Assessing the influences of different hydrological models on flood magnitudes within a discharge generator

Master Thesis

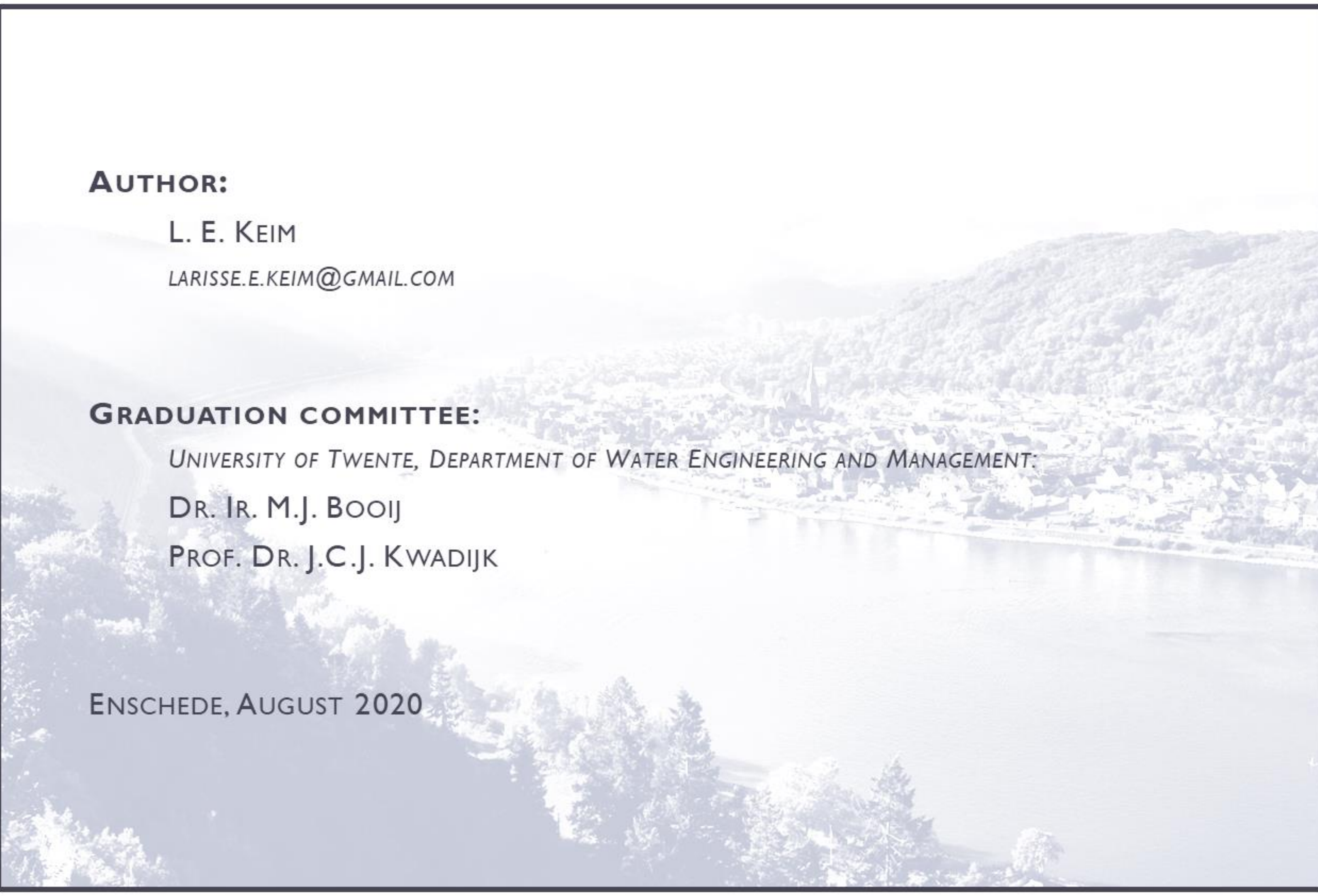
L. E. Keim

Civil Water Engineering and Management

University of Twente.

Assessing the influences of different hydrological models on flood magnitudes within a discharge generator

Master Thesis



Author:

L. E. Keim

larisse.e.keim@gmail.com

Graduation committee:

University of Twente, Department of Water Engineering and Management:

Dr. Ir. M.J. Booij

Prof. Dr. J.C.J. Kwadijk

Enschede, August 2020

Civil Water Engineering and Management

University of Twente.

SUMMARY

Until recently, the design discharges for the river Rhine were based on historical discharge data, with the return period of low probability high discharges based on a statistical analysis of a limited number of measured extreme events. Current research focuses on the use of resampled weather data, generated by a weather generator, as input for a more robust hydrological simulation. This approach is also followed in this study, where annual peak discharges for a 50.000-year synthetic data series simulated by two hydrological models are compared. This in order to assess whether the choice of a hydrological model within the discharge generator affects the annual maximum discharges of the Moselle basin, and therefore predicting different extreme discharges at large return periods.

First, a fifteen-year historical series was used to calibrate and validate the hydrological models HBV and GR4J using an automatic calibration method: SCEM-UA. This calibration method optimizes for objective function y, which combines the Nash-Sutcliffe coefficient (NS) and Relative Volume Error (RVE) metrics. It drives the model to simulate the high peaks, as well as the low flows correctly. The schematized Moselle area consists of 26 subcatchments, of which 21 contain a discharge station. The five remaining subcatchments use median parameter values as a substitute for calibration.

During calibration it was found that these five uncalibrated subcatchments influence optimal parameter sets of downstream catchments. These still perform well during validation, but the optimal parameter sets were found close to the limits of the parameter ranges. The model tends to compensate downstream by selecting extreme parameter values that are not realistic for the sub-catchment. Overall, GR4J outperformed HBV, particularly on the NS metric.

The last step in the process was combining the synthetic climate data with the previously calibrated hydrological models. The annual peaks of the 50.000 years series are visualized in flood frequency curves. The main finding here is that for areas with high quality data and no up-stream uncalibrated subcatchments, both GR4J and HBV roughly follow the same curve (figure 1). For catchments where data quality was an issue, with uncalibrated upstream subcatchments, results of GR4J and HBV deviated from each other significantly. Specifically, HBV shows more extreme discharges in the catchments downstream of uncalibrated areas (figure 2).

In conclusion, both GR4J and HBV perform well regarding observed climate data in the calibration and validation. When combined with the weather generator's synthetic series, the HBV model shows a sensitivity for uncalibrated upstream areas. The discharge generator could benefit from including multiple hydrological models, especially in areas with scarce data. Based on this research, HBV appears to be more susceptible for incomplete data than GR4J, but further research should confirm.

PREFACE

As this is a public version of the report, the personal preface has been removed.

