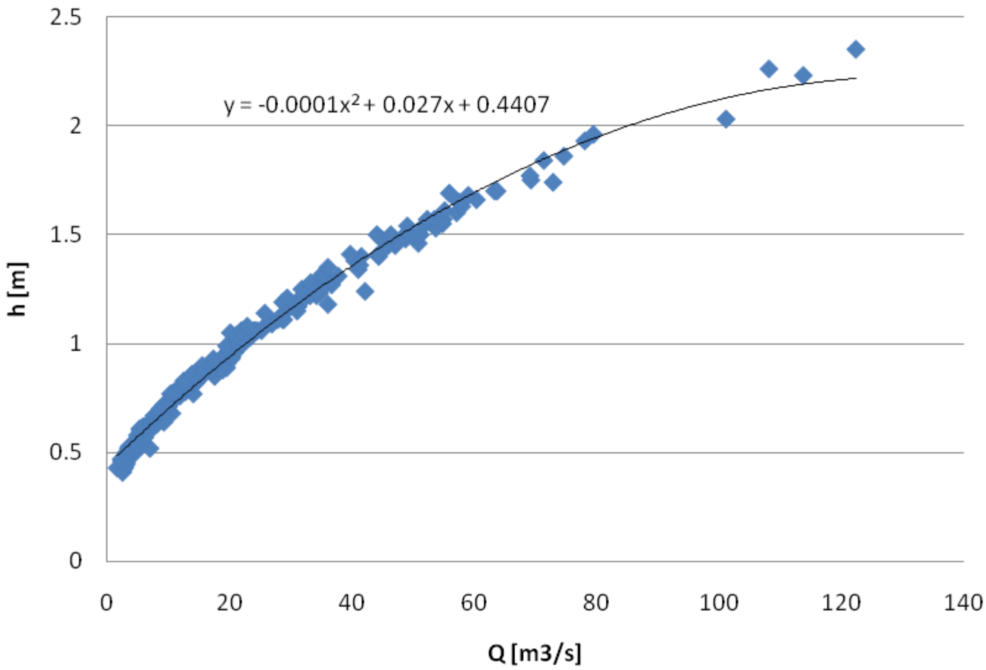


Figure 15: Q-h relation Ourthe

For the Chiers basin, no water level data are known, so for the Chiers basin it is not possible to set up such a Q-h relation. In equation (9), the value of the water level is not important. Only the change in water level is of importance. Because the Ourthe and Chiers river have discharges of comparable magnitude, an assumption is made that a certain increase in discharge in the Chiers river, has a similar increase of water level compared to the Ourthe river with the same increase in discharge. In other words, the Q-h relation for the Ourthe is also used for the Chiers river.

Properties of a high water event

To integrate this error, the discharge series are adapted. Assumptions are made which determine whether there actually is a wave and whether this wave is steep and short or long and gradually increasing. Also the overestimation and underestimation has to be defined. In Table 9, the used characteristics for the waves are shown. Combinations of the extreme values are used in the construction of the artificial discharge series, resulting in eight different discharge data series.

An overestimation is inserted if a gradually increasing wave is present, so if the discharge is above the *threshold of a high water level* and if the increase or decrease of the water level is smaller than the *maximum gradient of the gradually increasing wave*. An underestimation is added if there is a high peak with a large increase, so if the discharge is higher than the *threshold of the peak of the wave*, and the gradient is larger than the *minimal gradient steep wave*.

Table 9: Key characteristics of flood waves

Key values	***Ourthe***	***Chiers***
*Threshold peak of the wave [m^3/s]**	200	150
*Threshold high water level [m^3/s]***	100	75
*Minimum gradient steep wave [$m^3/s/day$]****	15 – 20	15 – 20
*Maximum gradient gradually increasing wave [$m^3/s/day$]****	5 – 10	5 – 10
*Underestimation****	10% - 20%	10% - 20%
*Overestimation ****	10% - 20%	10% - 20%

* These values are chosen because these discharges represent a peak wave which occurs about one time a year.

** Significant differences only occur at high water levels. An assumption is made that high water is defined as half the discharge of a peak which occurs once a year on average.

*** Assumptions are made regarding the values of the maximum and minimum gradient and the over- and underestimation because there is no information about these parameters.

3.2.4 Error source 4: outdated Q-h relation

In this type of error, the used Q-h relation is outdated and will contain certain errors. The determination of the water level *h* is performed well, but the corresponding discharge *Q* is not realistic due to changes in the conditions of the river, such as changes in the river bed. To simulate this kind of errors a systematic error can be introduced. This systematic error starts to increase just after a revision of the Q-h relation. The errors increase to a certain maximum systematic error just before a new Q-h relation is set up. The systematic errors are assumed to be absolute deviations from the original values, because it is assumed that the expiration of the Q-h relation is caused by changes in the cross section. In that case, it is more logical to insert an absolute value for the systematic error than to insert a relative error. Other assumptions are that the systematic error is positive and that the Q-h relation is revised every 5 years. This assumption is an estimation of the period between two revisions and is based on the research of Jansen (2007) in which the period of five years has proved to be a current practice in the Meuse River.

The maximum systematic errors that are used, are 1, 5, 10 and 15 m^3/s. In Figure 16 a maximum systematic error of 5 m^3/s due to an outdated Q-h relation is shown. This figure indicates that the systematic error is zero every time the Q-h relation is revised and increases until the next revision which takes place 5 years later. The calibration period is 15 years so three revisions take place in this period. The original time series start on 1 January 1968 with a systematic error of 0. After that, every five years a revision takes place. The calibration period starts on January 1st 1984, one year after the Q-h relation revision on January 1st 1983.

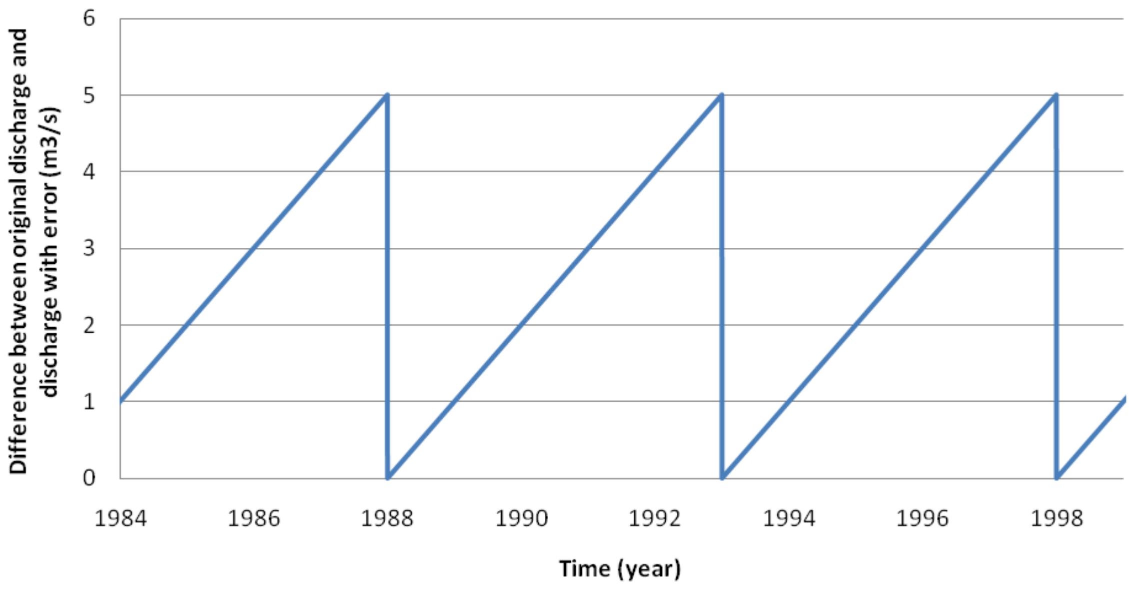


Figure 16: Systematic error due to an outdated Q-h relation

3.3 Realistic scenario

A part from the different sources of errors, a scenario will be developed which contains errors in the discharge that are realistic in the two chosen sub basins. In these two sub basins it is expected that the Q-h relation is used for discharge determination. Which of the error sources are used in this scenario, will be decided after the model results are presented, in paragraph 4.6.

3.4 Classification of quality artificial discharge time series

To assess the quality of the constructed discharge time series which contain one or more errors, two quality functions are used. These functions can be used to compare the quality of the discharge series with the objective function after calibration. The first one is aimed at the quality of the shape of the hydrograph and is comparable with the Nash-Sutcliffe coefficient (Nash and Sutcliffe, 1970) which assesses the quality of the model. This function is called Quality Of Discharge (*QOD)* and is shown in equation (11). With a perfect match of original and constructed data, this function will have a value of 1.

$$QOD = 1 - \frac{\sum_{i=1}^{N}\left(Q_{art}(t_i) - Q_{org}(t_i)\right)^2}{\sum_{i=1}^{N}\left(Q_{org}(t_i) - \overline{Q_{org}}\right)^2} \qquad \textbf{(11)}$$

In which Q_{art} = the artificially constructed discharge time series
Q_{org} = the original discharge time series

The second function to assess the quality of the constructed discharge data series, is called *BALANCE* (equation(12)).

$$BALANCE = \frac{\sum_{i=1}^{N} \left(Q_{art}(t_i) - Q_{org}(t_i) \right)}{\sum_{i=1}^{N} \left(Q_{org}(t_i) \right)} \qquad \textbf{(12)}$$

It represents the water balance between the original discharge data and the new, constructed data series. If in the entire calibration period the same amount of water is discharged through the system, the *BALANCE* function will have its optimal value of 0. The *QOD* and *BALANCE* functions are based on similar quality functions which are used in the research of Andréassian et al. (2001). These functions, named *GORE* and *BALANCE* were used to assess the quality of rainfall time series.

4 Model results

In this chapter the influence of the errors in the different sources on the model performance is discussed, as well as the influence of these error sources on the parameter estimation. For every error source, the values of the objective function *y* are related to the discharge quality functions *BALANCE* and *QOD* and presented graphically. Furthermore, the relation between the discharge quality functions and the model performance, expressed in the objective function *y* is investigated. Also the differences between the Ourthe and Chiers basin in model performances are observed.

A part from the influence of the errors on model performance, the influence of the error sources on the model parameters for both basins is analyzed. After that, the influence on the parameter values is interpreted taking into account the physical meaning of the parameter. For the calibrations, the same parameter ranges and optimization properties are used as in the original optimizations, in section 2.4.3.

4.1 Error source 1: Combination of systematic and random errors

In this source of error, combinations of systematic and random errors are inserted into the discharge data. These two kinds of errors are discussed in section 3.2.1. The values for the spread of the random errors are 5% and 10%, and the systematic errors have values of -25%, -10%, -5%, 5%, 10%, 25% of the original value. The random and systematic errors are combined. These combinations lead to 12 different discharge data series and consequently 12 calibrations.

4.1.1 Influence on model performance

In Figure 17 and Figure 18 the model performance with the discharge data including error source 1 is shown. The different symbols represent different systematic errors with a 5% spread of the random error. The triangles stand for a systematic error of 5%, the circle for 10% and the square for 25%. If the spread of the random error is equal to 10%, a dot is shown. The '*' sign indicates the values of the objective function and the discharge quality functions in the original situation. In that case, the *QOD* and *BALANCE* functions have the optimal value. This is because the discharge series are not altered yet and thus are identical to the original situation.

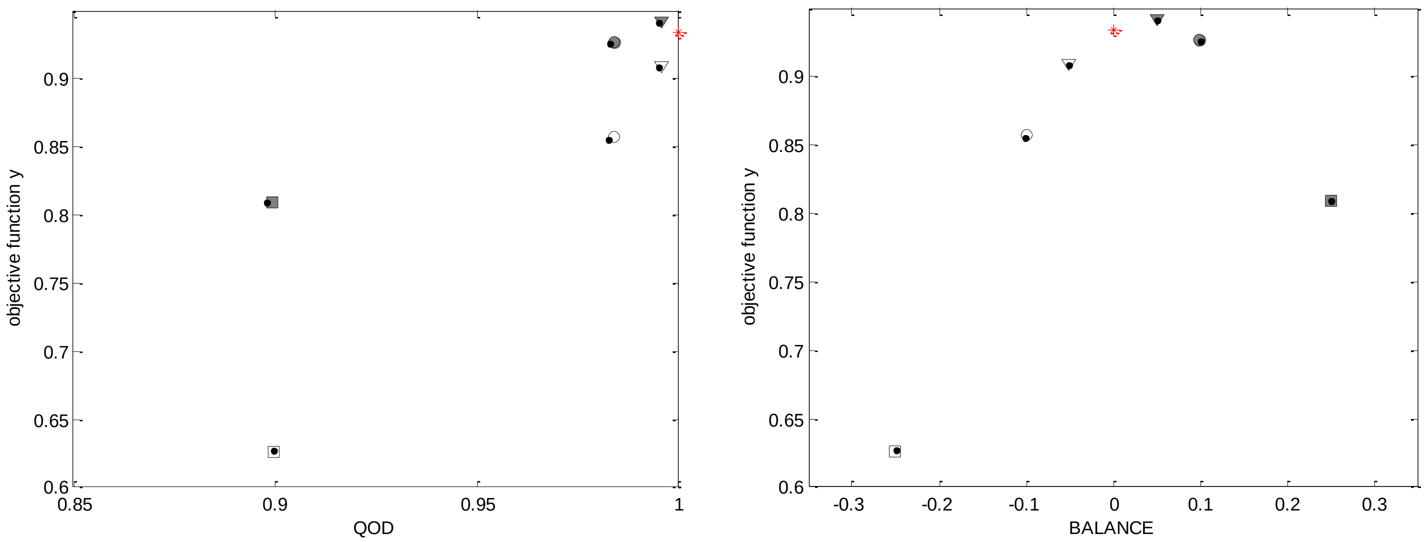


Figure 17: Influence of error source 1 on model performance Ourthe basin

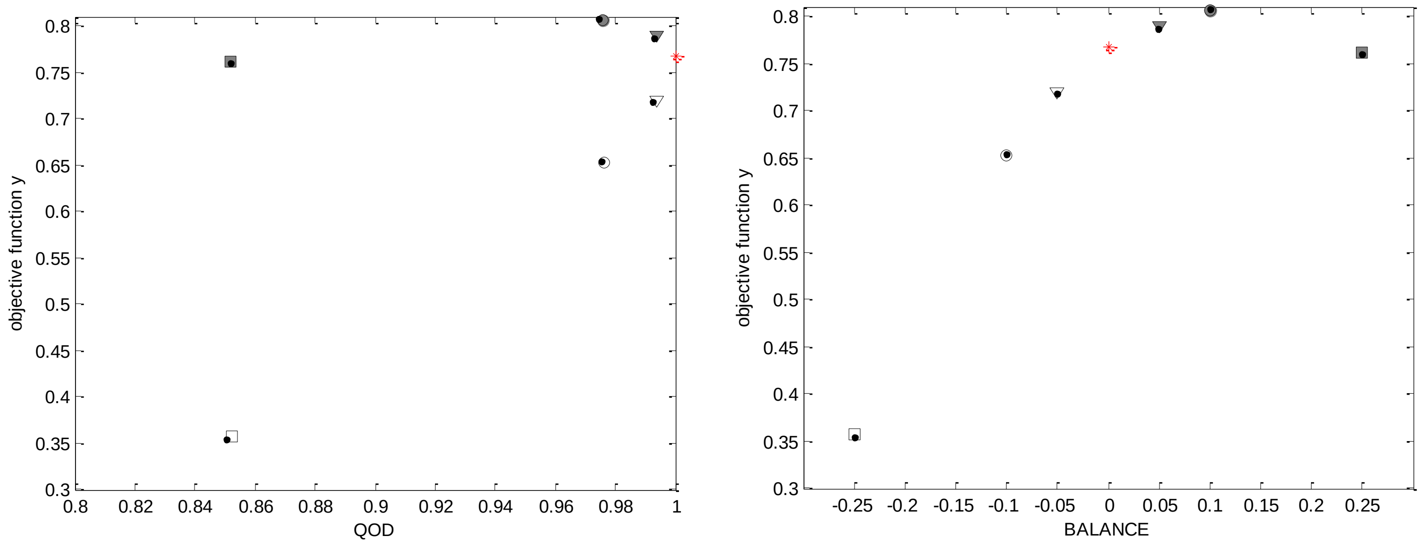


Figure 18: Influence of error source 1 on model performance Chiers basin

The figures show that the influence of the random error is small, even if the spread is quite large, like 10%. This can be concluded because the dots (10% spread) do not differ much from the '∇', 'O' and '□' signs (all 5% spread). Furthermore, the expectation would be that with an increasing positive or negative systematic error, the value of the objective function *y* would decrease.

Figure 17 and Figure 18 show that if the objective function is expressed against *BALANCE*, a parabola arises. However, the peak of this parabola is not situated at *BALANCE*=0, but the highest value for the objective function in both basins is reached for a certain positive systematic error. In the Ourthe basin the peak is found at around +5 m^3/s, and in the Chiers basin this optimum appears to be between +10 m^3/s and +25 m^3/s. This means that the discharge series containing a certain positive systematic error can better be approached by the HBV model than the original discharge series.

4.1.2 Influence on model parameters

Figure 19 and Figure 20 show in what way the parameters in the two sub basins are affected by the errors of error source 1. The vertical axis represents the parameter value. The ranges of the axes are equal to the used parameter ranges. The horizontal axis represents the objective function *y*. The open markers indicate a negative systematic error, while the filled markers indicate that the systematic error is positive. The triangle stands for a systematic error of 5%, the circle for 10% and the square for 25%. Because the magnitude of the random error does not have much effect on the model performance or parameters, these symbols are used for both the random errors of 5% and 10%. The '*' sign shows the parameter value and the value of the objective function *y* in the original situation, i.e. the situation without errors in the discharge series.