Contents

CHAPTER 1

INTRODUCTION

Humankind's fight against water is one of all times. Even our ancient predecessors had already some form of defence against high water levels, for instance the terps, to prevent immediate drowning. In more recent times, water defence focuses more on preventative measures, keeping all our lands dry, either from the sea or peak discharges in our rivers.

The Netherlands is a delta of some of the larger rivers in Europe. Particularly the Rhine and Meuse disembogue into the North Sea after running through our 'low lands'. As such, our country always has been at risk of river floods. Its inhabitants have been working together in this fight against the water for ages; our water boards, in charge of local water defence, have been founded over 750 years ago.

Until recently, the design discharges for the river Rhine were based on historical discharge data. The return period of low probability high discharges is based on a statistical analysis of a limited amount of measured extreme events. Current research focuses on the use of resampled weather data, generated by a weather generator, as input for a more robust hydrological simulation, as compared to pure statistical analysis on historical discharges. One of the advantages of this modelling method is that the effect of upstream flooding can be accounted for. A limitation is that the most extreme flood magnitudes predicted by GRADE cannot be validated directly against measured data. (Hegnauer et al., 2014).

This research applies this approach as well; a generator of rainfall and discharge extremes (GRADE) is used to estimate the frequency of extreme discharges in the Netherlands, for the Rhine basin. The outcome of this discharge generator is used to design flood defences in the Netherlands.

GRADE is composed of three components: the Royal Dutch Meteorological Institute (KNMI) rainfall generator, a HBV model for precipitation to discharge simulations, and an 1D SOBEK schematization for hydrodynamic simulations. Hegnauer et al. (2014) assess uncertainties within the GRADE instrument in multiple areas of the instrument. The components, as well as the overall uncertainty in the GRADE simulations are quantified. Although thorough assessment of uncertainties is already done in multiple components, the impact of uncertainties as a result of the hydrological model structure has not yet been comprehensively explored.

Prior research shows various examples of a weather generator combined with a hydrological model (e.g.; Chen et al., 2010; Hegnauer et al., 2014; Falter et al., 2015; Khalili et al., 2011). However, a comparison study between different hydrological models, in combination with a stochastic weather generator to forecast extreme flood discharges, has not been extensively covered in literature yet at the onset of this research. Exactly this gap will be the focus of this research.

Goal of the research is to assess whether the choice of hydrological model within the discharge generator affects the annual maximum discharges of the Moselle basin, and therefore predicting different extreme discharges at certain return periods. To reach this goal, the following research questions were formulated.

1. *Which hydrological model shows the best performance after calibration/validation, regarding observed climate data?*

   In order to combine the hydrological models with the weather generator in the next step of this research, the hydrological models first have to be calibrated and validated. The parameters within the models **should** be set to simulate the discharge correctly based on the 15 years of available climate data for the Moselle catchment.

2. *What influence do different hydrological models combined with the weather generator have on the flood magnitudes?*

   The second part of this research will combine the weather generator with the previously calibrated hydrological models. 50 000 Years of synthetic climate data, resampled by the weather generator, will be used as input of the models. Then the models will be compared using the annual maxima, assessing the difference between flood magnitudes at their respective return periods.

Two different rainfall runoff models will be selected and compared using such a stochastic weather generator for the Mosel catchment. The two hydrological models used in this research are HBV, GR4J. The models will be calibrated using the SCEM-UA algorithm (Vrugt et al., 2003).

The outcomes of this research may lead to an improved insight in extreme discharge events, as the impact of hydrological model choice will be quantified. The remainder of this thesis is structured as follows. Chapter two describes the study area, available historical data, and introduces the discharge generator and the hydrological models applied. Chapter three focusses on the applied methodology, of which the results are described and discussed in chapter four. This thesis concludes with conclusions and recommendations in chapter five.

CHAPTER 2

STUDY AREA AND DATA

In this chapter background information of this research is given. Section 2.1 gives an overview of the study area. The historic climate data used will be described in section 2.2, and the synthetic data from the discharge generator is covered in the last part.

2.1 STUDY AREA

The Netherlands is located in a delta and is partly located below sea level. Large European rivers, such as the Rhine and Meuse, disembogue into the North Sea after running through the 'low lands'. Within this delta, nearly half of the population of the Netherlands is situated between Rotterdam and Amsterdam, which is predominantly located below sea level.

The study area, the Moselle, is one of the fifteen major sub-basins of the Rhine river. For the most part it is located within Germany, smaller upstream catchments of the Moselle are shared with neighbouring countries France and Luxembourg (figure 2.1). The Moselle was named after the large nearby river the Meuse. Mosella means the 'little Meuse'. In the Netherlands it is referred as 'de Moezel', and in Germany and France it is called 'die Mosel' and 'la Moselle'. The Moselle springs in the Vosges mountains, located in France near the German border, and meanders between German cities. The tributary connects to the main Rhine channel in Koblenz.

The catchment size of the Moselle is 29000 km^2, the length of the river is 545 km, and the average discharge measures 325 m^3/s. The river Moselle contains various barriers influencing its flow characteristics; the river is regulated by locks, dams, and weirs.

The Moselle was selected for this research for the following reasons:

- Quality of data is adequate, e.g., 80% of the schematized subcatchments have a discharge station.
- The area is precipitation dominated. This allows for benchmarking model response on precipitation.
- Snow is of low impact; not all used models include a snow routine.
- Characteristics of the Moselle river are akin to the Meuse. Results of this research may be applicable to the Meuse are as well, and studies for both areas can be compared for both historical and synthetic time series.

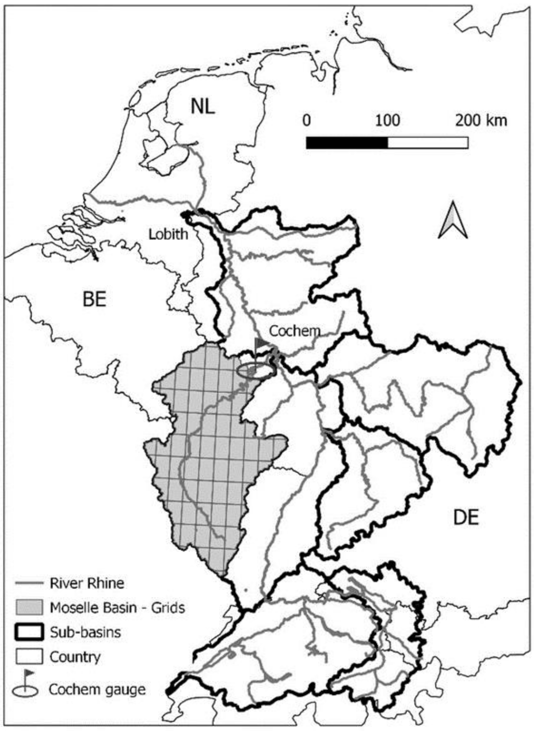


FIGURE 2.1 - MOSELLE CATCHMENT VISUALIZED WITHIN THE RHINE CATCHMENT (DEMIREL ET AL., 2019)

2.2 AVAILABLE OBSERVED CLIMATE DATA

2.2.1 AVAILABLE DATA SETS

The datasets used for this research are described by Winsemius et al. (2013). Observed daily precipitation and temperature data (variables P, T), as shown in table 2.1 from the German Bundesanstalt fur Gewasserkunde (BfG) was retrieved via Deltares. These observations entail the period 1951 - 2006. Potential evapotranspiration (variable PET) is retrieved from Deltares, based on deviations from long term average temperatures and evaporation profiles, as shown in equation 2.1. This is the default approach in HBV model.

Daily discharge measurements (variable Q) is from a collection of corrected data series from the HYMOG (Hydrologische Modellierungsgrundlagen im Rheingebiet), for the period 1990-2007. (Steinrücke, 2012)

From available data sets, year 1990-2006 will be the period available for calibration and validation purposes, as it is the overlapping timeframe between the four described data sets. Data sets are summarized in table 2.1 below.

TABLE 2.1 - OVERVIEW OF AVAILABLE OBSERVED DATA

Variable	Name	Number of stations/ sub-basins	Period	Spatial resolution
Q	Discharge	21	1990-2007	Point
P	Precipitation	26	1951-2006	Basin average
T	Temperature	26	1951-2006	Basin average
PET	Potential EvapoTranspiration	26	1951-2006	Basin average

2.2.2 Data clean-up

Provided data for variables P, T, PET was fully available, without gaps, and used as-is. For this study, measured discharge data (variable: Q) was used of 21 different sub-catchments, from the period 1990 through 2007. For a majority of these discharge stations, data was mostly complete, however some areas (i.e., Alzette, Prims) had significant periods with data missing.

Provided data has an hourly resolution, which was aggregated to daily values for the use in this study. Alongside this process, various clean-up activities have taken place to ensure only valid data points are fed into the model. These are described below.

Missing data points

By carefully evaluating the available data, some patterns were recognized. Some sets would intermittently have a few hours of data missing on various day. As data is aggregated from hourly to daily resolution, some missing datapoints would be acceptable. If at least half of the reported hours in a given day were available, the data point would be reported as the average of the available data points. If less than half of the data points would be available, the full day would be reported as missing, and thus not taken into account for calibration and validation.

Irregular data patterns

Another pattern was found in some of the data sets: repeating sequences of data showing the same value, up to weeks in a row. This could indicate a form of interpolation in the data, and/or measurement errors.

Another pattern that was found in the data were sets of steadily increasing/decreasing series, again for weeks in a row. This could again be indicating interpolation and/or measurement errors. As a smaller, but more robust data set would be preferable over a slightly larger one with suspect data series, filtering of irregular data patterns is performed.

To ensure suspect data series / interpolations would not impact model calibration/verification, a set of rules was chosen to cut off these long sequences. After some trials the following ruleset was chosen. A series would be suspect if for a period of more than 72hrs in a row, the data would not, or not significantly (<1%) change between hourly data points. After 72hrs of no significant change, the remaining values would be invalidated, until change did appear.

First 72 hours would remain in the data set. A risk with this approach would be that low flow periods might have the tendency to show exactly this behaviour and could be impacted by this rule. Hence, impact of this rule on the data sets was checked; it correctly cut off the manually identified suspicious series of data, while not impacting the regular low flow periods.

2.3 Synthetic climate data

Next to calibration and validation based on measured data, the second part of this research focuses on applying the models using a synthetic climatic time series as input. This paragraph describes this generated set, prepared by the KNMI. The weather generator uses historical data from the 56-year period 1951-2006 to simulate a daily weather series of 20 000 – 50 000 years (Hegnauer, 2014b).

The weather generator 'randomly' selects a one-day rainfall event from the available historical data. The choice is restricted to a 121-day window around the calendar day of interest, to ensure the seasonal variation within the simulated series. The selection of nearest neighbors will logically not lead to different one-day rainfall amounts from the historical data; however, the multi-day rainfall amounts can lead to not previously observed events (figure 2.2).

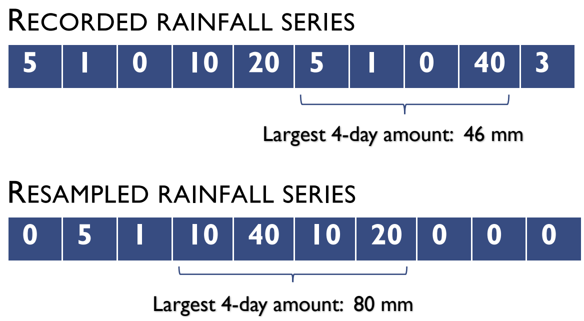


FIGURE 2.2 - SCHEMATIC REPRESENTATION OF RESAMPLING (HEGNAUER, 2014B)

The dataset contains the variables precipitation and temperature. Potential evapotranspiration was not supplied. Hence, this PET has been derived from the temperature parameter using the HBV-standard approach set with a correction factor equal to 0.1 $^{\circ}C^{-1}$ (Siebert, 2005).

$$PET(t) = \left(1 + C_{ET}(T(t) - \bar{T})\right)\overline{PET} \qquad \text{while: } 2\,\overline{PET} \geq PET(t) \geq 0 \qquad [2.1]$$

TABLE 2.2 - VARIABLE DESCRIPTION OF FORMULA 2.1

Variable	*Description*	*Unit*
$\boldsymbol{PET(t)}$	Potential evaporation at day t	[mm/day]
$\boldsymbol{\overline{PET}}$	Long-term mean potential evaporation for this day of the year	[mm/day]
$\boldsymbol{C_{ET}}$	Correction factor	[1/°C]
$\boldsymbol{T(t)}$	Temperature at day t	[°C]
$\boldsymbol{\bar{T}}$	Long-term mean temperature for this day of the year	[°C]

CHAPTER 3

METHODS

An in-depth overview of the methodology is given in this chapter, starting with an overview of the model structures. This is followed by the schematization of the study area, after which the calibration method will be discussed. Then, the implementation and validation of the chosen calibrated models within the discharge generator are explained. The chapter concludes with a validation of the synthetic data.

3.1 MODEL STRUCTURES

Hydrological models are used in practice for different aims: to improve the fundamental understanding of existing hydrological systems and assessing the impact of change on water resources, to develop new models or improve old models for management decisions on current and future catchment hydrology, and to extrapolate point measurements in both space and time (Pechlivanidis et al., 2011; Singh and Woolhiser, 2002).