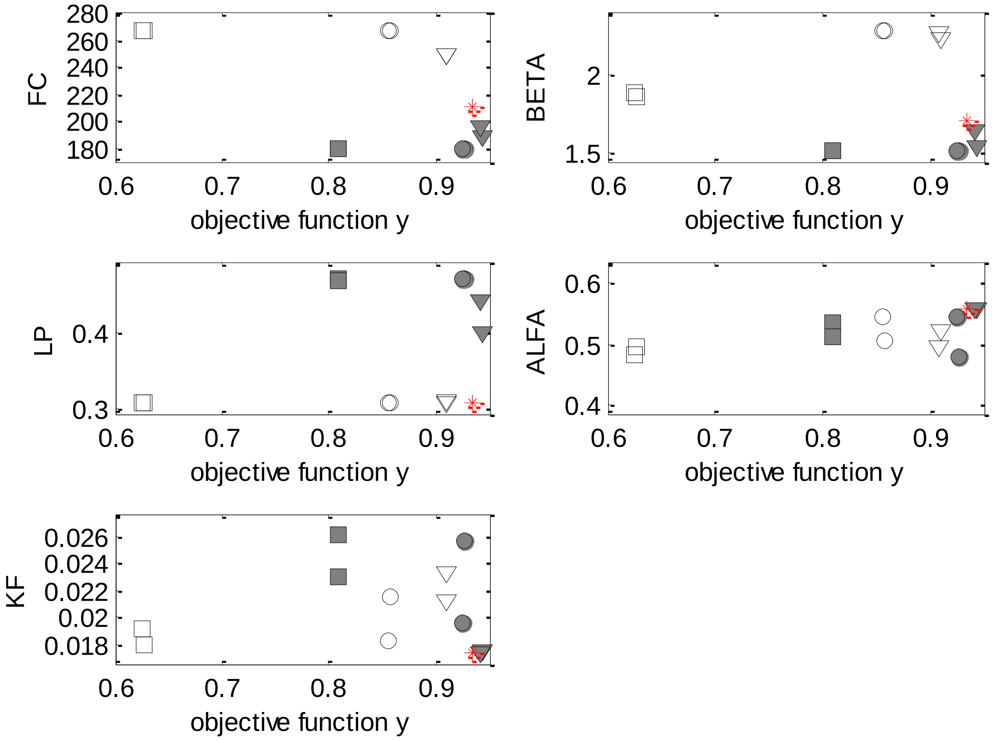


Figure 19: Influence of error source 1 on model parameters Ourthe basin

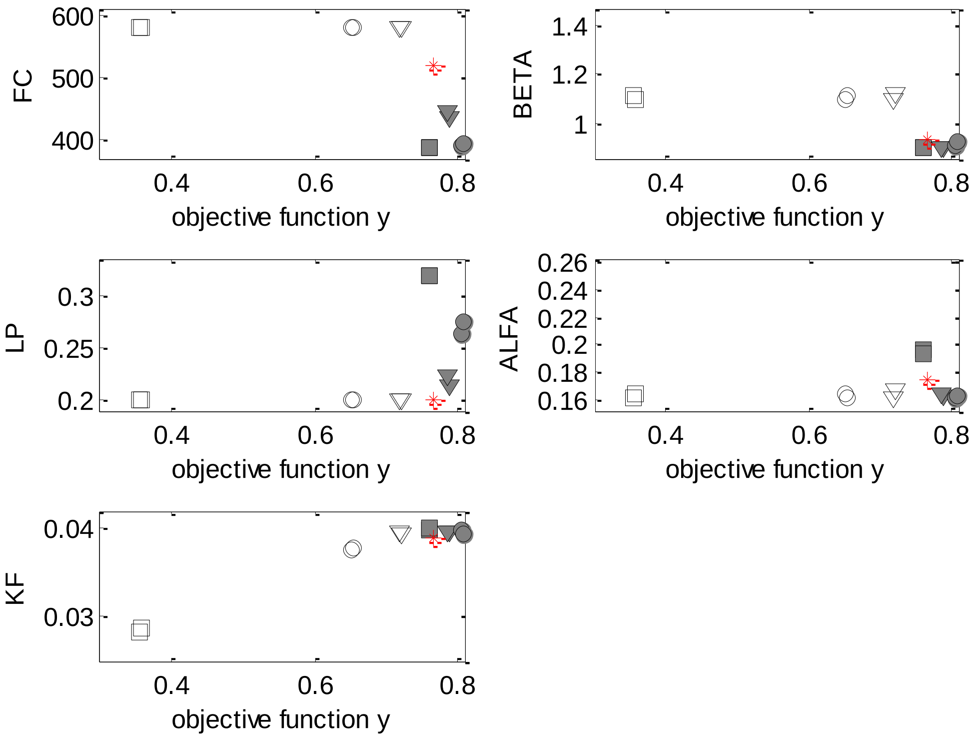


Figure 20: Influence of error source 1 on model parameters Chiers basin

The first remarkable fact is that it is clear that systematic errors have a big influence on some of the model parameters in both basins. In both basins this is clear for the parameters *FC*, *BETA* and *LP*. For the parameters *ALFA* and *KF* no patterns are visible. Only in the Chiers basin parameter *KF* shows a deviation if *BALANCE* is very small (-25%). The parameters *FC*, *BETA* and *LP* show a certain pattern in both basins. The values of *FC* and *BETA* increase if the systematic error is negative, while the values decrease with a positive systematic error. *LP* shows similar behavior, but then the other way around: positive systematic errors result in an increase in parameter value, while a negative systematic error makes *LP* decrease. These phenomena are found in both the Ourthe and Chiers basins.

The observed patterns can be explained by the physical meanings of the parameters. *FC* decreases with a positive systematic error, because the model has to generate more runoff than in the original situation during the entire calibration period. The HBV model simulates this by decreasing the capacity of the soil box. In that case, the recharge to the upper response box increases, so the storage in this response box increases. This is the reason that there is more quick runoff over the entire calibration period. If the systematic error is negative, an opposite behavior of *FC* is visible.

In both basins *BETA* decreases with a positive systematic error and vice versa. This can be explained by the fact that with a decreasing *BETA*, the contribution from precipitation to recharge is larger. That is why the upper response box will be filled faster and therefore the quick runoff increases.

The value of *LP* increases if the systematic error is positive. If the value of *LP* increases, the fraction of the field capacity above which the actual evapotranspiration is equal to the potential evapotranspiration increases. This means that if *LP* increases, there is less evapotranspiration during the calibration period. In that case, there will be more soil moisture in the soil box. This means that the recharge to the upper response box increases and thereby the (quick) response as well.

Parameters *ALFA* and *KF*, which both have influence on the quick response, have no marked pattern in which the values change by the systematic errors in error source 1. Only if the systematic error is very large (-25%) in the Chiers basin, *KF* decreases significantly. This is because the quick response drastically decreases in that case. The parameters *FC*, *LP* and *ALFA* have reached the border of their parameter range in this case. The model however still tries to decrease the total runoff to approach the discharge data. Because only *KF* and *BETA* have not reached their range border yet, these two parameters have to compensate for the low discharge value. This is done by decreasing the value of *KF*.

However *ALFA* and *KF* do not seem to be influenced by the systematic errors, the random errors appear to have some effects on these parameters. This can be concluded because the pairs of the same markers are not situated very close to each other. This is mainly visible in the Ourthe basin. The random errors have less effect on the model parameters in the Chiers basin.

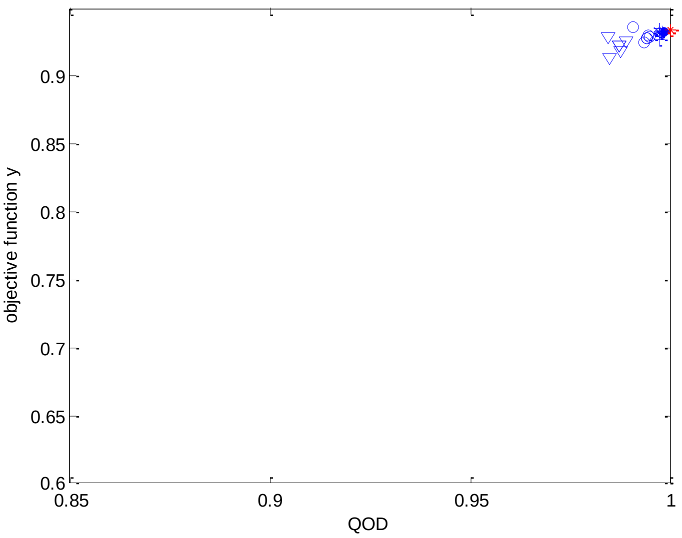
4.2 Error source 2: Random errors with autocorrelation in time

This error source contains random errors with autocorrelation in time. In paragraph 3.2.2 an explanation of this type of error is presented. Equations (6) and (7) are used to adapt the discharge data. For the autocorrelation coefficient α the values 0.5, 0.7, 0.8, 0.9 and 0.95 are used. The fixed value Δ is set to 0.05, because it is expected that the spread of the random error is around 10%.

4.2.1 Influence on model performance

In Figure 21 and Figure 22 the influence of error source 2 on the model performance is shown. The different symbols represent the different values of α:

.: $\alpha = 0.5$
+: $\alpha = 0.7$
x: $\alpha = 0.8$
o: $\alpha = 0.9$
∇: $\alpha = 0.95$



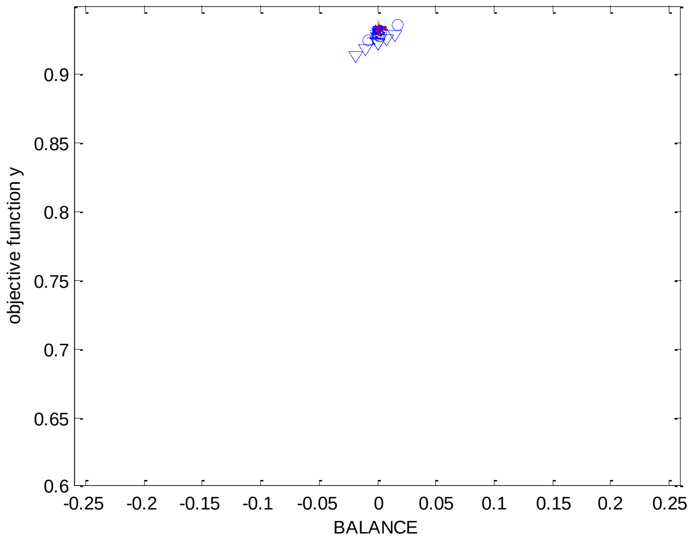
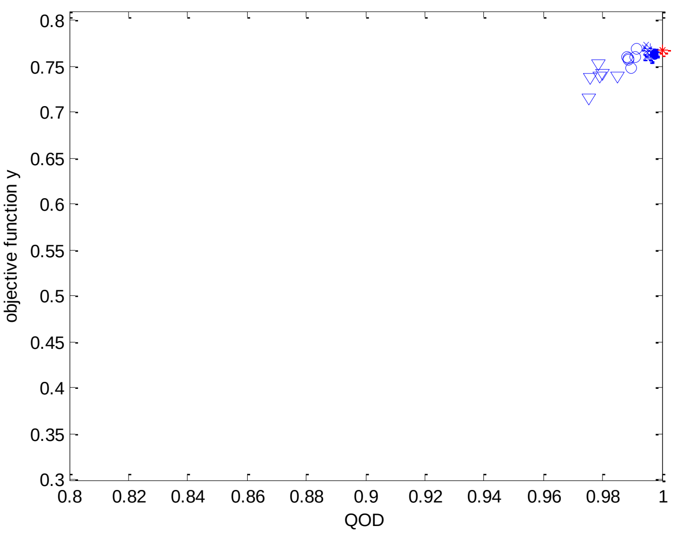


Figure 21: Influence of error source 2 on model performance Ourthe basin



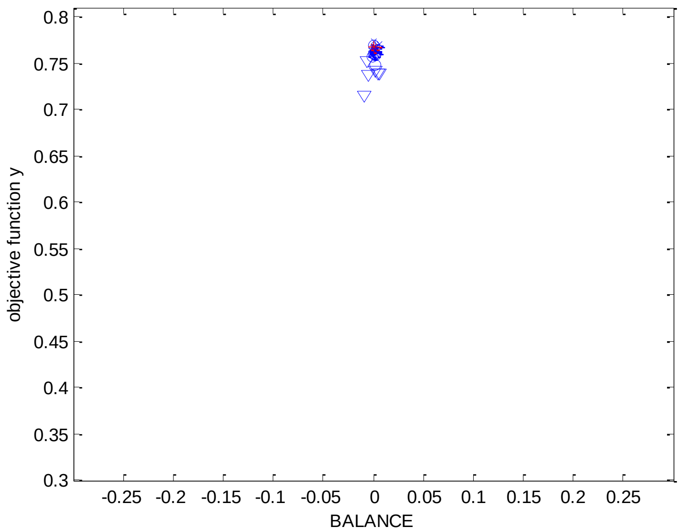


Figure 22: Influence of error source 2 on model performance Chiers basin

The two figures show that the different values of α result in some small deviations from the original values for *BALANCE*/*QOD* and *y* in both the Ourthe and Chiers basin. The larger α, the larger the deviation from the original discharge data and the original functions. With an increasing α, the objective function decreases and the discharge quality functions vary from the optimal value. The spread of the random errors that are a result of the changes in discharge data in the Ourthe and Chiers basins can be found in Table 10. This table shows that the calibrations with the largest values for α (0.9 and 0.95) result in large values for the spread, which are not realistic (see Table 5 in section

3.1.1). In general it can be concluded that this error source has a small influence on the model performance.

Table 10: Spread due to different values for *α* in Ourthe and Chiers basins and the corresponding values for *y*

Alpha value	*Ourthe*		*Chiers*	
	Average spread	*Average value objective function y*	*Average spread*	*Average value objective function y*
0.5	7%	0.933	7%	0.764
0.7	8%	0.932	8%	0.762
0.8	10%	0.932	10%	0.766
0.9	15%	0.929	14%	0.759
0.95	28%	0.922	21%	0.737

4.2.2 Influence on model parameters

The influence of the error source 2 on the model parameters in the Ourthe and Chiers basin are shown in the graphs in respectively Figure 23 and Figure 24.

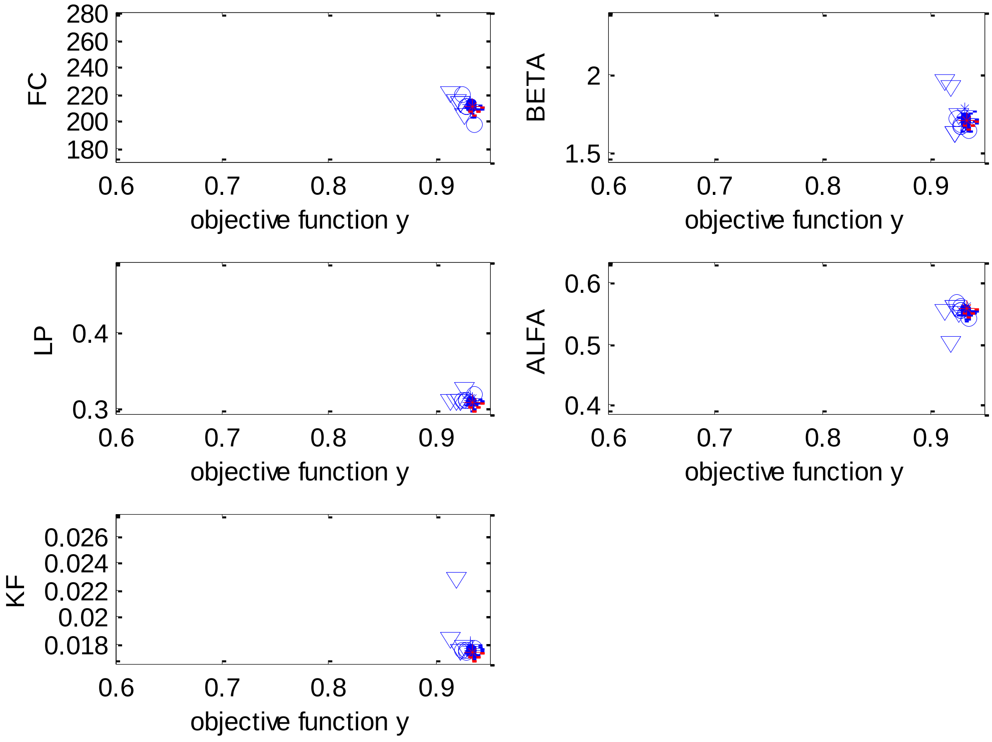


Figure 23: Influence of error source 2 on model parameters Ourthe basin

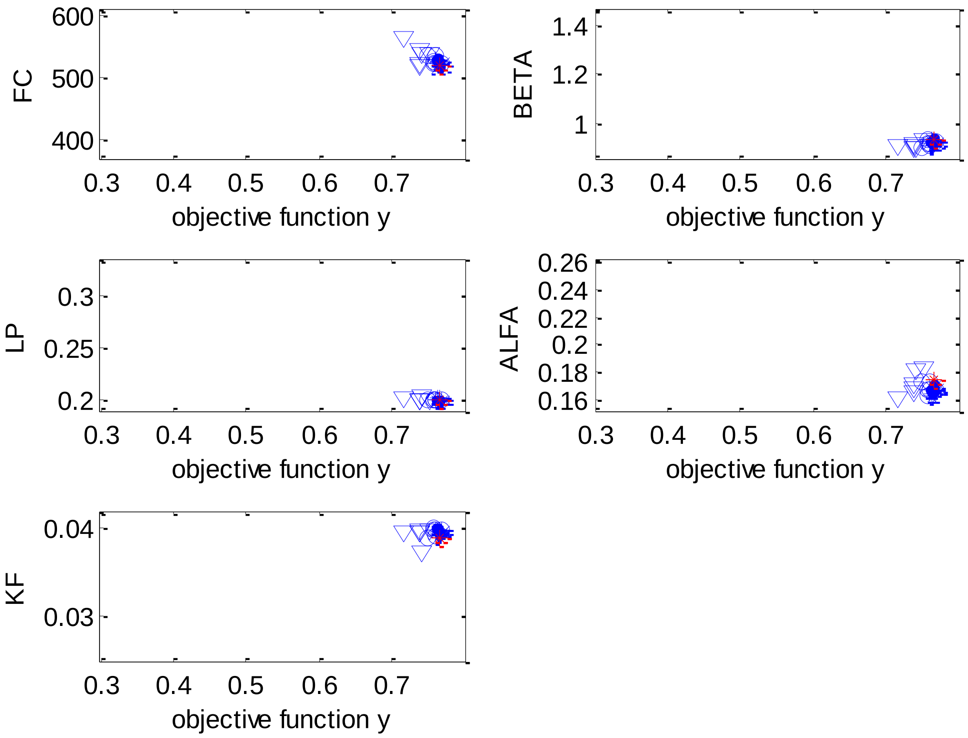


Figure 24: Influence of error source 2 on model parameters Chiers basin

The above figures show that the random errors do not have much influence on the model parameters in both basins. Only if the autocorrelation is very high (‘∇’: α=0.95), there are some small changes in parameter values. In the Ourthe basin, some changes occur for the parameters *BETA*, *ALFA* and *KF*. These changes do not have much effect on the value of the objective function. In the Chiers basin the changes in parameter values are minimal and do not influence the objective function either, even if the autocorrelation is very high (α=0.95). It is concluded that this error source has a small influence on the model parameters in the two basins.

4.3 Error source 3: Using the Q-h relation; hysteresis and properties high water event

Error source 3 contains a combination of two errors that are a result of using the Q-h relation. These errors are considered because they influence the determination of the discharge in a way that a certain water level *h* not always represents the same discharge.

Hysteresis effect

First of all, the hysteresis phenomenon is inserted into the discharge data. In paragraph 3.2.3 the method of introducing these errors into the discharge data is described. This is done by using equation (9) and (10). Therefore an assumption for the flow velocity has to be made. The flow velocity is assumed to be 1.5 m/s in both the Ourthe and the Chiers river. This value for the velocity is quite high and this value is chosen because in paragraph 3.2.3 it turned out that only the highest flow velocities have some influence on the model calibration.

Properties of a high water event

The properties of a high water event can also cause errors into discharge determination when using the Q-h relation. To simulate this kind of error, some key characteristics are assumed. The key characteristics used for this error source are shown in

Table 9 in paragraph 3.2.3. Combinations of the extreme situations lead to eight calibrations for this error source. Figure 25 and Figure 26 show the influence of error source 3 on the model performance.

4.3.1 Influence on model performance

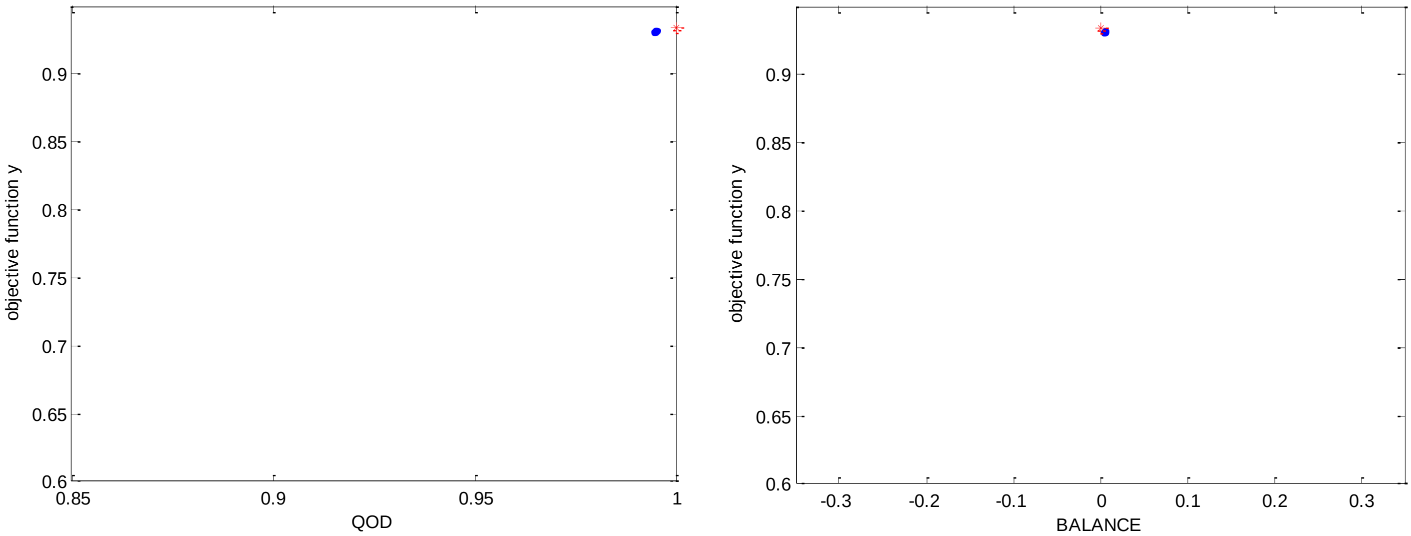


Figure 25: Influence of error source 3 on model performance Ourthe basin

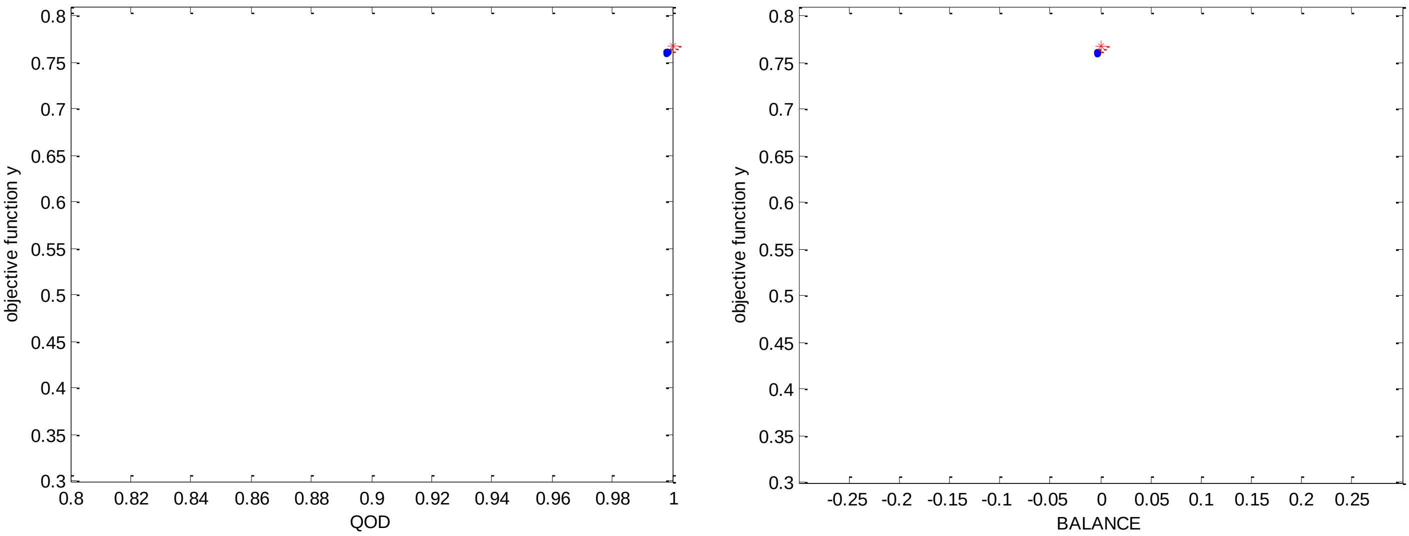


Figure 26: Influence of error source 3 on model performance Chiers basin

In the Ourthe basin error source 3 shows values of the objective function *y* of around 0.932, while the original discharge data of the Ourthe results in an objective function *y* of 0.934. The discharge quality functions *QOD* and *BALANCE* show some small effects. Error source 3 seems to have some influence on *QOD* and no significant effect on *BALANCE*. This is due to the fact that the errors change the shape of the hydrograph. This change in shape affects the quality function *QOD*. The errors that are a result of this source however do not cause much effect on the *BALANCE* function because the errors counterbalance each other as they contain both positive and negative deviations from the original series. In general, the conclusion can be drawn that the changes in both the *BALANCE* and *QOD* functions do not cause large deviations from the objective function *y*.

The discharge quality functions in the Chiers basin show a similar behavior. Both the *QOD* and *BALANCE* do not have large deviations from the optimal values. *BALANCE* again does not differ much from the original situation because the errors caused by error source 3 have positive and negative deviations and these deviations compensate for each other.

4.3.2 Influence on model parameters

In Figure 27 and Figure 28 the influence of error source 3 on model performance is shown for respectively the Ourthe and the Chiers basins.

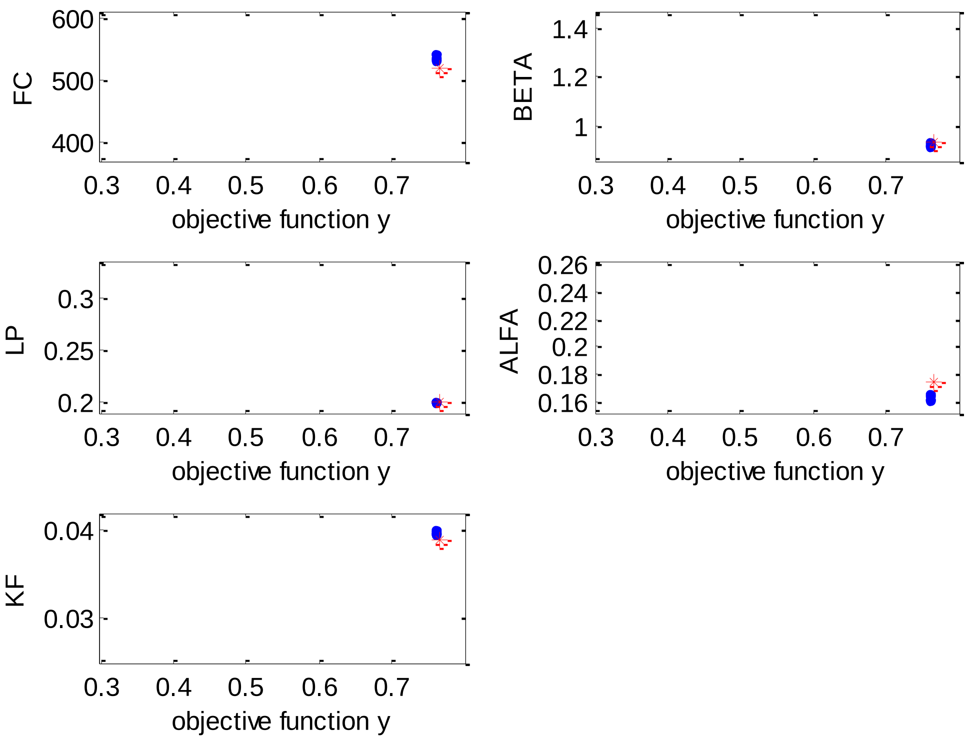


Figure 27: Influence of error source 3 on model parameters Ourthe basin

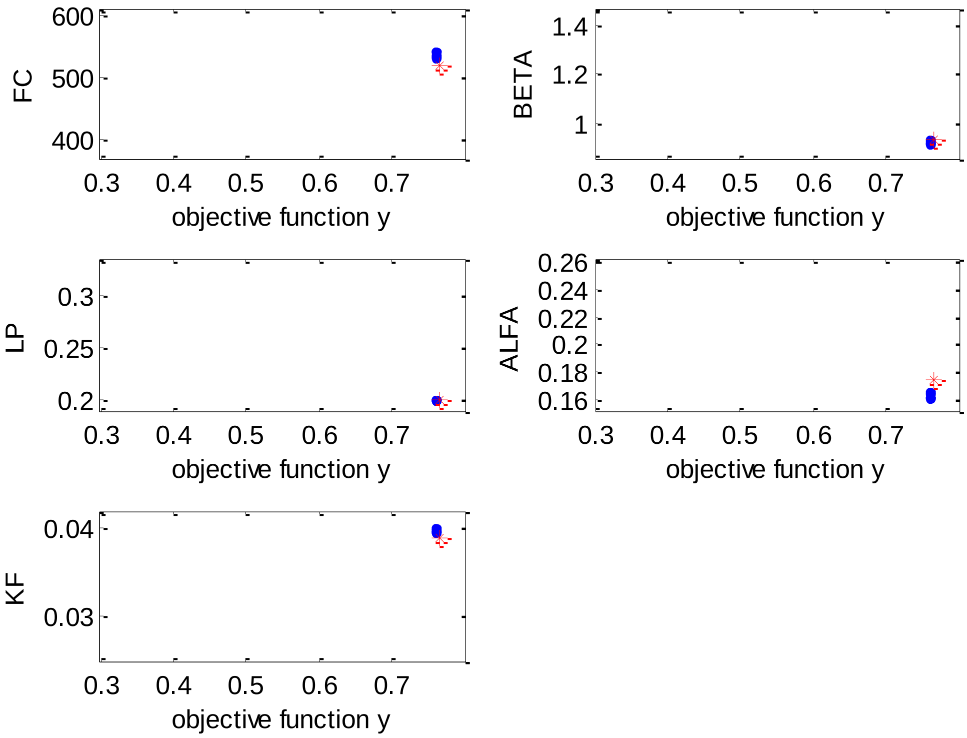


Figure 28: Influence of error source 3 on model parameters Chiers basin

The figures show that the deviations in discharge do not seem to have much influence on the parameter values in both basins. Only the parameters *FC* (small increase) and *ALFA* (small decrease) are a little affected by the errors. This means that the maximum soil storage increases a little and the quick response has a more linear behavior related to the storage in the upper response box. The reasons for these changes are unknown and difficult to clarify.

4.4 Error source 4: Using Q-h relation; Outdated Q-h relation

In paragraph 3.2.4 the implementation of this error into the original discharge data is explained. The used values for the maximum absolute systematic error are -15, -10, -5, -1, 1, 5, 10 and 15 m^3/s. In the figures in the following sections, the filled markers represent a positive systematic error and the open markers indicate that there is a negative systematic error. The following symbols are used as markers:

Triangle (∇): ± 1 m^3/s
Circle (O): ± 5 m^3/s
Square (□): ± 10 m^3/s
Diamond(◇): ± 15 m^3/s

4.4.1 Influence on model performance

The influence of error source 4 on model performance is shown in the following figures.
Figure 29 shows the influence on model performance for the Ourthe basin, while in Figure 30 the results for the Chiers basin are shown.

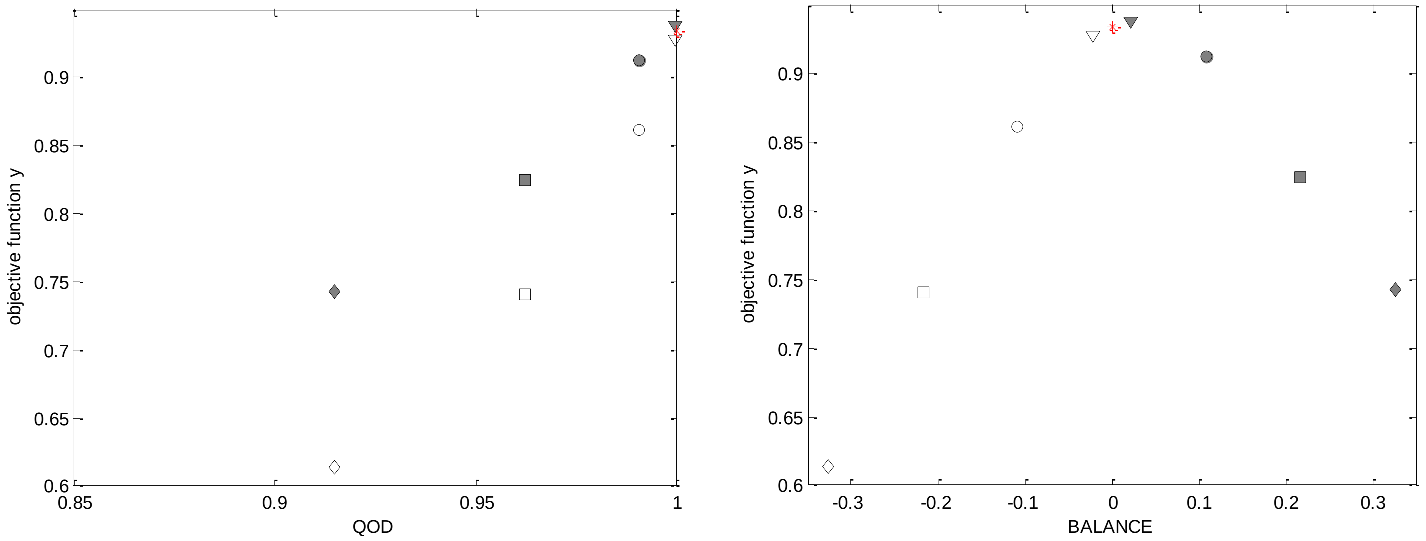


Figure 29: Influence of error source 4 on model performance Ourthe basin

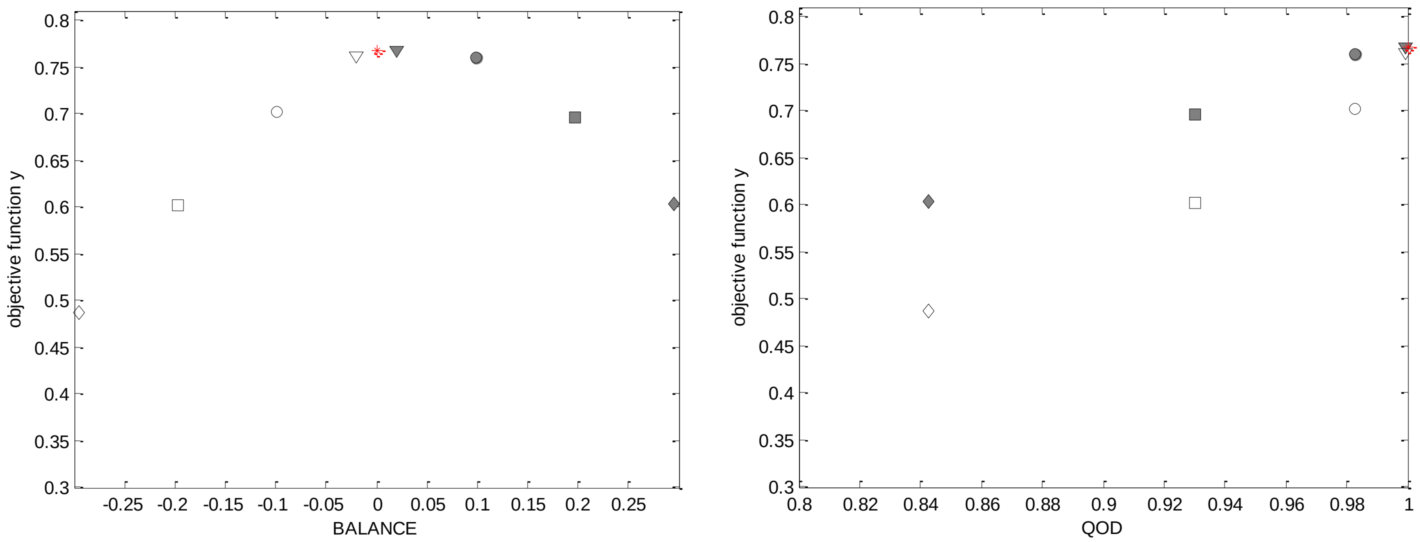


Figure 30: Influence of error source 4 on model performance Chiers basin

Error source 4 shows similar effects as error source 1. This is due to the fact that both error sources contain a type of systematic error. In general it shows that if the systematic error increases, the objective function *y* decreases. However, like in error source the highest value for the objective function is not found at *QOD*=1 and *BALANCE*=0, but if a small positive systematic error is present, both in the Ourthe and Chiers basin.

4.4.2 Influence on model parameters

Figure 31 and Figure 32 show the influence of error source 4 on the parameters in Ourthe and the Chiers basins.

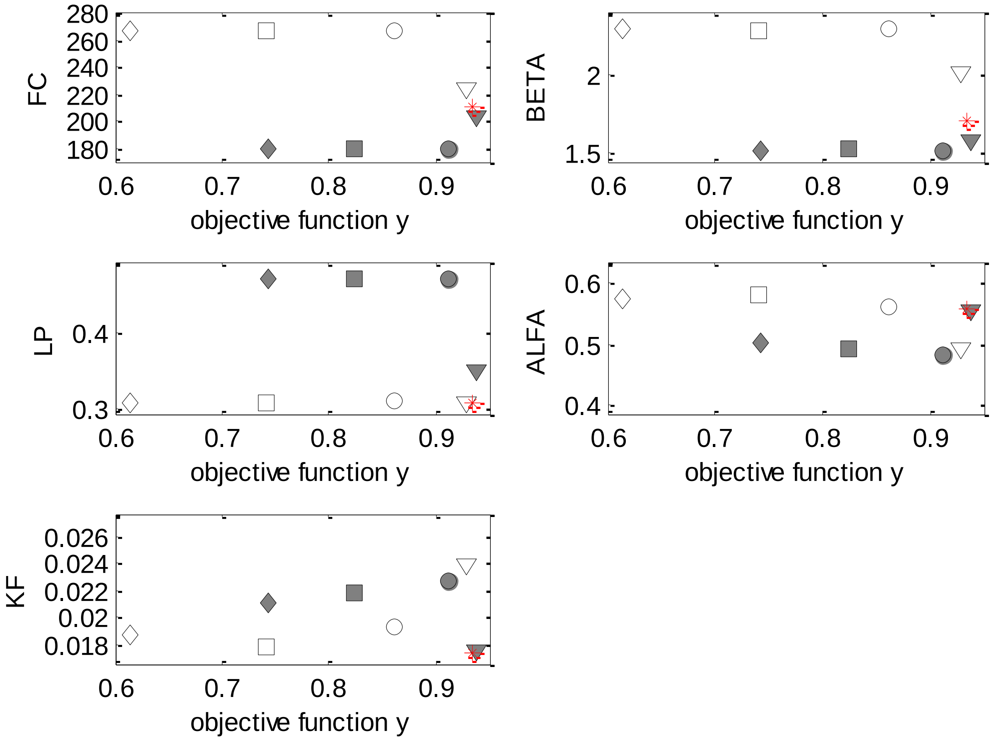


Figure 31: Influence of error source 4 on model parameters Ourthe basin

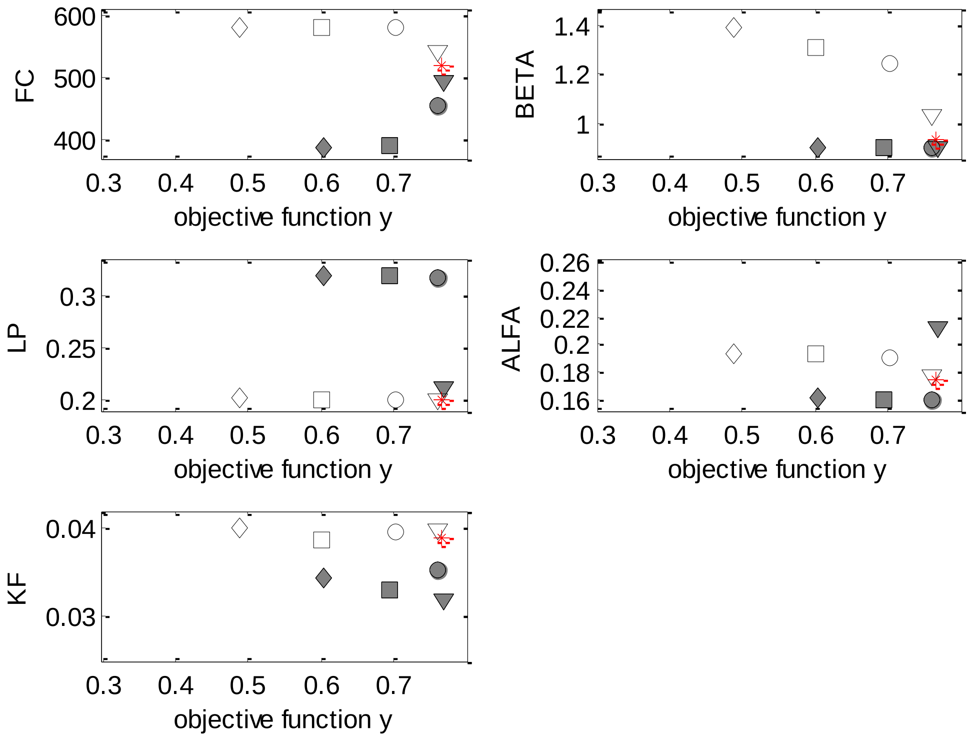


Figure 32: Influence of error source 4 on model parameters Chiers basin

These figures are similar to the systematic errors in error source 1. In both basins, the values for *FC* and *BETA* decrease with a positive systematic error and increase with a negative systematic error. The behavior of *LP* is also comparable with its behavior in error source 1, because a positive systematic error results in an increase of *LP*, whereas a negative systematic error has a decrease in the parameter value as a result. It can be seen that in this error source the borders of the parameter ranges are reached quite fast for the parameters *FC*, *BETA* and *LP*. In paragraph 5.2.1 this appearance is further discussed.