In this research, the second aim is mainly of relevance; management decisions are reliant on results from the hydrological model. The design discharges and associated flood hydrographs for the rivers Rhine and Meuse will be based on the outcome of the discharge generator, of which the hydrological model is an important component (appendix A).

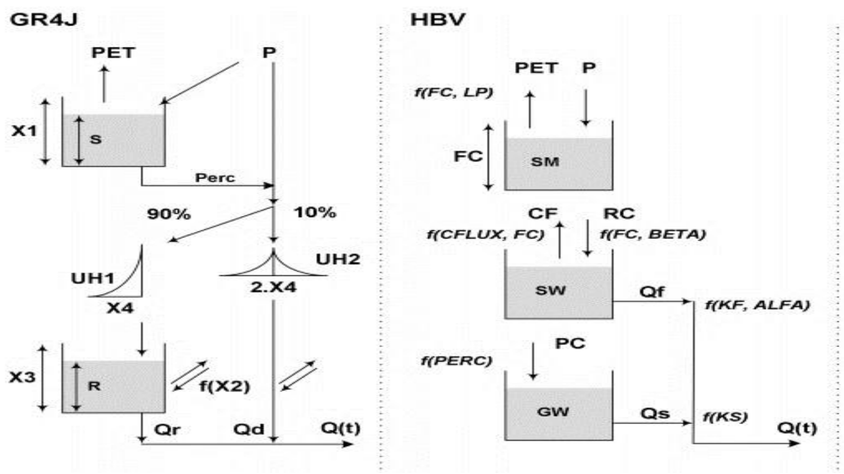


FIGURE 3.2 - SCHEMATIZATION OF MODEL STRUCTURES (DEMIREL ET AL, 2015)

To select fit for purpose hydrological models for this study, fitness requirements need to be set and benchmarked for the various models available in the field. For this purpose, an approach by Pechlivanidis et al. (2011) was chosen.

Pechlivanidis classifies models by five criteria: model structure, spatial distribution, stochasticity, temporal application, and spatial application. Based on provided climate input from GRADE and the already schematized study area, the classifications of a hydrological model fitting this research should be: conceptual, lumped, deterministic, continuous (daily), with respect to large catchment sizes, as this aligns with the selected study area.

In this paragraph, the two hydrological models used during the execution of the master thesis are described: HBV and GR4J (figure 3.1). Selection of these models is elaborated upon in 3.2.1. and 3.2.2.

3.2.1 GR4J

Génie Rural à 4 paramètres Journalier (GR4J), based on GR3J, is a continuous lumped rainfall-runoff model developed by Perrin (2003). The GR4J model is simple, with its limited four parameters (table 3.1). A promising model, with good model predictions regarding high peak flows (Zhang et al., 2015). GR4J uses a daily time-step (Journalier/Daily) and uses precipitation and potential evapotranspiration as input.

In a model comparison report by Pagano et al. (2010) the GR4J model outperformed every other model considered. Even though GR4J only has 4 tunable parameters, fewer than most rainfall-runoff models. Pagano even suggested that the model is capturing realistic hydrological behavior. In a study by van Esse (2012) GR4H is compared with 12 SUPERFLEX structures, conceptual hydrological models with different complexities, where GR4H performs best in most of the 237 French catchments.

3.2.2 HBV

HBV is a Swedish hydrological model, currently implemented within the discharge generator. The Hydrologiska Byråns Vattenbalansavdelning (HBV) model was developed by the Swedish Meteorological and Hydrological Institute. The model consists of four routines: precipitation routine, soil moisture routine, runoff routine, and a routing routine (Lindström et al., 1997). Input variables of this model are precipitation, temperature and potential evapotranspiration. Since the HBV model is currently implemented within GRADE it is an obvious choice to select this model as one of the two hydrological models to compare.

TABLE 3.1 – PARAMETER DESCRIPTIONS OF THE SELECTED MODELS

Parameter	*Description*	*Unit*
	GR4J	
x_1	Maximum capacity of the production storage	[mm]
x_2	Groundwater exchange coefficient	[mm]
x_3	One day ahead maximum capacity of routing store	[mm]
x_4	Time base of unit hydrograph	[day]
	HBV	
FC	Maximum soil moisture content	[mm]
β	Parameter in soil routine	[-]
LP	Limit for potential evapotranspiration	[-]
α	Response box parameter	[-]
Kf	Recession coefficient quick flow	[1/day]
Ks	Recession coefficient base flow	[1/day]
Perc	Percolation from upper to lower response box	[mm/day]
Cflux	Maximum value of capillary flow	[mm/day]

3.2 SCHEMATIZATION OF THE STUDY AREA

To be able to model all sub-catchment areas in the correct sequence, a map has been prepared detailing the consecution and properties for each of the 26 catchments. As a basis, the existing map from previous research by Winsemius et al. (2013) was used. Based on the actual geographical locations of the catchments (figure 3.2), the succession between areas is represented in the flow chart prepared for this study (figure 3.3). As compared to the previous study for this discharge generator, additional discharge stations are available for this study, allowing for six more catchments to be included for calibration.

Subcatchments without a discharge station will have their model parameters set to a median of previous calibrated catchments. The subcatchments are linked by a simple lag function. This lag function uses the distances between discharge stations and an average flow velocity of 2 m/s for the whole system. This value is partly chosen after evaluating flood hydrographs of multiple calibration runs with different flow velocities to match peaks of extreme events.

Catchment	Area [km^2]
Alzette	1204
Blies	1909
Kyll	816
Lieser	375
Nied	1368
Nims	241
Obsa	1854
Omos1	3353
Omos2	2912
Omos3	1852
Omos4	822
Orne	1272
Our	606
Prims	741
Pruem	602
Rest1	447
Ruwer	101
Sauer1	487
Sauer2	233
Seille	1287
Sure	947
Umos1	571
Umos2	1224
Umos3	812
Umos4	1051
Unsaar	1127