ALFA and *KF* have deviations in their parameter values, but comparable to error source 1, the changes in the parameter values are difficult to explain because no trend is visible. This is the case in both the Ourthe and Chiers basin.

4.5 Discussion: influence of error sources on model performance

In the previous paragraphs the different error sources and their influence on model performance and model parameters have been treated. Looking at the results, some things can be noticed on the sensitivity of the calibrations to the error sources. The expectation of what is realistic and the influences of the different error sources form the basis for the definition of the realistic scenario. The differences in influence on model performance are explained in this paragraph.

Error sources 2 and 3 do not have much influence on model performance or parameter estimation. Especially the effect of error source 3 is negligible. Furthermore, error sources 1 and 4 do have large influences, especially in the extreme situations. This influence is mainly caused by the systematic errors. The random errors without autocorrelation have a very small influence, while the random errors with autocorrelation have more influence regarding the objective function *y*. In general, both basins show qualitatively similar behavior regarding the influence of the different errors on model performance.

The influences of error source 1 and 4 on the model performance are comparable in the Ourthe basin. A difference is that with a certain quality of discharge data, expressed in *BALANCE* and *QOD*, the objective function *y* shows higher values for error source 1 compared to error source 4, but the differences are very small.

In the Chiers basin, a difference occurs between the effects of the two error sources. A decreasing quality (decreasing *QOD* and increasing absolute value of *BALANCE*) result in a bigger influence on objective function *y* in error source 1 compared to error source 4. The difference can have different causes. First of all, it may say something about the discharge regime of the Chiers basin compared to the Ourthe basin. Another reason can be that the high water peaks and the low flows in the Chiers basin occur at unfavorable moments in the calibration period. If for example a Q-h revision takes place during a certain high water event, HBV will find some difficulties in the simulation of this phenomenon. So the value of the objective function will decrease drastically if the systematic errors in this error source become larger.

4.6 Realistic scenario

As indicated in paragraph 3.3, a realistic scenario is set up to investigate the influence of the errors which are expected to occur. It is assumed that the Q-h relation is used in discharge determination in the two sub basins. That is why error source 3 and 4 can be present. Error source 3 (hysteresis effect and properties of a high water event) however, has a very small influence on the model performance, even in extreme situations. Therefore error source 3 will not be used in the realistic scenario. Error source 4 (outdated Q-h relation) will be present in this scenario. When using the Q-h relation, measurements of the water level *h* have to be performed. These measurements can contain some random errors. That is the reason why also random errors with autocorrelation are assumed to be present in the realistic scenario. The influence of these random errors with autocorrelation is quite small, but they are included into the scenario anyway to achieve completeness of the realistic scenario.

4.6.1 Implementation

To implement the realistic scenario into the discharge data, first the effects of an outdating Q-h relation are inserted and after that, the random errors with autocorrelation are added.

The outdating of the Q-h relation is simulated in a similar way as in error source 4. The difference is that in error source 4 the maximum systematic error after each Q-h revision was constant, while in the realistic scenario each time a systematic error with different magnitude is assumed. The maximum systematic error is uniformly distributed over the interval [-20,20] m^3/s. The Q-h relation in this scenario is also revised after each 5 years. This means that in the calibration period of 15 years, three revisions take place.

The random error with autocorrelation is simulated in the same way as in error source 2. The chosen α is equal to 0.5 in both the Ourthe and the Chiers basin, because then in both basins the spread of the random error is around 7%. According to Jansen (2007) the spread in random errors in discharge when using the Q-h relation is between 5% and 10%. The choice for α = 0.5 results in random errors which do not have much autocorrelation in time. Because of the relatively low value for the auto correlation α, the influence of a certain error is noticeable for just a couple of days. This is realistic for certain phenomena which cause random errors that persist a couple of days. For errors for which the influence persists longer, this auto correlation coefficient does not function well. This is further discussed in paragraph 5.1.1.

Together these two kinds of errors are combined into the realistic scenario. In Figure 33 a possible development of the combination of errors is shown. The vertical axis shows the deviation of the modified discharge data to the original data. The increasing systematic error due to the outdating Q-h relation and the random errors are clearly visible in this figure. The horizontal axis shows the calibration time period of 15 years (1984 – 1998).

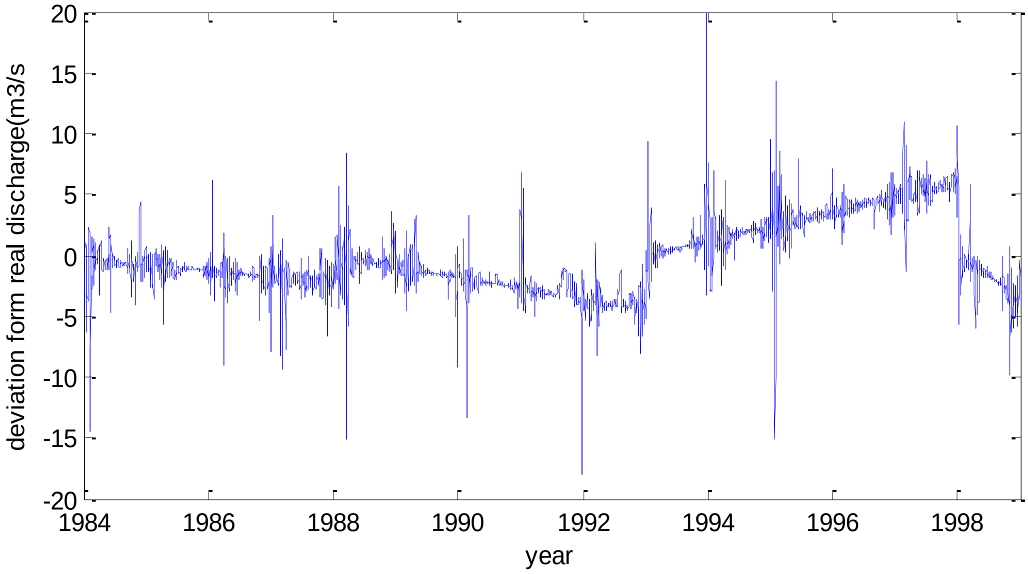


Figure 33: Graph of possible error in realistic scenario

Thirty different discharge series are constructed and with each of these series, a calibration is performed. Each of these calibrations is unique because of the randomness of the systematic error after the period of 5 years and the random errors with autocorrelation.

4.6.2 Influence on model performance

In Figure 34 and Figure 35 the influence on model performance is shown. In these figures different symbols are used. If *BALANCE* has a positive value, then the markers are filled, otherwise the markers are open. Furthermore, if *BALANCE*>0.1 or *BALANCE*<-0.1, circles are used. If -0.1<*BALANCE*<0.1, a triangle is shown. These markers can be useful in the figures with the *QOD* function, because they can make clear what kind of relation is present.

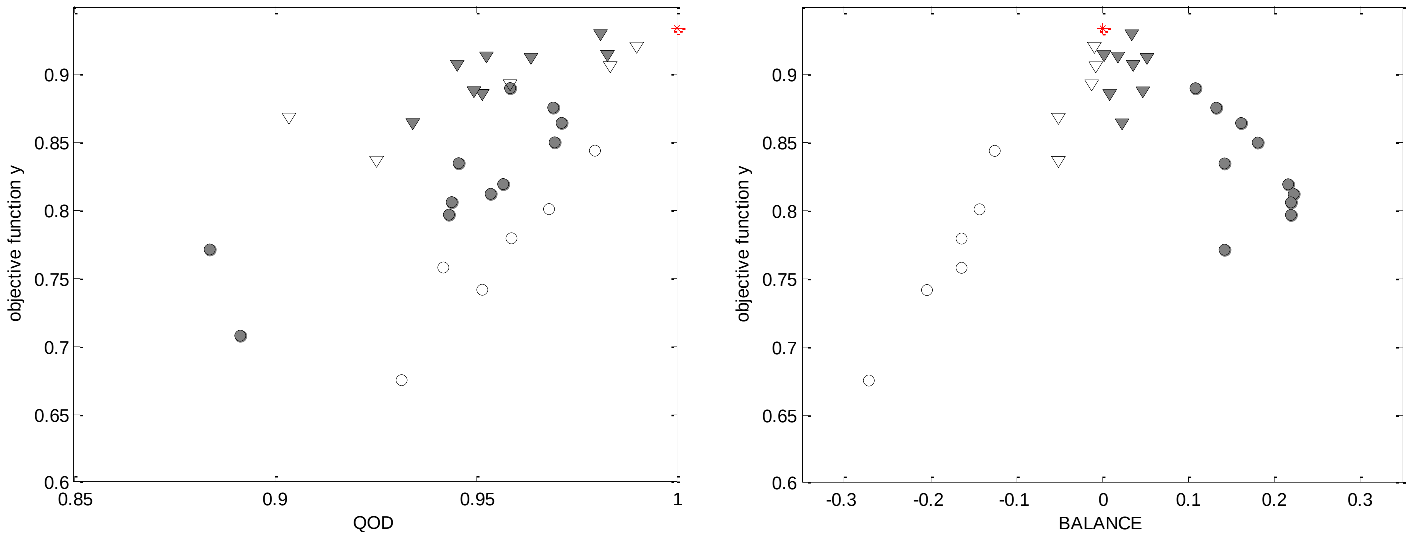


Figure 34: Influence of the realistic scenario on model performance Ourthe basin

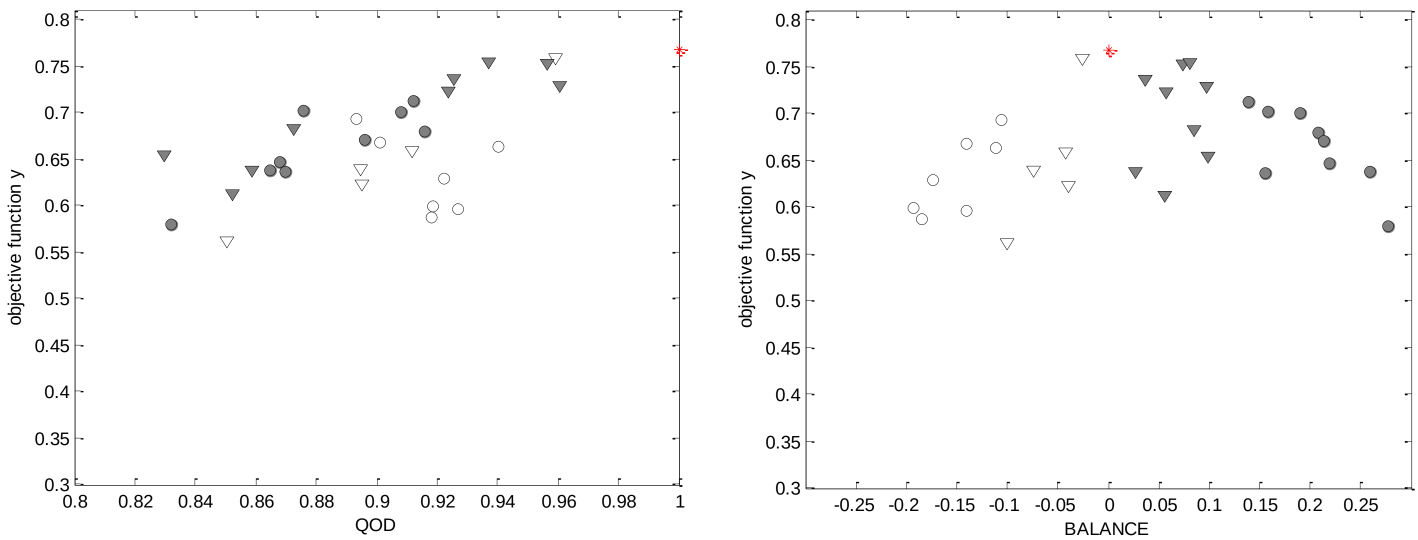


Figure 35: Influence of the realistic scenario on model performance Chiers basin

Figure 34 and Figure 35 show the influence of the different runs of the realistic scenario on model performance. In general, a lower value of *QOD* results in a lower model performance in both basins. In both figures the observation is made that the runs with a negative *BALANCE* (open markers) result in a lower value for the objective function *y*. In the Ourthe basin the circle markers have lower values for the objective function compared to the triangle markers. In the Chiers basin this relation is less obvious.

Deviations from the optimal value of *BALANCE* result in worse model performances. The figures with the values for *BALANCE* show a parabola. This was also observed in error source 1 and 4. In this realistic scenario however, the objective function never exceeds the objective function of the original calibration. This may be caused by the fact that the modified discharge data do not contain a positive systematic error over the entire period, like in error source 1 and 4. There are some calibration runs though in which all systematic errors just before revision of the Q-h relation are positive in the calibration period of 15 years. However, also in these cases the original objective function is never exceeded. The reason for this may be the fact that the systematic errors just before the revision of the Q-h relation are not of the same magnitude. This makes that the HBV model may find some difficulties because it has to generate a single set of parameters over the entire calibration period, with in total three Q-h relation revisions in it, each with a different systematic error just before the revision. Another possible reason for the lower values for the objective function than in error source 1 and 4 is that the random errors with auto correlation are the cause of the decrease in *y* value. This seems not to be the dominant effect for this phenomenon however, because of the fact that the decrease in objective function in error source 2 (random errors with autocorrelation) is quite small.

In the Ourthe basin a pattern can be seen that the highest value for the objective function in the thirty calibrations of the realistic scenario is found with a small positive value for *BALANCE*. This is also the case in error sources 1 and 4. In the Chiers basin this phenomenon is not so clear in the right figure. In the left figure however can be seen that the calibration runs with a positive value for *BALANCE* in general produce the highest value for *y*. The obtained parabolas do not form a line like in error source 1 and 4, but a cloud of points in a shape of a parabola. This is because of the fact that the value for *BALANCE* is mainly composed by the systematic errors due to the outdating of the Q-h relation. This means that the systematic errors between the periods of revision are in most cases not entirely positive or negative. As a result the positive and negative systematic errors in most cases partly counterbalance each other. The value for *BALANCE* consequently approaches 0, but the HBV model still has a lot of difficulties in fitting a discharge data set to these new discharge data so the objective function *y* decreases.

4.6.3 Influence on model parameters

Figure 36 and Figure 37 show the influence of the errors in the realistic scenario on the model parameters. The same markers are used as in the figures in section 4.6.2.

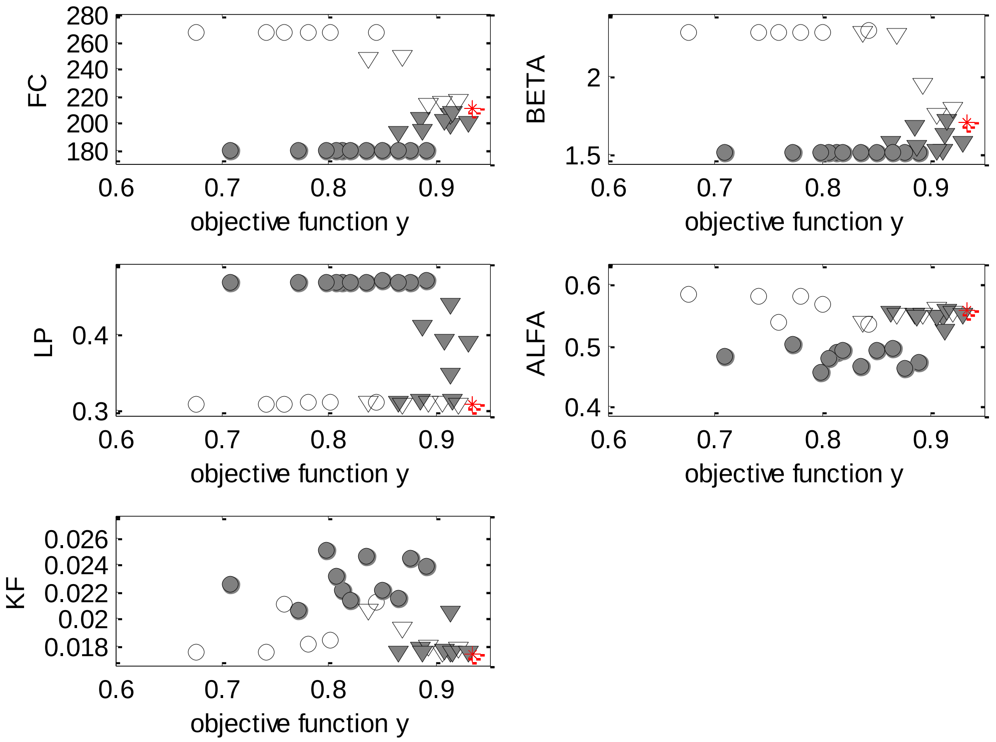


Figure 36: Influence of realistic scenario on model parameters Ourthe basin

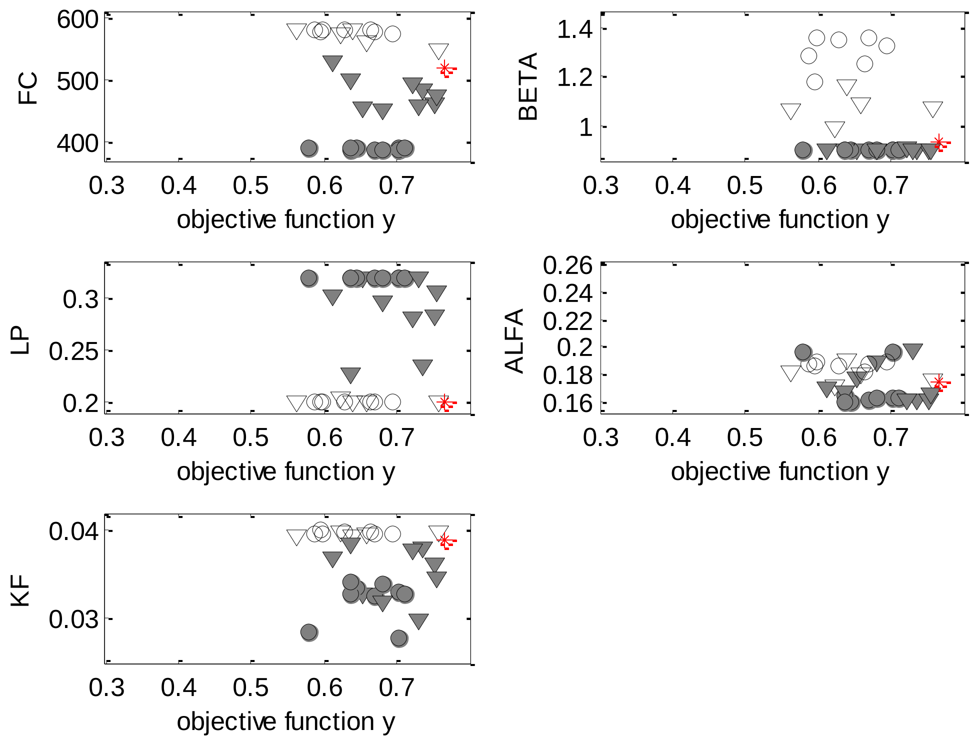


Figure 37: Influence of realistic scenario on model parameters Chiers basin

These figures show similar patterns for the parameters *FC*, *BETA* and *LP*, compared to error source 1 and 4. The cases in which *BALANCE*<0, the parameters *FC* and *BETA* show an increase in objective function, while parameter *LP* decreases in that case. The behavior of these three parameters can be explained by their physical representation.

The other two parameters show a somewhat other behavior. Parameter *ALFA* shows no pattern in the Chiers basin. This was also the case in error source 1 and 4. In the Ourthe basin small deviations from the original data also do not have much influence on the parameter value. With extreme conditions however, some deviations are visible. Very high values for *BALANCE* (>0.1) result in a decrease of *ALFA*, while very low values for *BALANCE* (<-0.1) lead to an increase of *ALFA*. An increase of the *ALFA* value means that at a certain change in storage in the upper response box the fast runoff has a more non-linear behavior than in the original situation. The fact that with low negative values for *BALANCE* the *ALFA* value increases is difficult to clarify, because it would be expected that it would behave the opposite direction as the increase in *ALFA* value will have an increase in the fast runoff as a result.

Parameter *KF* does not show any patterns in error sources 1 and 4 in neither the Ourthe nor the Chiers basin. In the realistic scenario an unexpected phenomenon occurs in the Chiers basin. Positive values for *BALANCE* lead to a decrease in parameter value, while negative values for *BALANCE* lead to an increase of *KF*. In the Ourthe basin this relation is not found. In the Ourthe basin it seems that extreme high values for *BALANCE* result in an increase, while extreme low values for *BALANCE* lead to a small increase or a decrease in parameter value. The expectation would be that a positive value for *BALANCE* results in an increase in *KF* value, because this would mean that at a certain storage in the upper response box an increase of fast runoff would occur. The model in the Ourthe basin largely behaves that way, but the Chiers basin for some reason shows opposite behavior.

An explanation for the unexpected behavior of both *ALFA* and *KF* could be that these parameters are mainly influenced by random errors. This means that a change with a systematic character has more influence on *FC*, *BETA* and *LP*. The influence of the random errors in the realistic scenario is difficult to determine because the systematic errors also have random character.

4.6.4 Uncertainty in parameter values

The parameter values which are found after calibration of different model runs of the realistic scenario may be uncertain. There are two types of uncertainty that can be distinguished. First of all, there can be uncertainty in the parameter value due to the calibration method. The reason for this is that the maximum number of iterations of the SCEM-UA algorithm is set to 1500 or that it is difficult to find a single optimum. The uncertainty in the parameter due to the calibration method is defined as the spread in parameter values of the last 300 iterations (20% of the total amount of iterations). The spread is defined as the 95% confidence interval of the errors. The uncertainty due to calibration method can be identified in the following figures by analyzing the size of the error bars. If a certain error bar is large, the uncertainty due to the calibration method is large.

Secondly, there can be uncertainty in the parameter values due to uncertainties in the discharge data. The realistic scenario contains a set of different possible situations. This leads to a set of 30 different discharge data series. Each data series produces a certain parameter set. In the following figures, this uncertainty can be found if parameter values significantly differ from each other and there is no overlap in error bars. It is assumed that if there is no overlap between the error bars, there is a significant deviation in parameter value.

The errors with a systematic character have influence on the parameters *FC*, *BETA* and *LP*. The uncertainty due to the calibration method is quite small in these parameters, i.e. the last 20% of the iterations do not show a large spread and these parameters are therefore well identifiable. The patterns of the parameters caused by differences in discharge series are already discussed in section 4.6.4. In these parameters almost all deviations in parameter values are significant. Therefore the figures with the uncertainties of the parameters *FC*, *BETA* and *LP* are not presented in this section. The figures that show the uncertainties in the parameter value can be found in appendix 3.

The errors in the realistic scenario that have a random character are less predictable and the model therefore has difficulties in finding optimal values for the parameters *ALFA* and *KF*. In the following figures these two kinds of uncertainties in *ALFA* and *KF* values are shown. In these figures, the dots represent the parameter values of the 30 calibrations of the realistic scenario after the 1500 iterations and the error bars represent the 95% confidence interval. The objective function and parameter value in the original situation are indicated with a marker '*'. Figure 38 and Figure 39 show the uncertainties in *KF* and *ALFA* in the Ourthe basin.

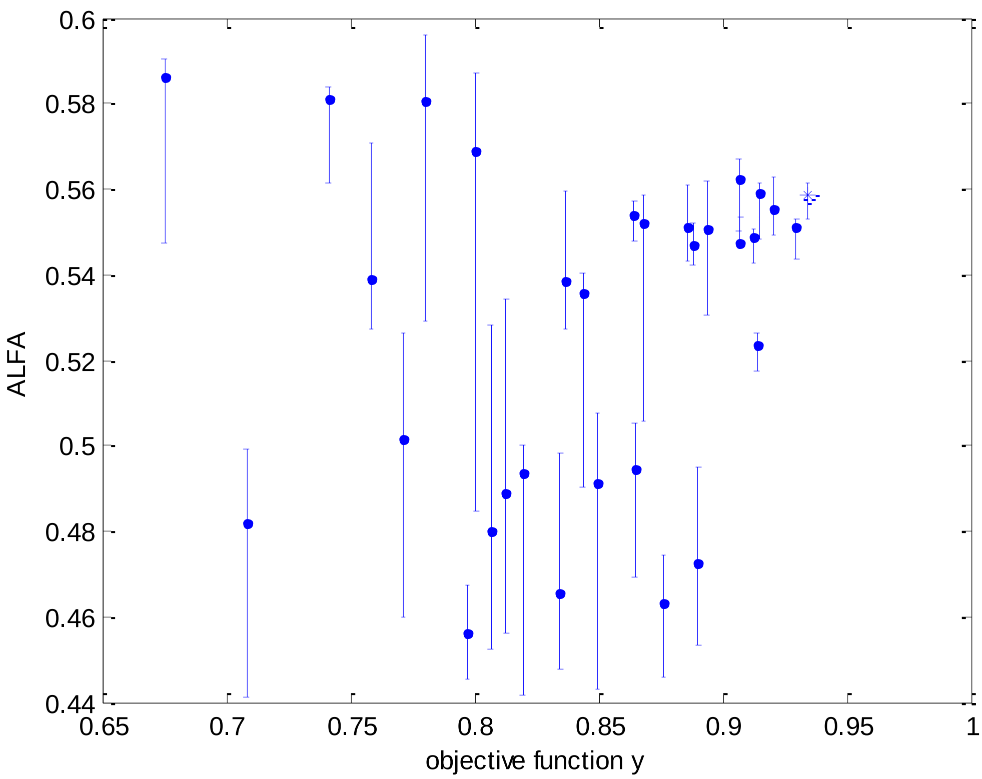


Figure 38: Values for *ALFA* in realistic scenario with 95% confidence interval error bars, Ourthe basin

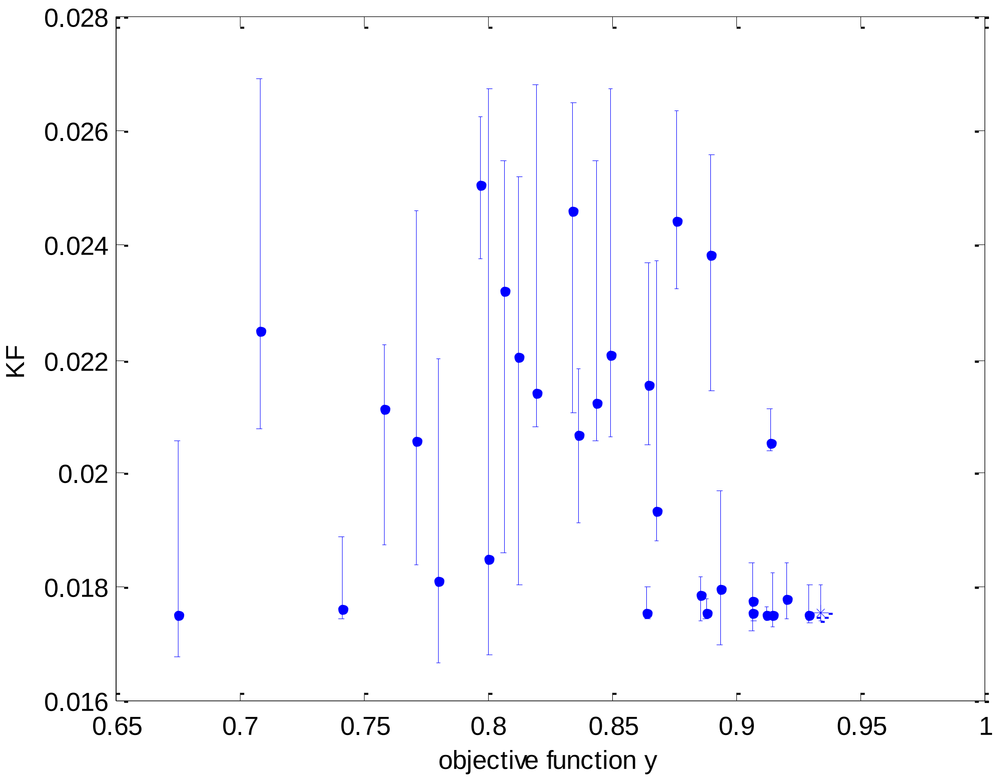


Figure 39: Values for *KF* in realistic scenario with 95% confidence interval error bars, Ourthe basin

Uncertainty due to calibration method, Ourthe basin

Figure 38 and Figure 39 show that the original situation with no errors in the discharge data has a relatively small spread in parameter values of *KF* and *ALFA* in the last 20% of the iterations and therefore are well identifiable. In general it can be concluded that with a lower value for the objective function the uncertainty in the parameter values due to the calibration method increases in the Ourthe basin. This can be explained by the fact that the model has some difficulties in fitting the model parameters to the adapted discharge data. In general it can be concluded that if $y>0.9$, the uncertainty in parameter value is very small. If $0.85<y<0.9$, the uncertainty is somewhat increasing and if $y<0.85$, the uncertainty due to the calibration method is large. This is the case for both *ALFA* and *KF*.

Uncertainty due to discharge data, Ourthe basin

The uncertainty due to the used discharge data in the Ourthe basin can also be found in Figure 38 and Figure 39. Regarding *ALFA*, a convergence towards the original parameter value is shown. One significant deviation from the original value occurs if $y=0.92$. If $y<0.90$, more significant deviations from the original value are present. If $y<0.85$ about half of the calibrations show significant deviations.

Regarding *KF* a large uncertainty in the parameter value appears for $y<0.90$. In that case, about half the calibrations deliver a parameter value which deviates significantly from the original value. In case of $y>0.90$ just one calibration delivers a significant deviation from the original value (at $y=0.92$). That particular calibration also had a deviation in *ALFA* value.

Figure 40 and Figure 41 show the uncertainties in *KF* and *ALFA* values in the Chiers basin.

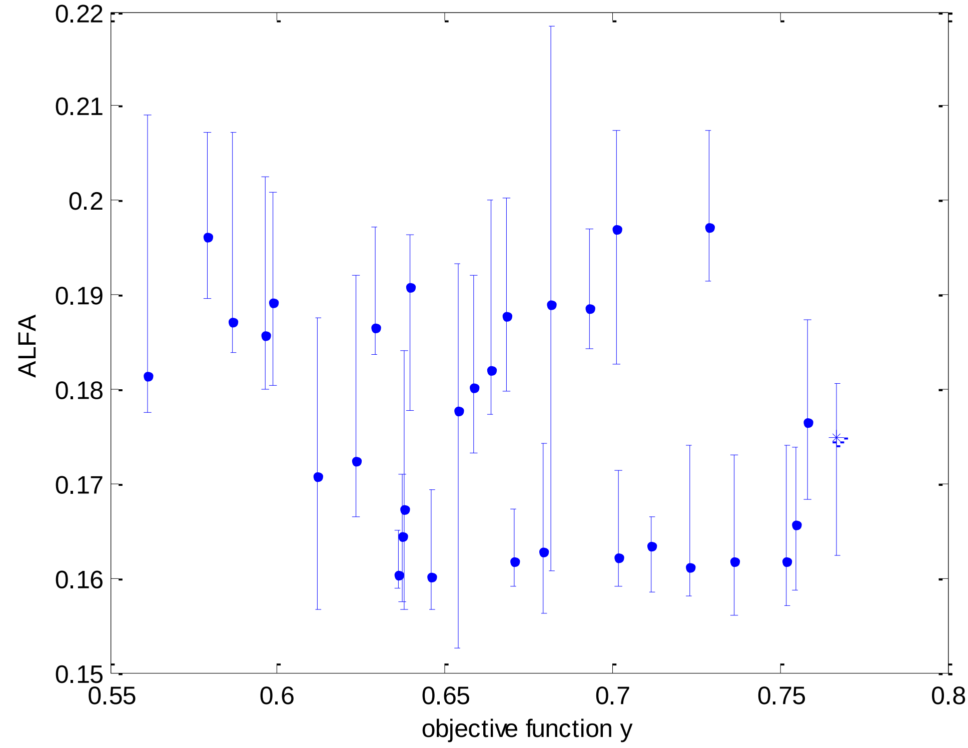


Figure 40: Values for *ALFA* in realistic scenario with 95% confidence interval error bars, Chiers basin

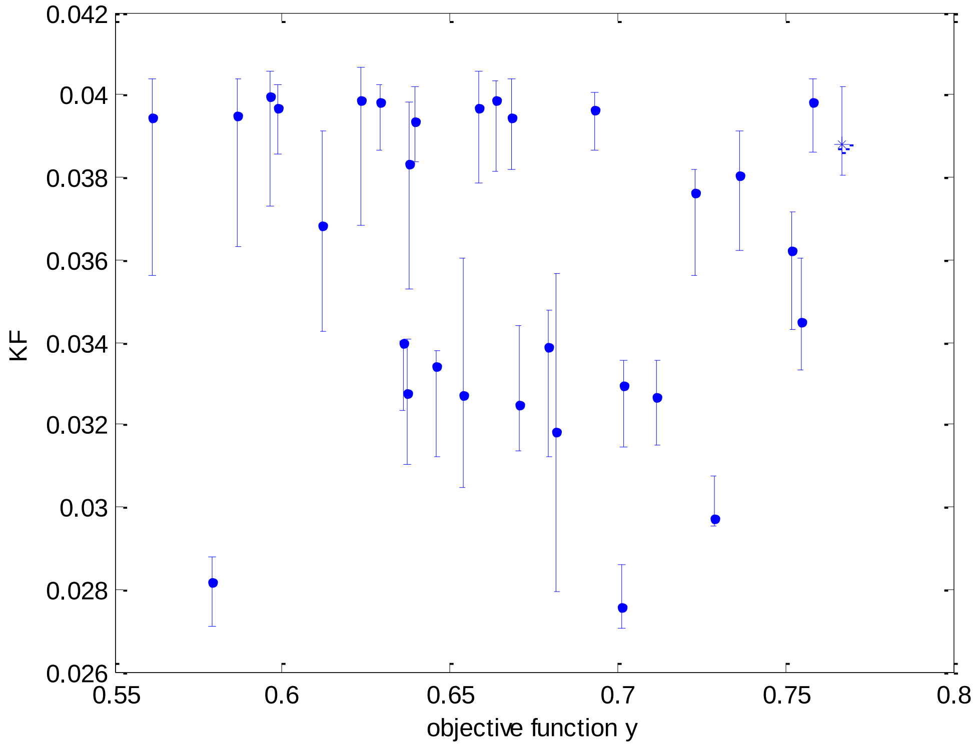


Figure 41: Values for *KF* in realistic scenario with 95% confidence interval error bars, Chiers basin

Uncertainty due to calibration method, Chiers basin

About the uncertainty due to the calibration method there are a couple of observations. The first thing that is remarkable is that the original situation in the Chiers basin has a larger spread around the parameter value compared to the spread of the original situation in the Ourthe basin. This can be explained by to the fact that the HBV model has more difficulties in finding the optimal solution in the Chiers basin, also considering the lower values for the objective function *y* compared to the values in the Ourthe basin in all calibrations.

Another notable fact is that the value of the objective function does not have influence on the uncertainty in parameter value due to the calibration method in the Chiers basin. Contrary to the Ourthe basin, where a lower value for *y* leads to a larger uncertainty due to the calibration method, in the Chiers basin the uncertainty is not depending on the objective function. For every value for the objective function the parameters are bad identifiable. This can be explained by the fact that the uncertainty in parameter values *ALFA* and *KF* in the Ourthe basin are relatively small if $y>0.90$. In the Chiers basin *y* never exceeds the value of 0.90. This can be a reason that the parameter values *ALFA* and *KF* in the Chiers basin are uncertain, even at the highest values for *y*.

Uncertainty due to discharge data, Chiers basin

The uncertainty due to differences in discharge data does not show a clear pattern for the parameter *ALFA*. If $y<0.70$ the value significantly deviates from the original parameter value in about half of the calibration. If $y>0.70$ just one calibration significantly deviates ($y=0.73$). So a certain convergence is present for this parameter.

For parameter *KF* no clear pattern is visible. There are significant deviations over the entire domain of *y*. Even if the objective function decreases a little, some deviations occur. Furthermore, a lot of values do not deviate significantly from the original value only because the maximum value in the parameter range is reached. It can be assumed that if the maximum value of the parameter range is higher, a lot more significant deviations would be present. This means that the value of *KF* due to the discharge data is quite uncertain.

Comparison uncertainty in parameter values Ourthe and Chiers

The difference between the Ourthe and Chiers basin regarding the uncertainty in parameter values of *ALFA* and *KF* due to the calibration method is that in the Ourthe basin the uncertainty is depending on the value of the objective function. If the value for *y* is small, the parameter value is uncertain. If the value for *y* is close to the original value, the parameter value is not uncertain. This applies to both *KF* and *ALFA*. The expected reason for this is that the objective function in the original situation is higher in the Ourthe compared to the Chiers basin.

Regarding the uncertainty in parameter value due to the uncertainty in discharge data the Ourthe basin clearly shows a convergence in parameter value with an increasing value for the objective function for both parameters. This means that in situations in which the objective function value for the realistic scenario does not deviate much from the original objective function value, the parameter values of *ALFA* and *KF* are quite certain. In the Chiers basin this convergence cannot be distinguished in any of the parameters. For both parameters many significant deviations over the entire domain of *y* are present. This means that both parameters are very sensitive to the different situations of the realistic scenario and therefore the parameter values of *KF* and *ALFA* are uncertain in the Chiers basin due to the differences in discharge data. The reason for the uncertainty in the Chiers basin can be the fact that the model performance in the Chiers is anyway lower than in the Ourthe basin.

Comparison of the two types of uncertainty

There seems to be a connection between the two different types of uncertainties in the Ourthe basin. In the Ourthe basin the uncertainty due to the calibration method increases if the uncertainty in discharge data also increases. In the Chiers basin however, this phenomenon is not visible. The reason for this might be that both types of uncertainties are quite large in this basin. The expectation is that in general with increasing uncertainty in discharge data, the uncertainty due to the calibration method also increases.

5 Discussion of methodology and results

In chapter 4 the influence of different sources of errors, as well as a realistic scenario of uncertainty in discharge determination is examined. In this chapter, some elements of the research are critically reviewed. In paragraph 5.1 the choices made in the error sources are discussed. Paragraph 5.2 contains some discussions about the used model calibration. In paragraph 5.3 the differences between the sub basins are treated. In paragraph 5.4 an additional uncertainty in the research is discussed, namely the phenomenon that a presence of a systematic error leads to a rise in model performance in both basins.

5.1 Error sources

In the four error sources assessed in the current study, different elements of their representation are debatable. In this paragraph these elements are discussed.

5.1.1 Random errors with autocorrelation

An assumption in the realistic scenario is that the measurements contain random errors with autocorrelation. According to the research of Jansen (2007) the spread of the random error when using a Q-h relation is between 5% and 10% in the discharge values. In this research it is assumed that the random errors are auto correlated in time. There are no references found that support that assumption, but it is still used because it seems plausible due to the possible phenomena that may cause the errors. Some of the errors will persist a couple of days, while other errors are just present in one measurement. For example, errors which are caused by impoundment from a heavy long lasting storm may have influences over a longer period than errors which are caused by a single misuse of the measuring equipment. The different periods over which the errors persist, request for different auto correlation coefficients α. In this research just one value for α is chosen in the realistic scenario. The autocorrelation coefficient α is set to a value of 0.5 in both sub basins, because that leads to a spread of around 7% (between 5% and 10%) in the deviations of the discharge for both the Ourthe and Chiers basin. This value of the autocorrelation coefficient means that the time scale in which a certain error fades away is a couple of days and that errors with different time scales are not included in this case. The chosen value for α can be justified because it could be an average of possible values for α.

A part from the autocorrelation coefficient a choice of absolute instead of relative random errors could have been made as well to simulate kinds of errors with different character. This is not done because in the research of Jansen (2007) the spread of random errors is defined as a percentage of the real value.