FIGURE 3.2 - SUBCATCHMENTS OF THE MOSELLE WITH CORRESPONDING DISCHARGE STATIONS AND RESPECTIVE SIZE

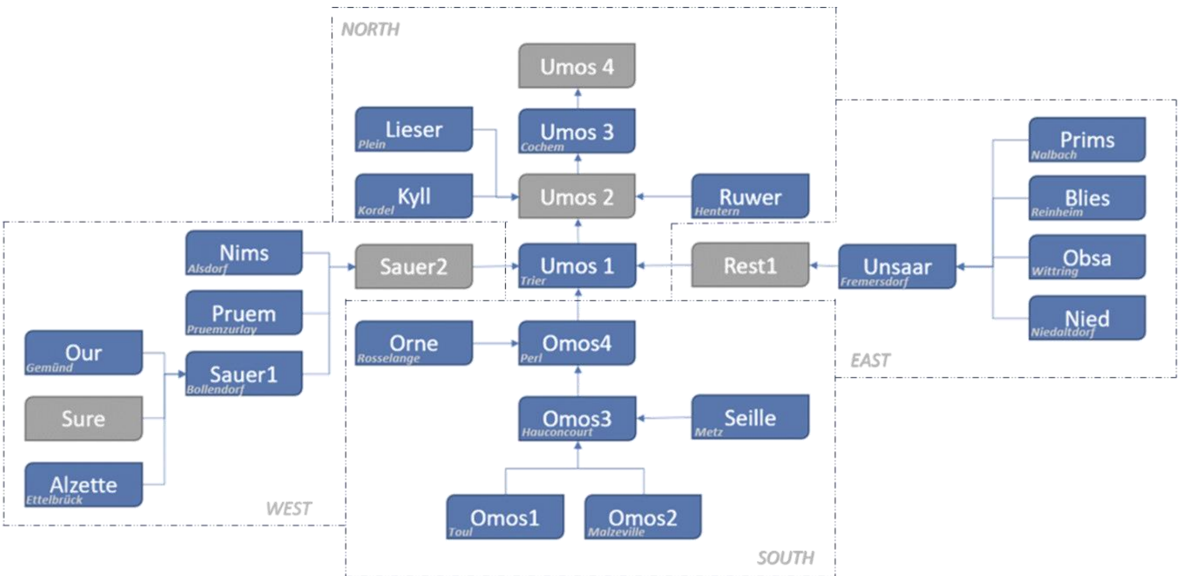


FIGURE 3.3 - FLOW CHART OF THE STUDY AREA, GREY SUBCATCHMENTS DO NOT HAVE A DISCHARGE STATION

3.3 CALIBRATION

Before automated methods for model calibration were introduced, one would manually adjust the parameter values to find a good fit. Manual adjustment of parameter values by the modeler is time consuming, and the chance of finding an optimal parameter set is small. Dawdy and O'Donnel (1965) reported the first steps of automatic calibration, and automatic calibration methods started to improve. Still, finding an optimal global parameter set would remain to be difficult (e.g., Duan et al., 1992).

One validates the model with other data to conclude if the found parameter set is indeed a good fit to simulate the catchment and predict correct discharges. The data used in this research is split in half for validation and calibration. A limitation of the generator of rainfall and discharge extremes as stated by Hegnauer et al. (2014) is that the hydrological model simulates discharges far above the highest recorded discharge. Therefore, for calibration the time period of 1996-2006 was chosen, and for validation 1990-1995. The years 1994 and 1995 have 2 of the highest peaks in recorded history. Some data was found to be missing from the data set, below figure 3.4 shows the proportion of data available for calibration/validation per subcatchment. The goal of this research is to simulate even higher discharges than historically measured, so the calibration and validation process should fit this purpose.

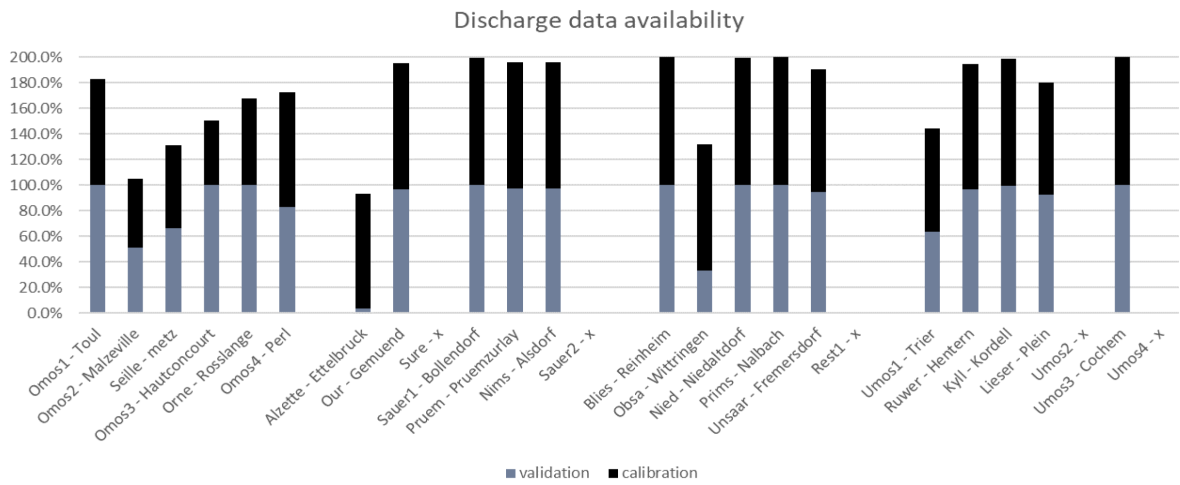


FIGURE 3.4 - PROPORTION OF DATA AVAILABLE PER SUBCATCHMENT FOR CALIBRATION/VALIDATION

3.3.1 SCEM-UA

According to Vrugt et al. (2003) SCE-UA by Duan et al. (1992) is a powerful robust and efficient global optimization procedure. "*The SCE-UA algorithm is consistent, effective and efficient in locating the optimal model parameters of hydrological model.*" But still, the algorithm by Duan has difficulties with finding a unique 'best' parameter set.

SCEM-UA is a combination of SCE-UA and the Metropolis-Hastings algorithm. The Metropolis algorithm is the basis of classical MC^2 methods, it will rapidly explore the parameter space. However, when a proposal distribution is poorly chosen, it will slowly converge, and a limited distribution will be found. Hydrological models lack often a priori knowledge, which limits the MC^2 samplers. The SCEM-UA algorithm merges the strengths of the Metropolis-Hastings algorithm, controlled random search, competitive evolution, and complex shuffling. (Vrugt et al., 2003)

The SECM optimization can be tweaked by the modeler to fit the study area and hydrological model. The variable values chosen during this procedure are shown in table 3.2. Variables q and s are chosen based on the recommendations by Vrugt et al. Variable ndraw was chosen at 5000 as early runs of the model converged well within 5000 iterations. A more detailed description of the SCEM-UA method can be found in Vrugt et al. (2003).

TABLE 3.2 - SCEM VARIABLES

Variable	*Description*	*Value*
n	Number of parameters to be optimized in the hydrological model	8 (HBV) - 4 (GR4J)
q	Number of complexes	5
s	Number of random samples	100
ndraw	Number of iterations	5000

3.3.2 OBJECTIVE FUNCTION

While the focus of this research is on low probability high discharges, it is expedient to also choose parameters that simulate the whole system, including low flows. To ensure a representative hydrological model with, i.a., a correct water balance.

Different objective functions were investigated. Multiple of these functions caused equifinality, when many sets of parameter values give similar results after calibration, and therefore would not converge speedily, or not converge at all, to a global optimal parameter set.

Previous research by Akhtar et al. (2009) suggested a new objective function 'y' (formula 3.1); a combination of the Nash-Sutcliffe coefficient (NS) (1970) and the Relative Volume Error (RVE). Research by van den Tillaart et al. (2013) implemented this objective function in combination with SCEM-UA and this resulted in an effective calibration. The NS gives an indication of the overall performance of the model, given formula 3.2 where Q and Q_s are the observed and simulated discharge at a given time (t) [m^3/s]. The NS is situated between 1 and $-\infty$, where a score of one characterizes a perfect model. The RVE can score between -1 and 1, where a score of zero indicates no volume error. Thus, the objective function y scores a one for a flawless model.

Although this research focusses on low probability extreme high discharges, its objective function does not solely. The objective function y drives the model to simulate the high peaks, as well as the low flows correctly. This to ensure that the hydrological models portray the whole discharge series as good as possible.

$$y = \frac{NS}{1 + |RVE|} \qquad [3.1]$$

where

$$NS = 1 - [\frac{\sum[Q_s(t) - Q(t)]^2}{\sum[Q(t) - \bar{Q}]^2}] \qquad [3.2]$$

$$RVE = \frac{\sum[Q_s(t) - Q(t)]}{\sum Q(t)} \qquad [3.3]$$

3.4 VALIDATION - OBSERVED CLIMATE DATA

By purposely withholding some of the data from the optimization algorithm, an opportunity is created to benchmark (validate) the performance of the calibration algorithm. The same objective function as applied by calibration is used to score the model in the validation period. For 21 of the 26 subcatchments a discharge station was available. For the Alzette catchment there was not enough data available for validation, so only 20 subcatchments were scored

The data from the years 1990-1995 is used for the validation of the models. This period was purposely chosen, as it contains the two most extreme events in the dataset. By selecting this period for validation, some insight could be gained how well the calibrated model would cope with events more extreme than in the calibration period. This is particularly relevant for this study as in the synthetic dataset where the model is applied to as well, even more extreme events are present.

3.5 VALIDATION - SYNTHETIC CLIMATE DATA

Validation as applied on the measured climatic data is not possible for the synthetic climate data set. As this set does not contain an actual observed series of discharges, there is no data to validate against.

As an alternative, the complete series of simulated synthetic discharge has to be benchmarked using other techniques focusing on internal consistency of the series (ratio $\sum Q$ / $\sum P$ per sub-catchment), statistical properties like mean and variance of the data set, and most importantly the flood frequency curve as compared to the extrapolated curves from the observed climate data and modelled observed climate data.

CHAPTER 4

RESULTS

In this chapter the results will be presented and discussed. Starting with the calibration results including an in-depth view on the best parameter sets for all catchments, where after the simulation results of both GR4J and HBV are reviewed. The chapter concludes with the differences between both models regarding the synthetic 50.000 year series.

4.1 CALIBRATION AND VALIDATION RESULTS - OBSERVED CLIMATE DATA

In table 4.1 the parameter ranges and calibrated median parameters values for all subcatchments are shown. Most best parameter sets fit well within the parameter ranges set for the automatic SCEM calibration procedure, as visualized in detail in appendix B. However, some subcatchments, for example Umos3, converge towards the set boundaries. This can be explained due to the inflow of its upstream areas, part of those upstream areas was not calibrated and could therefore contain deviations of the actual discharge. The model tries to compensate for the errors made upstream, and hence the model turns out to find its solution in extreme parameter values that are not realistic for the models.

TABLE 4.1 - SCEM PARAMETER RANGES AND CALIBRATED MEDIAN VALUES - PARTLY BASED ON DEMIREL ET AL. (2013)

Parameter	*Model parameter ranges*		*Calibrated median value*	*Unit*
	GR4J			
x_1	0	3000	291	[mm]
x_2	-10	10	0.37	[mm]
x_3	10	500	47.3	[mm]
x_4	0.6	5	1.17	[day]
	HBV			
FC	10	1000	259	[mm]
β	1	6	1.69	[-]
LP	0.1	1	0.78	[-]
α	0.1	3	0.22	[-]
Kf	0.0005	0.3	0.15	[1/day]
Ks	0.0005	0.3	0.05	[1/day]
Perc	0.001	6	1.87	[mm/day]
Cflux	0	6	1.41	[mm/day]

Tables 4.2 and 4.3 display for GR4J and HBV respectively, parameter sets with the highest scores for the objective function. The order or the subcatchments is the same order in which they were calibrated, starting with Omos1, and ending with Umos4.

The subcatchments with its name coloured light blue are subcatchments without discharge station. Therefore, these subcatchments obtained a median parameter set of previously calibrated catchments. these are not exclusively upstream. The background of parameter values has been conditionally coloured in blue shades when the parameter value is close to the set boundary as mentioned in table 4.1. It has no colour when it is close to the median value, therefore the uncalibrated catchments show no colouring in this table. This does not mean that parameters for these areas are a good fit; they are just set to these values as a best estimate by lack of data to calibrate for. An interesting effect surfaces in subcatchments downstream of these uncalibrated catchments. These are more likely to have parameter values close to the boundaries. A similar effect is visible with catchments with a multitude of upstream catchments entering the same area (specifically: Unsaar).

These catchments that show parameters close to the boundaries seem to compensate for discharge errors in the simulation upstream. Another phenomenon is visible around succeeding catchments Umos1, Umos2, Umos3. Umos1 is already compensating for upstream areas and has its best parameter set touch many boundaries. Umos2 is uncalibrated. Umos3 in turn compensates again, but now some parameters sway to the other boundary. For GR4J this is x1, for HBV its parameters Beta, Kf, Ks, Cflux.

TABLE 4.2 - GR4J CALIBRATED PARAMETER SETS PER SUBCATCHMENT IN ORDER OF CALIBRATION

	SOUTH						*WEST*						
	Omos1	**Omos2**	**Seille**	**Omos3**	**Orne**	**Omos4**	**Alzette**	**Our**	**Sure**	**Sauer1**	**Pruem**	**Nims**	**Sauer2**
x1 [mm]	337.14	311.72	192.61	1252.74	255.85	2.90	508.39	133.55	283.79	36.37	77.53	298.05	255.85
x2 [mm]	0.63	0.52	1.00	5.70	0.46	-1.54	-0.44	0.34	0.49	1.58	-0.25	-0.17	0.46
x3 [mm]	50.31	48.30	65.94	382.52	27.98	45.26	31.14	55.13	49.31	47.91	52.53	36.45	48.30
x4 [day]	1.20	1.25	1.20	4.73	1.32	1.28	0.93	1.17	1.23	1.02	1.13	0.85	1.20
	EAST						*NORTH*						
	Blies	**Obsa**	**Nied**	**Prims**	**Unsaar**	**Rest1**	**Umos1**	**Ruwer**	**Kyll**	**Lieser**	**Umos2**	**Umos3**	**Umos4**
x1 [mm]	823.04	291.09	210.80	546.31	2212.03	294.57	1538.59	270.37	390.32	174.90	294.57	6.58	291.09
x2 [mm]	-0.24	-0.02	0.52	0.10	8.39	0.40	-10.00	0.41	-0.27	-0.65	0.22	1.24	0.34
x3 [mm]	40.93	24.51	34.79	70.07	499.69	48.11	10.00	83.36	43.51	41.41	46.59	10.00	45.26
x4 [day]	1.17	1.18	1.25	0.79	0.60	1.18	1.15	0.71	1.07	1.06	1.16	1.43	1.17