Another shortcoming in simulating random errors is that every calibration run is different because of the unpredictability of the error. This means that every calibration contains a situation which could be the case in reality, but a realistic description of the error is difficult to construct. Even if for example the spread in random error in every situation is 7%, this will lead to different hydrographs. This means that the real situation is difficult to simulate. Only a set of possible realistic scenarios can be constructed and it is not sure whether the real situation is present in that set of possibilities. This problem can be solved by simulating a very large number of possible situations. In the current research the realistic scenario contains 30 possibilities. To get a more complete analysis it is recommended to increase this number a lot (for example 1000 runs), so that a wide spectrum of possible outcomes arises.

5.1.2 Presence and magnitude of systematic errors

Different kinds of systematic errors are present in error source 1 (fixed relative deviation), error source 4 and the realistic scenario (both absolute deviations due to expiration of a Q-h relation). Whether a systematic error is a relative or absolute deviation from the real value, or a combination, is not known. The choice for the kind of deviation of the systematic error is made as follows: in error source 4 and the realistic scenario the systematic error arises from the outdating of the Q-h relation. This is often caused by processes like, for example, changes in the geometry of the cross section or the roughness of the river bed. That is why a choice is made for an absolute value for the systematic error. In error source 1 however the cause of the systematic error is not known, so a relative deviation from the original value is assumed. This is done because of the fact that if the discharge is high, the possible deviation in discharge from the real value will be higher than with a low discharge, so this option seems to be the most logical one. For example in the research of Jansen (2007) only relative deviations occur.

In general, it is difficult to identify systematic errors in discharge determination. If a certain systematic error would be known, it would be easy to take that error out of the discharge data and no systematic error would be present any more.

The magnitude of the systematic error after an expiration of a Q-h relation (error source) is an assumption that is not based on literature. Because the maximum systematic error is ± 20 m^3/s after 5 years and the average discharges in the Ourthe and Chiers rivers are respectively 23 m^3/s and 25 m^3/s in the calibration period, this can lead to some unrealistic situations in which periods with no discharge occur, as well as periods in which the discharge is twice as high compared to the original situation. These situations however are exceptional and only occur in a few of the calibrations. These calibrations are used as simulations of an extreme situation.

5.1.3 Use and outdating of Q-h relation

Another assumption is that a Q-h relation is used for discharge determination in both rivers. This assumption is not based on known characteristics of the measurement station, but on the fact that the use of the Q-h relation is likely to have taken place. Furthermore, in the error source an assumption is made that a revision of the Q-h relation takes place after a fixed period of 5 years. This is not always the case. Revisions of Q-h relations often take place after a high water event because a lot of erosion or sedimentation can take place in such event (Jansen, 2007). The revision of the Q-h relation is done to make the relation up to date again. An assumption is made that on average the time between two revisions is five years.

Also the choice of the moments that a Q-h relation revision takes place is arbitrary. In this research the first revision takes place after 4 years from the beginning of the calibration period, on 1 January 1988. This can have some influence on the model calibrations because it gives other outcomes if a certain high water event always is located just before or just after a revision of a Q-h relation. The influence of this can be examined by changing the date of the first revision.

5.1.4 Uncertainties in parameter values *KF* and *ALFA* after calibration

The parameters *KF* and *ALFA* have two different uncertainties. Firstly, different discharge data result in different values for these parameters so the parameter value due to measurement uncertainties is large. This can be explained by the fact that the objective function is sensitive to errors that have a random character, because they are closely related to the shape of the hydrograph.

A part from uncertainties in parameter value due to measurement uncertainties also the uncertainty in parameter value after 1500 iterations of the SCEM-UA algorithm is quite large. This means that the parameter values are quite uncertain due to the calibration method. This problem could be solved by increasing the maximum of iterations. This gives the chance to the SCEM-UA algorithm to converge to a single parameter value instead of a range of parameter values.

5.2 Model calibration

5.2.1 Parameter ranges

For the execution of the calibrations, the same parameter ranges are used as in the original calibration. Because with this method the optimum is often found at the border of one or several parameter ranges, the research would possibly be more accurate if the parameter ranges were adjusted in every calibration. This means that after a calibration is performed, an analysis would have to be made whether the global optimum is located in the used parameter space. Because in this research a lot of calibrations had to be performed, it would have cost a lot of calculation time if every calibration had to be executed several times. A part from this, it might not be correct if every parameter could be varied over all values, because the parameter represent a physical value. Some situations would be unrealistic, for example negative values for the parameters. It is recommended to expand the ranges in Table 3 somewhat, in a way that the borders of the range are not reached too fast, but on the other hand, the minimum and maximum value still have to be realistic values.

5.2.2 Use of SCEM-UA optimization algorithm

The SCEM-UA algorithm is used in this research which is a global optimization method, developed by Vrugt et al. (2003a). More information about the method can be found in appendix 1. The SCEM-UA calibration method performed very well in this research. In this research the maximum number of calibrations was set to 1500. In some cases this number of iterations was not enough to determine the optimum, with an uncertainty in some parameter values as a result.

Furthermore, the SCEM-UA algorithm meanwhile has a successor. The DREAM-UA algorithm (Vrugt et al., 2008) is also a global optimization procedure that performs even better than the SCEM-UA algorithm. This algorithm would have found the optimum faster than SCEM-UA and the use of DREAM-UA would result in less uncertainty in the parameter values due to the calibration method.

5.2.3 Choice of optimization of 5 out of 8 parameters

In the sensitivity analysis in section 2.4.2 a choice is made for an optimization of 5 out of the 8 parameters of this version of the HBV model. The other three parameters (*CFLUX*, *PERC* and *KS*) have a fixed value in all calibration. The choice for five parameters instead of eight is made because it resulted in a decrease in calculation time and the selection of the five used parameters is based on the sensitivity of the eight parameters in the calibration with the original discharge data. However, it could be that certain deviations due to uncertainty in discharge determination result in an optimization in which the unused parameters *CFLUX*, *PERC* and *KS* could have been useful in the calibration, to approach the new discharge data. *PERC* and *KS* for example characterize the inflow and outflow of the lower response box. If these parameters were used in the calibration, perhaps the low flows could have been adapted more easily if for example a systematic error was present in the discharge data.

5.3 Differences between Ourthe and Chiers

5.3.1 Model performance

The main difference between the Ourthe and Chiers basins is that in all calibrations the objective function *y* is significantly higher in the Ourthe basin than in the Chiers basin. This result is also found in other studies, like the research of van Deursen (2004).This can either be caused by a better quality of data of the Ourthe basin, but can also be caused by the fact that the HBV model is more suitable for discharge regimes like in the Ourthe basin and that therefore the discharge in the Ourthe river can better be simulated by the model. The reason for this could be that the HBV model can handle a discharge regime with low base flow and high peaks better than a discharge with high base flow and lower peaks. This seems not to be the case as in literature this conclusion never has been drawn. As seen in Figure 5, the hydrograph in the Ourthe basin contains high peaks and low base flows compared to the Chiers basin. The fact that the sub basins have these typical hydrographs might be caused by the fact that the Ourthe river has a steeper slope than the Chiers river. This means that the average streaming velocity of the Ourthe is higher and therefore the hydrograph is more capricious.

5.3.2 Parameter estimation

The influence of the error sources and realistic scenario on the model parameters is comparable in the two sub basins. The two types of uncertainty in parameter values are treated in section 4.6.4. The uncertainties in parameter values due to different discharge data do not differ much between the basins. The parameters which are related to the water balance *FC*, *BETA* and *LP* behave in a way that is expected according to their physical representation. Regarding the parameters *ALFA* and *KF* no clear pattern can be distinguished in both basins. This is because of the fact that these parameters are not influenced by the systematic errors, but on the random errors.

The uncertainties in parameter value due to the model calibration method differ between the sub basins for the parameters *ALFA* and *KF*. In the Ourthe basin, the uncertainty increases with a decrease in objective function, while in the Chiers basin the uncertainty due to the calibration method is large and independent on the value of the objective function.

Merz and Blöschl (2004) investigated the influence of some basin characteristics on parameters of a version of the HBV model. The basin average topographic slope has a significant influence on the parameters *BETA* and *LP*. If the parameters of the ‘base case’ calibration are considered, it is clear that these parameters differ significantly between the two basins, because the bottom slopes differ a lot between the basins ($3.7 * 10^{-3}$ for the Ourthe and $1.0 * 10^{-3}$ for the Chiers river).

5.4 Positive systematic error: rise of *y*

Up to a certain positive systematic error, the objective function *y* increases in both the Ourthe and Chiers basin. The reason for this is difficult to determine. Apparently this version of the HBV model can simulate the original data better with a positive systematic error than without the positive systematic error. The existence of this result can have several reasons. The reason can be a deficiency of the model, the use of this particular objective function or a bad quality of the present discharge or input data in the Ourthe and Chiers basins.

First of all, it could be a result of using the HBV model. It could be that the HBV model always performs better if a certain positive systematic error is present. This can be examined by performing the same research, but with a different hydrological model that does not have this problem.

A part from the model, also the use of this objective function can be the reason for this phenomenon. The use of the objective function *y* can be a cause for the appearance of the rise in model performance. The objective function *y* combines *RVE* and *NS* in a way that the optimal value of *y* is reached if *RVE* is close to 0. The model thus tries to approach a value of 0 for *RVE*. This means that if the model cannot simulate the changes or peaks in the hydrograph well, it tries to compensate for this in other periods. This makes that low flows tend to get too low, because the HBV model always has some difficulties in simulating high peaks, as seen in paragraph 2.4.3. It could be that if a certain positive error is added, the model performance increases because it can compensate more easily for the high flows because the low flows are somewhat higher. The use of another objective function in which *RVE* has less or no importance can show if this is the case.

Another possibility is that in both the Ourthe and Chiers basin a certain systematic error is present in the input or output data. One option is the presence of a negative error in the discharge data. In this case, by adding a positive systematic error to the wrong, underestimated data the HBV model can simulate these discharge data better. This possibility can be examined by executing the same research on different sub basins. If the phenomenon does not occur in all other basins, it may be an indication that the present discharge data in the Ourthe and Chiers basins are underestimated. However, it can also be coincidental that the model has some difficulties in approaching the discharge data. If the phenomenon occurs at all basins, the conclusion can be drawn that there is a big chance that the HBV model does not function well and that the model always prefers an overestimation of discharge data.

A part from the discharge data, the problem can also be located in the input data, like the precipitation or potential evapotranspiration. It could be that the precipitation data series are overestimated or the potential evapotranspiration data contain an overall underestimation. In this case the discharge data with a positive systematic error perhaps are better approachable by a wrong input data set. This can be examined by adding a certain systematic error to the input data and see whether the model performance increases after calibration. Like in the discharge data however, it is difficult to determine whether there are systematic errors present in input data sets, so it is not easy to perform a realistic investigation.

6 Conclusions and Recommendations

The conclusions of the research are the answers to the most important research question. In paragraph 6.1 the conclusions of the research are summed up. After that, some recommendations for further research are listed in paragraph 6.2. These recommendations are derived from the discussion in chapter 5 and the conclusions in paragraph 6.1.

6.1 Conclusions

This research contains three research questions. The first two questions formed a foundation in order to be able to answer the third research question. The first two questions were:

1. *Which version and schematisation of the HBV model, which sub basins in the Meuse River Basin and which calibration procedure are most adequate for calibration?*

2. *What kind of uncertainties in discharge determination can be present and how can these errors be brought into existing discharge time series?*

The first question is answered in chapter 2, in which the data collection and research methodology is presented. The second question is treated in chapter 3, in which all uncertainties in discharge determination are discussed, as well as the implementation of these errors into existing discharge data series. Most attention is paid to systematic errors, random errors with autocorrelation and uncertainties in discharge due to an expiration of the Q-h relation.
The third question is the most important one in this research and it directly contributes to the research objective. The research objective was stated as follows:

The objective of this study is to investigate the influence of uncertainties in discharge determination on the estimation of the parameters and the performance of a lumped version the HBV model for two sub basins in the Meuse River, by applying an automatic global searching calibration method and using adapted observed discharge time series.

The most important research question that directly contributes to this objective was the following question:

3. *What is the effect of uncertainties in discharge determination on model performance and parameter estimation of the HBV model applied to different sub basins of the Meuse River?*

This research question has two aspects, namely the influence of uncertainties in discharge on model performance, and the influence of discharge uncertainties on parameter estimation. In the following sections these two aspects are treated. Section 6.1.1 treats the influence of discharge uncertainties on model performance, while in 6.1.2 the influence on model parameter estimation is pointed out.

6.1.1 Influence of discharge uncertainties on model performance

What is the effect of uncertainties in discharge determination on model performance of the HBV model, applied on different sub basins of the Meuse River basin?

The quality functions *QOD* and *BALANCE* are used to characterize the quality of the discharge data. Different sources of errors have different effects on the quality functions and the model performance, expressed in the combined objective function *y*. In general it can be concluded that unfavorable values for *QOD* and *BALANCE* lead to lower values for model performance than if the quality of the discharge is high.

Random errors with autocorrelation have some influence on model performance on the used HBV model in the Ourthe and Chiers basins. Random errors without autocorrelation have a marginal effect on the quality functions and model performance.

Systematic errors have much influence on the discharge quality functions *QOD* and *BALANCE*, as well as on model performance, expressed in the objective function *y*. Because it is difficult to detect systematic errors, the real magnitude of systematic errors is difficult to determine. In case of a certain small positive systematic error (+5 m^3/s in the Ourthe, +10 m^3/s in the Chiers basin) in the present discharge data, an increase in model performance, expressed in the objective function, occurs in both sub basins. The reason for this can be a bad quality of the input or output data, a shortcoming of the used HBV model or the choice for this particular objective function.

Error sources which are a result of a wrong discharge determination by using the Q-h relation, such as the hysteresis phenomenon or deviations due to the properties of a high water event, do not have any significant influence on the discharge quality, nor on the model performance of the HBV model applied on the Ourthe and Chiers basin. The expiration of a Q-h relation however, in which a systematic error arises in the period after a revision, has a large influence if the systematic error if the magnitude of the systematic errors is also large.

The realistic scenario contains a set of possible realistic errors in discharge determination. In this set of possible scenarios two kinds of errors are used: errors caused by an expiring Q-h relation and random errors in the water level which are auto correlated in time. The effects of the expiration of the Q-h relation and the random errors are different for every calibration. This leads to a large spread in outcomes of these calibrations. In general, unfavorable values for the discharge quality functions lead to a worse model performance. Only if systematic errors are present which counterbalance each other, *BALANCE* will not get influenced much, while the value of *QOD* is low. In this case the objective function will not have a high value as well. If the systematic errors caused by the Q-h relation are mainly positive (so if a positive value for *BALANCE* is present), the model performance in general is better than if they are negative.

QOD and in particular *BALANCE* give a good picture of the effects of the errors on model performance. Some patterns recur, particularly if model performance is expressed against *BALANCE*. The highest value for objective function is found if a certain positive value for *BALANCE* is present and the objective function value in general decreases with decreasing *QOD* value.

This version of the HBV model has a better model performance in the Ourthe basin than in the Chiers basin. This is concluded for not only the situations with the present available data, but for all error sources and the realistic scenario as well. There are different possible reasons for this. Firstly the quality of the data in the Ourthe basin could be better than in the Chiers basin. Another possibility is that the HBV model can perform better in basins which have a discharge regime with low base flow and high peaks like the Ourthe basin, compared to basins with a higher base flow and less high peaks.

6.1.2 Influence of discharge uncertainties on model performance

What is the effect of uncertainties in discharge determination on the estimation of the parameter set of the HBV model, applied to these sub basins?

Error sources which contain a systematic error, such as the combination of systematic and random errors without autocorrelation or an outdated Q-h relation and the developed realistic scenario have effects on model parameters *FC*, *BETA* and *LP*. These effects can be interpreted by their physical representation. For these parameters no big differences between the Ourthe an Chiers basins are found. The parameters *FC*, *BETA* and *LP* have a small uncertainty due to model calibration after the 1500 iterations of the SCEM-UA algorithm and are therefore well-identified. The uncertainty due to the used discharge data is quite large, because the parameter is rather sensitive to systematic errors. If a systematic error is present, the parameter value quickly reaches the borders of the parameter range.

The parameter estimation of *ALFA* and *KF* is more uncertain in both types of uncertainties. If the parameter estimation is expressed against the model performance, no clear obvious patterns can be distinguished. Therefore the uncertainty in parameter value due to the used discharge data is large, because values within the entire parameter ranges are found and no patterns are visible. Also the uncertainty due to limitations of the calibration method are present in both the Ourthe and the Chiers basin. Only if the objective function approached its value in the situation with the original data in the Ourthe basin, the uncertainty in parameter value is small. As soon as the objective function decreases, the uncertainty in parameter value of *ALFA* and *KF* increases. In the Chiers basin the uncertainty in parameter value of *ALFA* and *KF* is always significant. For the Chiers basin it is concluded that *ALFA* and *KF* are badly identifiable.

Comparable to the model performance, the discharge quality functions *QOD* and *BALANCE* give a good picture of the effect on model parameters. Especially the influence of the *BALANCE* function on the well-identified parameters (*FC*, *BETA* and *LP*) show logical patterns that recur if comparable systematic errors are present.

6.2 Recommendations

There are several recommendations for further research which are derived from the discussion and the conclusion of this research. Two issues can be identified that require some further research. First of all the possible error sources could be analyzed better and furthermore the calibration method could be improved. A part from recommendations for further research, a recommendation for the application of this research is given.

6.2.1 Error sources

It is recommended to extend the studying to the types of errors that can be present in discharge determination. For example the spread of the random errors with autocorrelation and the degree of autocorrelation is recommended to investigate. The choice for just one autocorrelation coefficient implies an existence of a kind of error that persists a certain period, based on the autocorrelation coefficient. Because there are different phenomena that have various periods of persistence, the choice for just one autocorrelation coefficient might not be the best option. It is recommended to insert random errors with autocorrelation that have different periods of persistence.

Besides the random errors, also more research is needed about the influence of an expiration Q-h relation. After every revision of a Q-h relation an estimation can be made of the systematic error in the time period between the two dates of revision of the relation. With interpolation techniques the outdating of a Q-h relation can be investigated in more detail.

In the model simulations the observation is done that a small positive systematic error leads to an increase in value of the objective function with the used data. It is not known what the reason for this behavior is. It is recommended to do some research about this remarkable phenomenon. It could be caused by a bad quality of discharge or climate data, the behavior of the HBV model or the use of the objective function. A way to investigate this is to perform a similar research in another sub basin from which it is known that it contains data with good quality, or with another rainfall runoff model, or with the use of another objective function.

6.2.2 Calibration

To achieve more knowledge about the influence of uncertainty in discharge determination on calibrations of a hydrological model, it possible to improve the calibration method, because some uncertainties in model parameters are still present. A recommendation to minimize the uncertainty due to the calibration method is to increase the number of maximum iterations or to use improved calibration schemes, such as the updated version of SCEM-UA, namely DREAM-UA (Vrugt, 2008).

Another recommendation is that it might be better to use somewhat larger parameter ranges to improve the model performances. The used parameter ranges were based on the original situation and narrowed for the specific sub basins. In situations including errors these parameter ranges were found to be too narrow, as many optimal parameter values were at the border of a range and the choice for wider parameter ranges should be made. It should be kept in mind however that parameter ranges should be kept realistically.

6.2.3 Recommendation for application

The importance of this research is that it is demonstrated that the often used assumption of the absence of uncertainties in discharge data for hydrological modeling is not always justified. This research approves that this version of the HBV model only can function well for the Ourthe and Chiers basins if it is known that there are not much uncertainties in the discharge data that were used for model calibration. Particularly the importance of systematic errors in discharge data is demonstrated in this research. Some more research is needed to investigate whether this is also the case in other basins and/or hydrological models. Furthermore it is recommended to ensure that good quality of discharge data is used for model calibration in order to do reliable simulations in hydrological modeling.