TABLE 4.3 - HBV CALIBRATED PARAMETER SETS PER SUBCATCHMENT IN ORDER OF CALIBRATION

	SOUTH						*WEST*						
	Omos1	**Omos2**	**Seille**	**Omos3**	**Orne**	**Omos4**	**Alzette**	**Our**	**Sure**	**Sauer1**	**Pruem**	**Nims**	**Sauer2**
FC [mm]	282.15	250.53	236.15	576.54	280.69	173.90	724.53	188.13	265.61	84.35	201.95	380.21	250.53
Beta [-]	1.52	1.69	2.35	2.72	3.12	1.00	1.20	2.02	1.85	1.00	1.60	2.42	1.69
LP [-]	0.71	0.68	1.00	0.48	0.95	1.00	0.43	0.80	0.76	1.00	0.62	0.78	0.78
Alfa [-]	0.10	0.21	0.10	0.11	0.10	0.63	0.41	0.25	0.16	0.74	0.27	0.23	0.23
Kf [1/day]	0.30	0.17	0.11	0.00	0.23	0.30	0.09	0.13	0.15	0.23	0.15	0.12	0.15
Ks [1/day]	0.11	0.03	0.04	0.28	0.03	0.19	0.03	0.05	0.04	0.28	0.18	0.00	0.05
PERC [mm/day]	5.99	2.56	0.75	2.00	1.00	2.24	1.37	0.98	1.69	5.91	1.04	1.04	1.37
Cflux [mm/day]	2.77	0.04	4.06	5.88	2.38	5.99	0.20	0.59	2.57	5.94	1.36	0.03	2.38
	EAST						*NORTH*						
	Blies	**Obsa**	**Nied**	**Prims**	**Unsaar**	**Rest1**	**Umos1**	**Ruwer**	**Kyll**	**Lieser**	**Umos2**	**Umos3**	**Umos4**
FC [mm]	586.84	339.16	252.28	333.88	180.88	266.49	233.69	325.99	490.13	186.97	266.49	118.15	252.28
Beta [-]	1.19	1.45	2.96	1.64	1.00	1.62	5.92	3.50	3.21	5.96	1.85	1.12	1.69
LP [-]	0.38	0.40	1.00	0.58	1.00	0.75	0.11	1.00	0.78	0.80	0.78	0.84	0.78
Alfa [-]	0.11	0.16	0.20	0.17	0.31	0.20	1.35	0.37	0.78	0.12	0.22	0.93	0.23
Kf [1/day]	0.28	0.30	0.15	0.15	0.30	0.16	0.08	0.06	0.02	0.00	0.15	0.30	0.15
Ks [1/day]	0.04	0.05	0.03	0.07	0.01	0.05	0.27	0.04	0.13	0.29	0.05	0.03	0.05
PERC [mm/day]	3.19	1.46	0.73	5.29	5.90	1.73	0.05	2.87	2.49	5.99	2.12	0.21	2.00
Cflux [mm/day]	0.01	0.05	1.41	0.01	5.67	1.38	0.70	1.63	0.00	5.80	1.38	5.81	1.41

In figure 4.1 the SCEM calibration is shown for two subcatchments, both unsorted and sorted iterations are visualized. During calibration, simultaneously multiple paths within the parameter space are explored for better objective function scores. Graphically displaying simulation results in the order they were calculated does not show intuitively to what degree the model has converged. By sorting simulations by objective function score, the convergence becomes clear visually in the graph. Omos1 is an upstream catchment, that converges towards the 'best' parameter set after few iterations. The sorted figure shows all the parameter sets ranked from left to right according the highest score for objective function y. In earlier calibration runs the catchments converged well within 5000 iterations. However, after some changes in handling uncalibrated areas it can be seen that for instance Umos1 has not fully converged after 5000 iterations. In a further study an adaptive iteration could be further investigated.

As mentioned before, most subcatchments converge within the set boundaries, for Umos1 its path however collides with the boundaries, the best parameter values are close or equal to the parameter ranges set (table 4.1). This may be an issue as parameter values outside the parameter ranges set might have resulted in an even better objective function score. But boundaries are set for a reason, to find model parameters within a 'realistic' range. Sorted SCEM convergence figures for HBV and GR4J for all catchments can be found in appendix C.