References

Akhtar, M., Ahmad, N., Booij, M.J.; *Use of regional climate model simulations as input for hydrological models for the Hindukash-Karakorum-Himalaya region*, Hydrol. Earth Syst. Sci., 13, 1075-1089, 2009

Andréassian, V., Perrin, C., Michel, C., Usart-Sanchez, I., Lavabre, J.; *Impact of imperfect rainfall knowledge on the efficiency and the parameters of watershed models,* Journal of Hydrology 250 pp206-223, 2001

Bárdossy, A., Das, T.; *Influence of rainfall observation network on model calibration and application,* Hydrol. Earth Syst. Sci. 12 pp77-89, 2008

Barneveld, H.J., Bastings, A.J.T.; Report: *QH-relatie Borgharen-Dorp, Ontwikkeling vanaf 1993*, Rijkswaterstaat Directie Limburg, 1998

Berger, H.E.J.; *Flow forecasting for the River Meuse,* Ph.D Thesis, Delft University of Technology, Delft, 1992

Bergström, S.; *Development and Application of a Conceptual Runoff Model for Scandinavian Watersheds*, SMHI Report RHO No.7, 1976

Bergström, S.; *The HBV-model, in: V.P. Singh (Ed.), Computer Models for Watershed Hydrology, Water Resources Publications*, pp. 443–476, 1995

Boiten, W.; *Hydrometry*. Universiteit Wageningen, 1986.

Booij, M.J.; *Impact of climate change on river flooding assessed with different spatial model resolutions*, Journal of Hydrology 303 pp. 176–198, 2005

Booij, M.J.; *Appropriate modeling of climate change impacts on river flooding,* Ph.D Thesis, University of Twente, Enschede, The Netherlands, 2002a

Booij, M.J.; *Modeling the effect of spatial and temporal resolution of rainfall and basin model on extreme river discharge*, Hydrological Sciences Journal 47(2) pp307-318, 2002b

Booij, M.J., Krol, M.S.; *Balance between calibration objectives in a conceptual hydrological model*, Hydrological Sciences Journal, submitted in 2010

Bormann, H.; *Impact of spatial data resolution on simulated catchment water balances and model performance of the multi-scale TOPLATS model*, Hydrol. Earth Syst. Sci. 10 pp165-179, 2006

Bras, R.L., Rodríguez-Iturbe, I.; *Random Functions and hydrology*, Dover Publications Inc., New York, U.S.A., 1993

Brath, A., Montanari, A., Toth, E.; *Analysis of the effects of different scenarios of historical data availability on the calibration of a spatially-distributed hydrological model,* Journal of Hydrology 291 pp232-253, 2004

Deursen, W. van; Report: *Afregelen HBV model Maasstroomgebied,* Carthago Consultancy, Rotterdam, January 2004

Dong, X., Dohmen-Janssen, M., Booij, M.J.; *Appropriate spatial sampling of rainfall for flow simulation,* Hydrological Sciences Journal 50(2) pp279-298, 2005

Jansen, P.C.; *Beoordeling van de kwaliteit van hoogwaterparameters berekend voor de Nederlandse Maas*, Universiteit Twente, Enschede, The Netherlands, 2007

Jansen, P. Ph., Bendegom, L. van, Berg, J. van den, Vries, M. de, Zanen, A.; *Principles of River Engineering, The non-tidal alluvial river*, Delftse Uitgevers Maatschappij, Delft, The Netherlands, ISBN 90 6562 146 6, 1979

De Kok, J.L., Booij, M.J.; *Deterministic-statistical Model Coupling in a DSS for River-Basin Management*, Environ Model Assess 14, pp 595-606, 2009

Liden, R., Harlin, J.; *Analysis of conceptual rainfall-runoff modeling performance in different climates*, Journal of Hydrology 238 pp231-247, 2000

Lindström, G., Johansson, B., Persson, M., Gardelin, M., Bergström, S.; *Development and test of the distributed HBV-96 hydrological model*, Journal of Hydrology 201 pp272-288, 1997

Merz, R., Blöschl, G.; *Regionalisation of catchment model parameters*, Journal of Hydrology 287 pp 95-123, 2004

Nash, J.E., Sutcliffe, J.V.; *River flow forecasting through conceptual models*, Journal of Hydrology 10, pp282-290, 1970

Région Wallonne; *Direction générale opérationnelle de la Mobilité et des Voies hydrauliques*, http://voies-hydrauliques.wallonie.be/, last visited on October 12, 2009

Riou vzw (Vereniging Zonder Winstoogmerk); *Mosa Natura Netwerk*; http://www.riou.be/, last visited on January 12, 2010

Singh, V.P.; *Computer Models of Watershed Hydrology*, Water Resource Publications, Baton Rouge, U.S.A., ISBN 0-918334-91-8, 1995

SMHI; *Integrated hydrological modeling system (IHMS) HBV Manual Version 4.5* Swedish Hydrological and Hydrological Institute, Norrköping, Sweden, 2003

Vrugt, J.A., Braak, C.J.F. ter, Clark, M.P., Hyman, J.M., Robinson, B.A.; *Treatment of input uncertainty in hydrologic modeling: Doing hydrology backward with Markov chain Monte Carlo simulation*, Water resources research, vol. 44, 2008

Vrugt, J.A, Gupta, H., Bouten, W., Sorooshian, S.; *A Shuffled Complex Evolution Metropolis algorithm for optimization and uncertainty assessment of hydrological parameters. Calibration of Watershed Models,* Water Science and Application 6, 2003a

Vrugt, J.A., Gupta, H.V., Bouten, W., Sorooshian, S.; *Shuffled complex evolution metropolis (SCEM-UA) algorithm manual version 1.0.*, University of Amsterdam, Amsterdam, the Netherlands, 2003b

Appendices

Appendix 1: Optimization procedure: SCEM-UA (Vrugt et al., 2003a)

The used method for model optimization is the SCEM-UA algorithm. This method is developed by Vrugt et al. (2003a). The SCEM-UA is global searching method which is based on the SCE-UA algorithm (Singh, 1995). Instead of using the Downhill Simplex method, an evolutionary Markov Chain Monte Carlo (MCMC) sampler is used. This means that a controlled random search is used to find the optimum set of parameter values in the parameter space. In Figure 42, a flowchart is shown which displays the different steps in the SCEM-UA algorithm. Also the name of the MATLAB implementation file is shown.

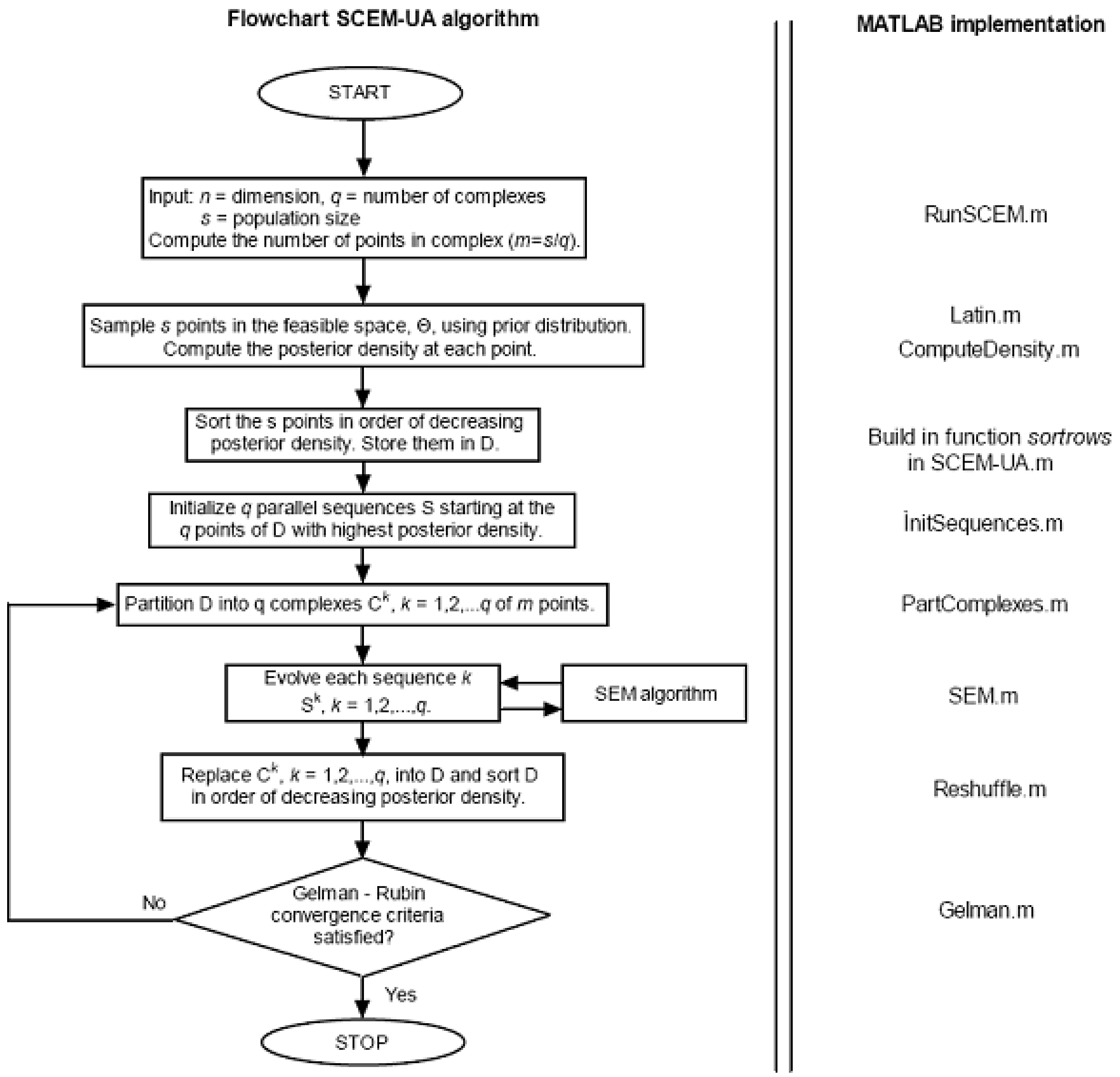


Figure 42: Flowchart of the sequential steps of the SCEM-UA algorithm (Vrugt et al., 2003b)

Before the SCEM algorithm starts, a number of choices has to be made. The number of parameters (n), the number of complexes (k), and the population size (s) have to be determined. After that, the

SCEM-UA algorithm starts with the user-specified number of random samples (s) of parameter sets (step 1). These samples are randomly placed within the specified parameter space. For the different parameter sets the posterior density, which is the value of the objective function, is determined. The s sets are ranked in a matrix D, in descending order of posterior density (step 2). Subsequently, the so called Markov Chains are initialized (step 3). This means, k independent sequences are initialized. The parameter set with the highest posterior density is sequence 1, the second best parameter set is sequence 2, and so on. After that, Matrix D is partitioned in a number of complexes (step 4). The number of points in one complex (m) is computed by dividing the population size by the number of complexes (m=s/k). The first complex contains the first m samples in matrix D. The second complex contains the samples in matrix D from 2 to m+1 and so on. Sequence 1 corresponds to the highest ranked parameter set of complex 1, sequence 2 corresponds to the highest ranked parameter set of complex 2 and so on. After this setup, the Sequence Evolution Metropolis (SEM) algorithm is started (step 5). In Figure 43 the sequential steps of this SEM algorithm are shown.

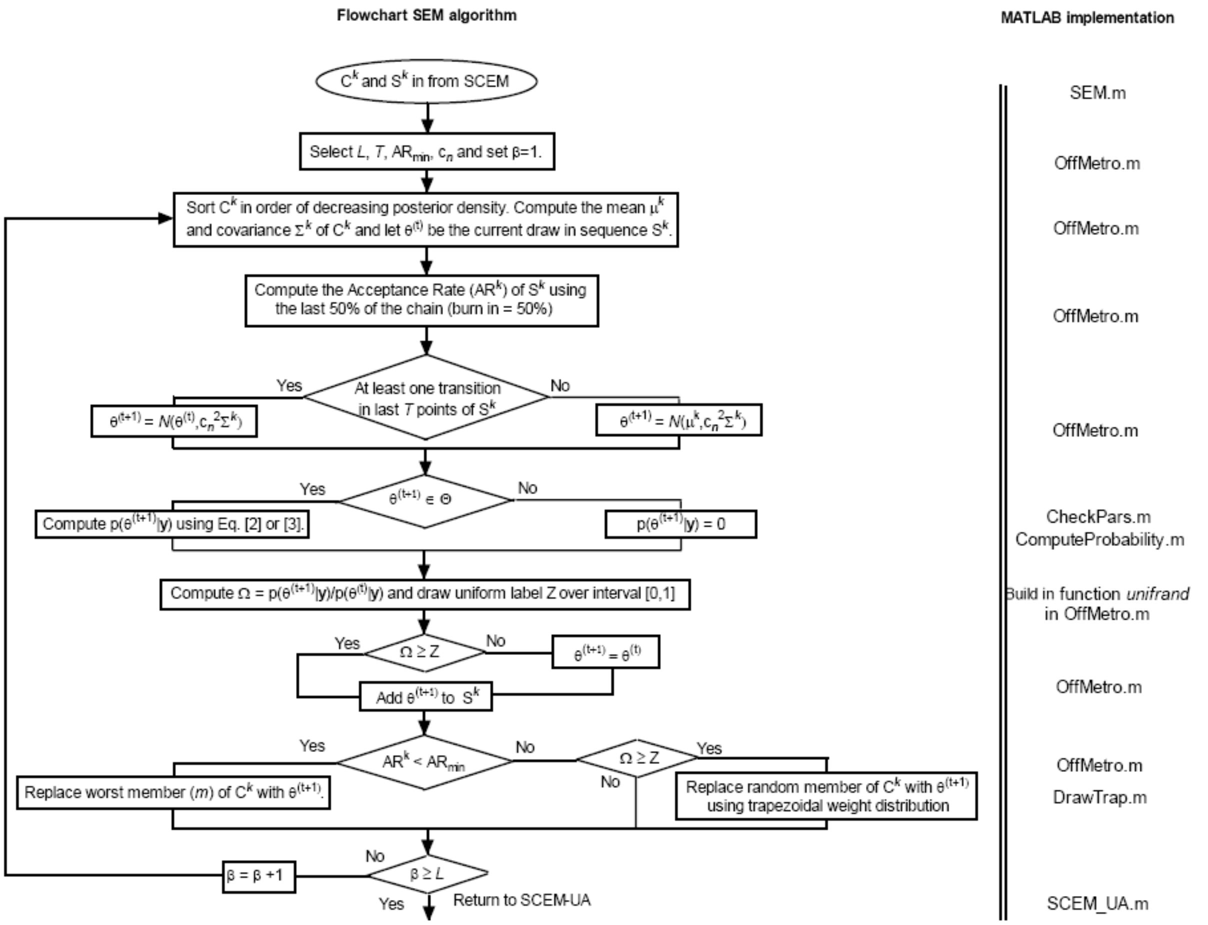


Figure 43: Flowchart of the sequential steps of the SEM algorithm (Vrugt et al., 2003b)

SEM generates and tests offspring of the parameter sets. The offspring consist of new candidate parameter sets which is derived following a certain procedure based on the existing parameter sets. Each sequence gets a new candidate parameter set, which is generated using multivariate normal distribution around the sequence parameter set of the mean of the points inside the corresponding complex. The candidate parameter set is generated using a predefined jump rate: $2.4/\sqrt{n}$ (Gelman et al., 1995). This jump rate, multiplied by the covariance of a calibration parameter in the complex and adding this to either the mean of the parameter values in the complex or the sequence parameter values, results in the offspring. Which of the two options is used, depends on whether there is a candidate point accepted over the last T points of the sequence. If a candidate parameter set is accepted, the sequence parameter set is used, else the mean of the complex. After this, the metropolis step begins in which the posterior density is calculated by running the model and computing the objective function. The parameter set is accepted and added to the sequence, if the ratio between the old posterior density and the computed posterior density is equal or higher than Z. Z is a random value between 0 and 1. If the ratio is smaller, the offspring parameter set is rejected. Z changes every start of a metropolis step.

If a parameter set is accepted, it has to replace a parameter set in the old complex. Therefore, the acceptance rate is calculated. This is done by dividing the number of accepted points in a sequence by the length of the sequence, using the last 50% of the generated points. If the acceptance rate is lower than a certain minimum value, the parameter set with the lowest posterior density is replaced. If the acceptance rate is higher than the minimum value, randomly a parameter set is replaced, using a trapezoidal probability distribution in which the best parameter set, in terms of posterior density, is has the highest chance to be replaced. After this metropolis step, all complexes are again sorted in a matrix D (step 6). In step 7, the Gelan and Rubin convergence is checked. If the convergence criteria are satisfied, the algorithm stops. Otherwise, steps 4, 5, 6 and 7 are repeated until the criteria are satisfied or until a user-specified number of iterations are done.

For the calibration, the following properties are used in the SCEM algorithm:

Table 11: SCEM-properties of the calibration

Property	***Interpretation***	***Value in calibration***
n	Dimension of the problem (number of parameters)*	8 and 5*
q	Number of complexes	5
s	Number of random samples each iteration	50
ndraw	Maximum number of iterations**	4000 and 1500**
Gamma	Kurtosis parameter Bayesian Inference Scheme	0
Option	How the model needs to interpret the model outcome and if any calculations need to be done afterwards to compute the posterior density	3 (non-informative prior)

* In the first calibration, the model was optimized using all eight parameters. After the sensitivity analysis, three parameters got a fixed value. The calibrations performed after the sensitivity analysis thereby got *n* = 5.

** In the first calibration, the model was optimized with a maximum number of iteration of 4000. After the sensitivity analysis, just five parameter had to be determined instead of eight. As a result of that, the optimum is found faster with five parameters compared to eight parameters. That is why the maximum number of iterations for all following calibrations is set to 1500.

Appendix 2: Sensitivity analyses

In this appendix chapter a sensitivity analysis is performed for both sub basins. In these analyses each parameter is varied one at a time, while the other seven are kept constant on its original value after the first calibration. The variations of the parameters influence the model in a way which results in a change in objective function. In the figure, scaled values of the parameters and the objective function are shown. A value of 1 indicates that the value is equal to the value of the original parameter or objective function.

Ourthe

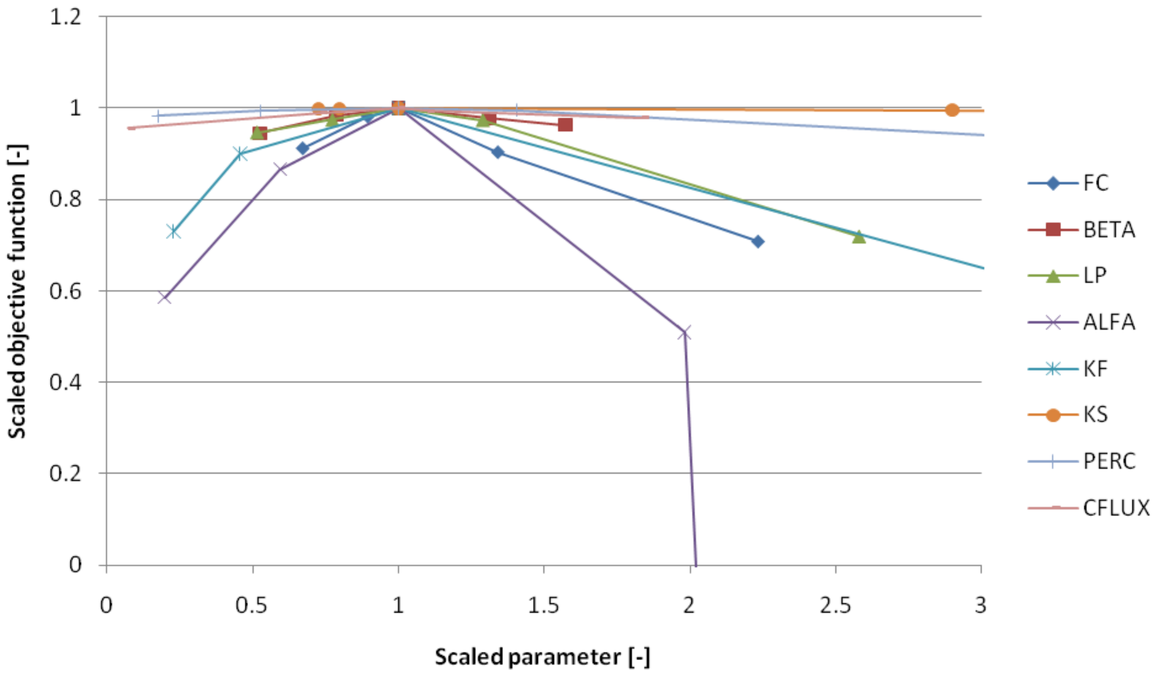


Figure 44: Sensitivity analysis Ourthe

Figure 44 shows the outcome of the sensitivity analysis of the Ourthe basin. It shows which parameters have variations in objective function if the value of the parameter is varied. If the graph has a steep slope, the parameter is sensitive on the objective function, because a small variation in the parameter value results in a relatively large deviation of the objective function. The most sensitive parameters are ALFA, FC, LP, BETA and KF. These parameters are chosen to optimize the calibration. CFLUX, PERC and KS are not sensitive and will get a fixed value in the calibration.

Chiers

In Figure 45 the outcome of the sensitivity analysis of the Chiers Basin is shown. The most sensitive parameters in this case are again BETA and FC. Less sensitive parameters are LP, KF and ALFA. CFLUX, PERC and KS are the least sensitive parameters. The parameters FC, BETA, LP, KF and ALFA are chosen for the calibration of the Chiers Basin. The other three parameters will get the a fixed value which was a result of the first calibration.

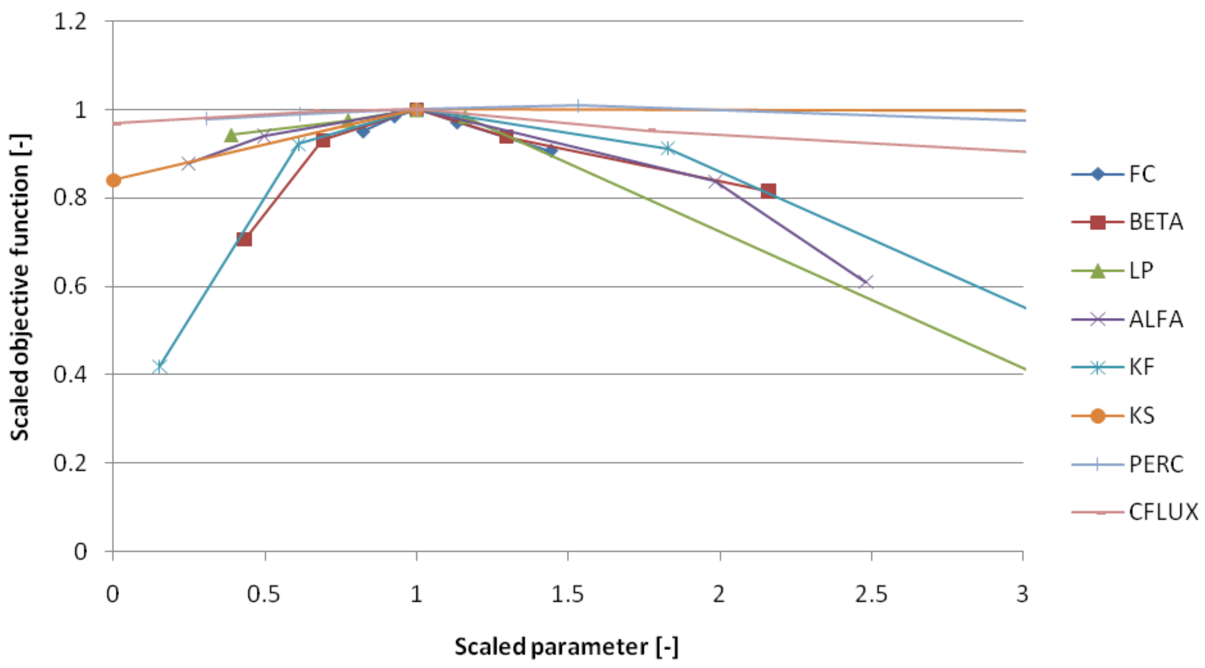


Figure 45: Sensitivity analysis Chiers

In a study of Booij and Krol (2009) the three parameters *ALFA*, *FC* and *LP* are considered the most identifiable for the Ourthe Basin and the Chiers Basin. This means that these parameters are most sensitive to a certain combined objective function in which the single objective functions *NS*, *RVE*, NS_L NS_H were included. In this research, the next most sensitive parameters in the two basins are *BETA* and *KF*. In the foregoing sensitivity analyses, *ALFA*, *FC*, *LP*, *BETA* and *KF* are also the most sensitive parameters.

Appendix 3: Uncertainties in parameter values

In the following figures, the uncertainty in model parameters after different calibrations of the realistic scenario is shown. In section 4.6.4 the uncertainties in model parameters *KF* and *ALFA* is shown. The other parameters do not show such behavior but are more certain. That can be concluded from the following figures. Figure 46, Figure 47 and Figure 48 show the uncertainties of the parameters *FC*, *BETA* and *LP* in the Ourthe basin. Figure 49, Figure 50 and Figure 51 illustrate how big the uncertainties of these parameters are in the Chiers basin. Overall it can be concluded that the parameters *KF* and *ALFA* are more uncertain after 1500 iterations compared to *FC*, *BETA* and *LP*.

Ourthe

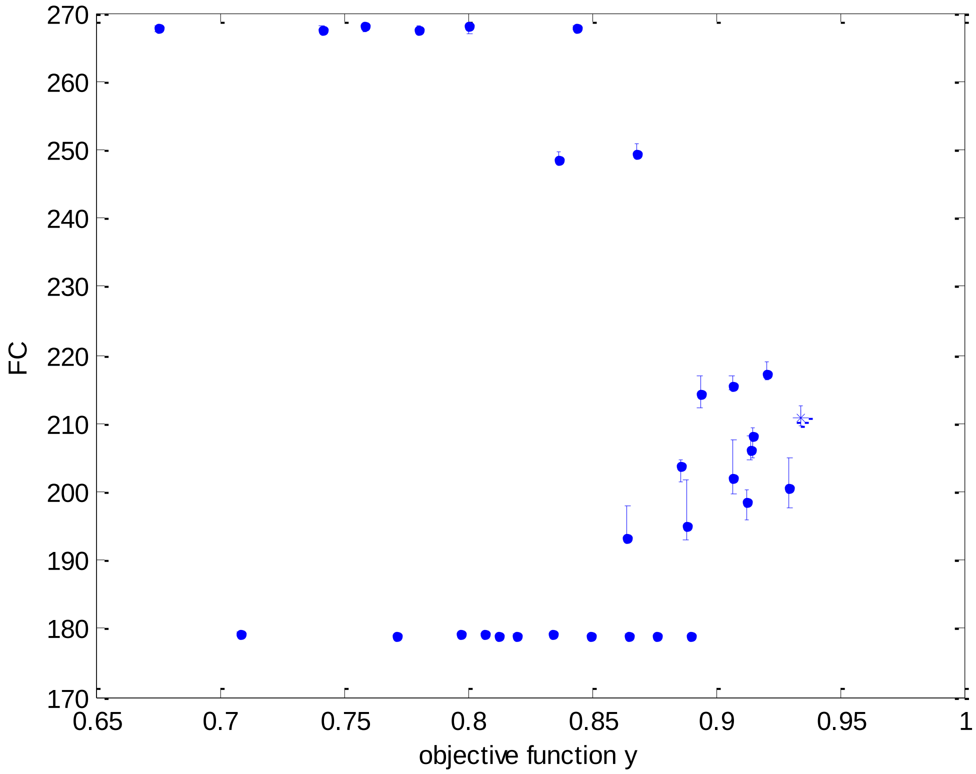


Figure 46: Values for FC in realistic scenario with 95% confidence interval error bars, Ourthe basin

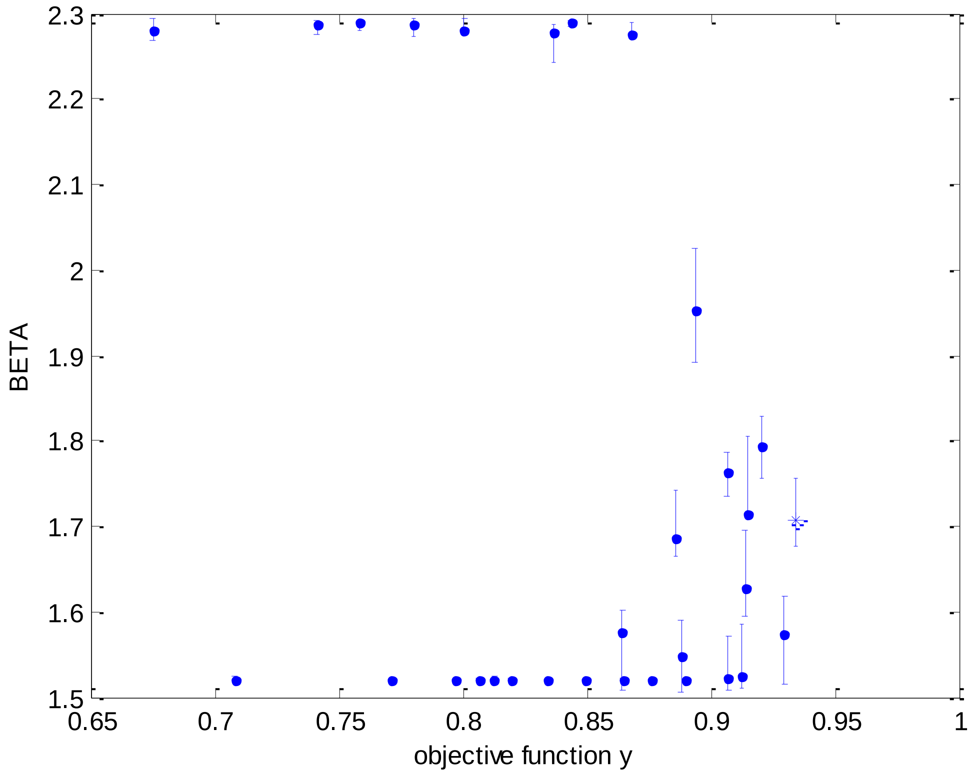


Figure 47: Values for BETA in realistic scenario with 95% confidence interval error bars, Ourthe basin

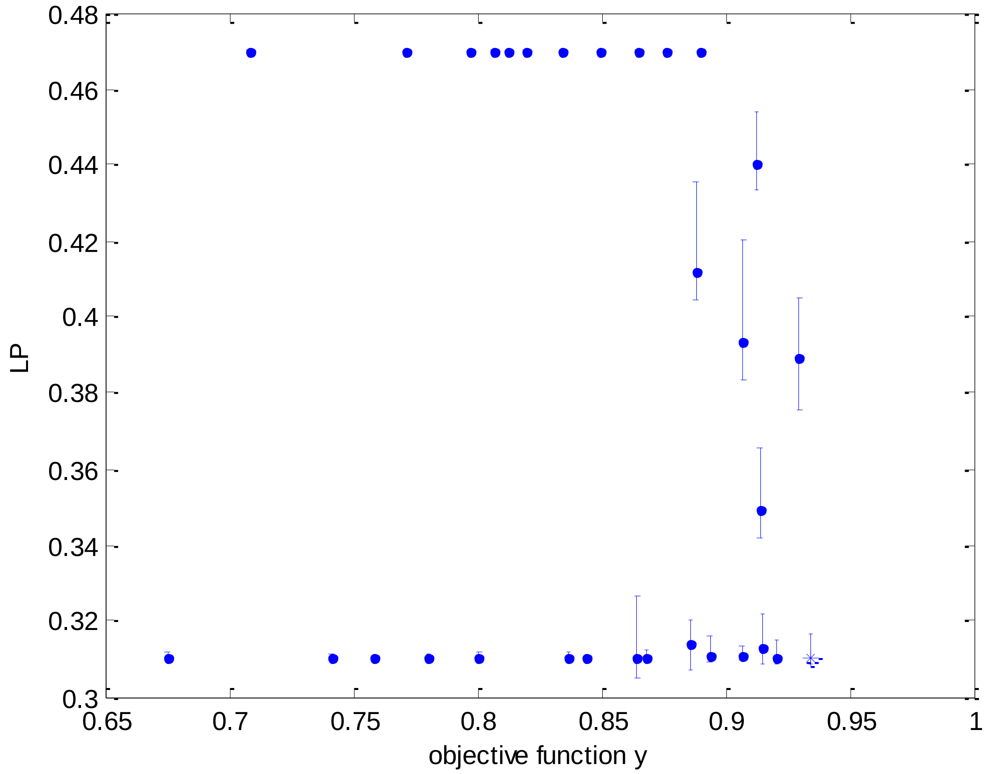


Figure 48: Values for LP in realistic scenario with 95% confidence interval error bars, Ourthe basin

Chiers

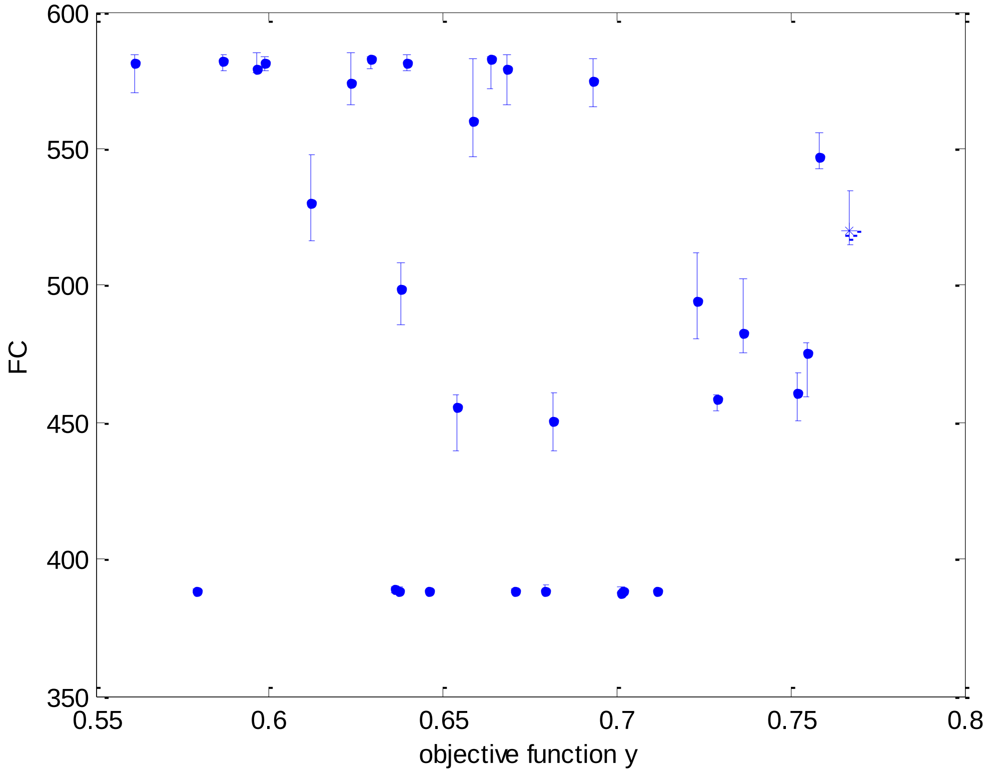


Figure 49: Values for FC in realistic scenario with 95% confidence interval error bars, Chiers basin

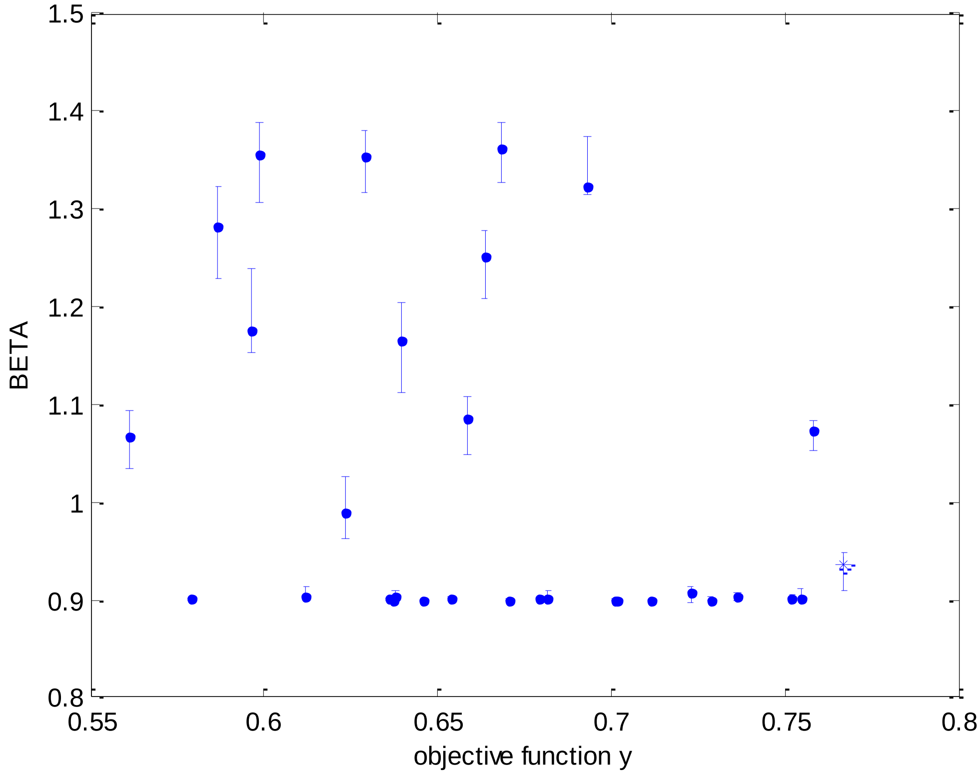


Figure 50: Values for BETA in realistic scenario with 95% confidence interval error bars, Chiers basin

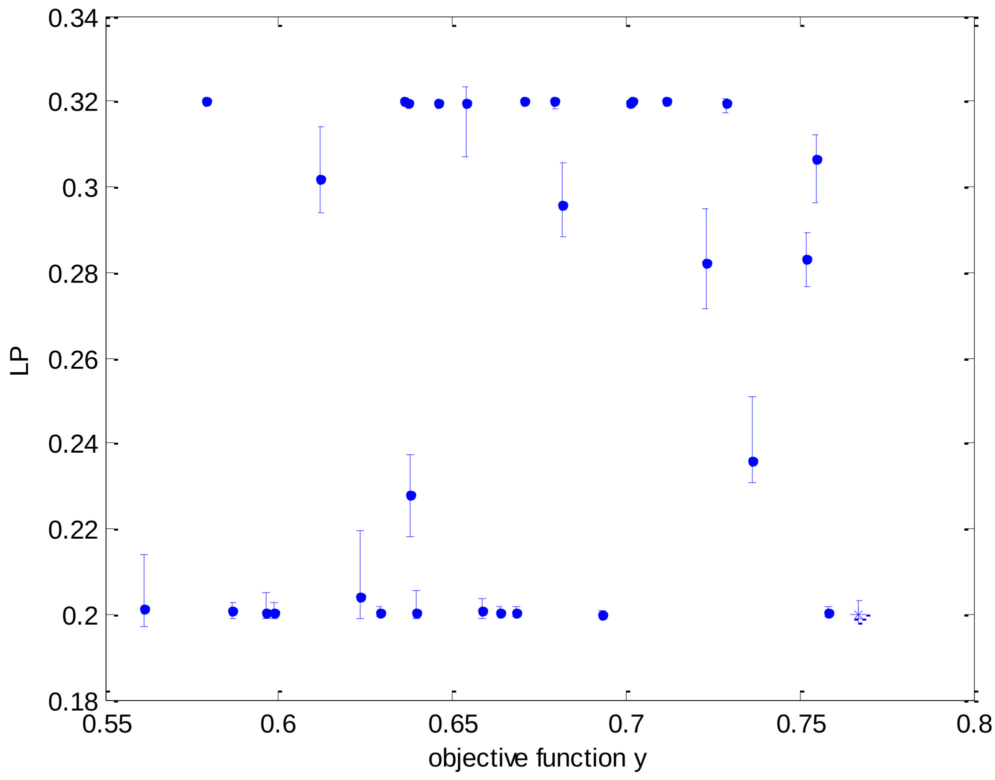


Figure 51: Values for LP in realistic scenario with 95% confidence interval error bars, Chiers basin