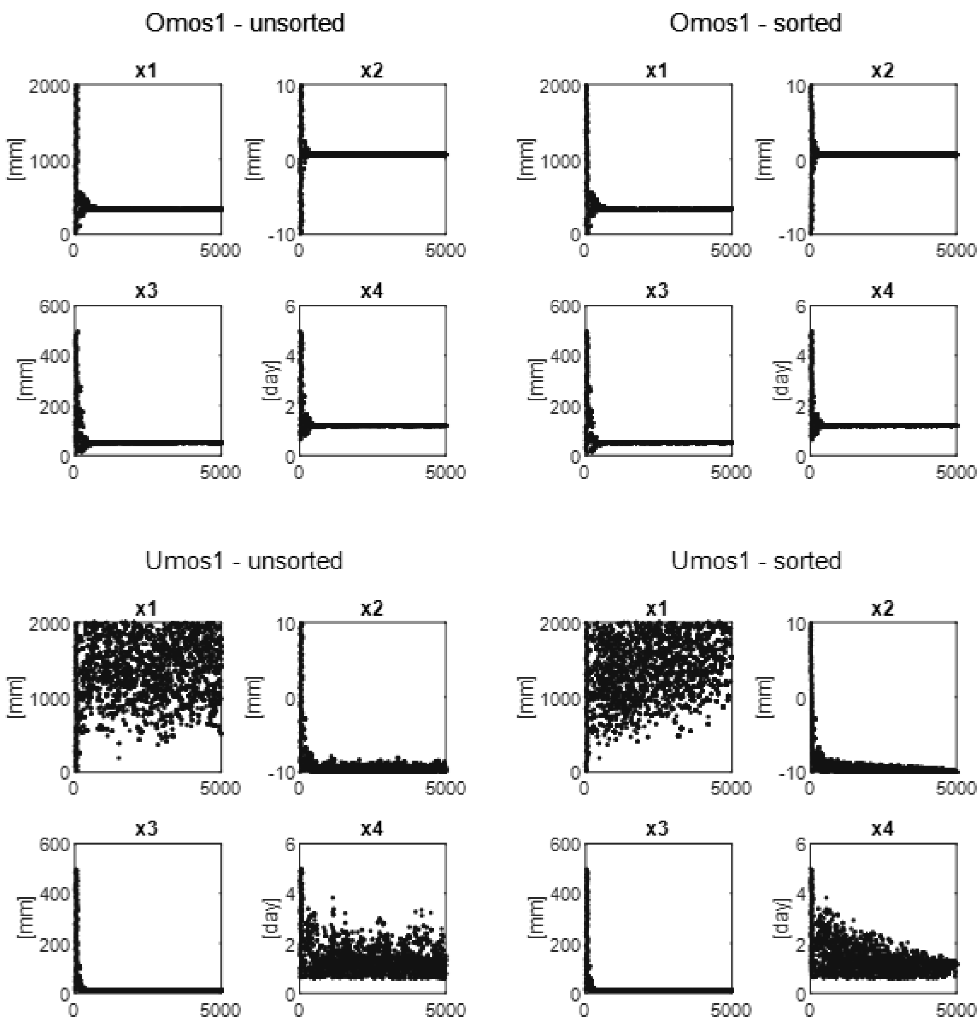


FIGURE 4.1 - GR4J SCEM CONVERGENCE GRAPHS FOR OMOS1 & UMOS1

The subcatchments are divided in four main areas corresponding to their location: south, west, east and lastly north, where the river disembogues into the main river Rhine. A schematic overview of the area can be seen in figure 3.3 (in the previous chapter). The catchments without a score do not have a discharge station with measured flow, thus could not be calibrated and subsequently scored.

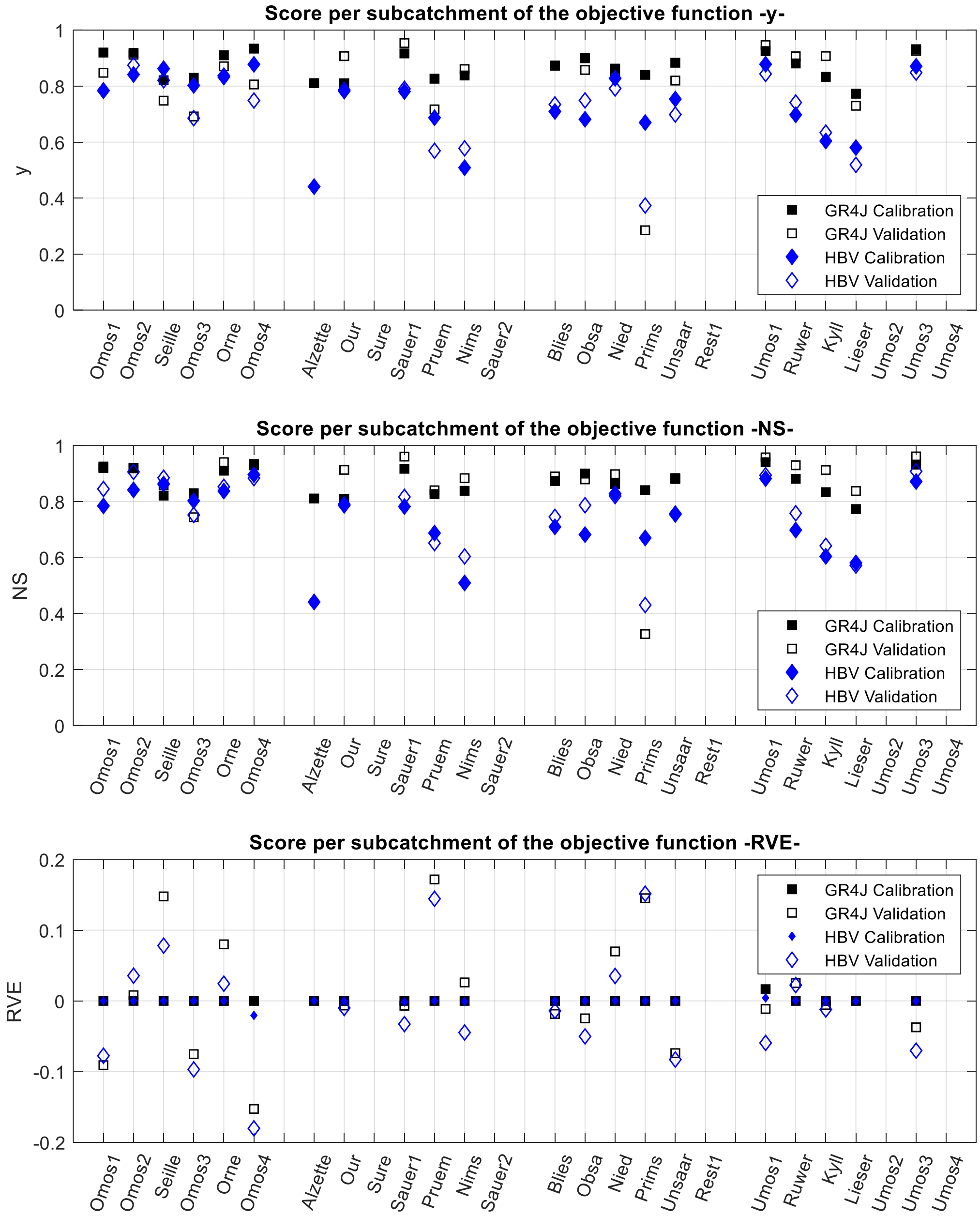


FIGURE 4.2 - SCORES PER SUBCATCHMENT FOR THE OBJECTIVE FUNCTIONS Y, NS, RVE

In figure 4.2 the calibration and validation scores are shown for objective function y, as well as the objective functions Nash Sutcliffe (NS) and relative volume error (RVE).

Overall GR4J scores better than HBV. Downstream area Umos3 (Cochem) has a high score for both models, with NS > 0.9, RVE < 10%. This is of importance for other studies, which apply a SOBEK model further downstream with discharges generated at Cochem as input.

In the previous paragraph an observation was made on multiple catchments with parameter sets close to the set boundaries. In this paragraph however, downstream catchments still score very well (i.e. Umos1), as the large inflows from upstream are the main contributing factor to the outflow, and hence the score.

Although most catchments show very promising results, there are exceptions. As could have been predicted upfront due to a lack of data as shown in figure 3.4, Alzette cannot be validated due to lack of discharge data for the validation period. Subcatchments with the lowest discharges are also scoring worst, especially for the HBV model.

The objective function y navigates for a perfect RVE. As can be seen in figure 4.2c, most calibrated RVE values are equal to 0. However, the validated volume error varies between -19% and +18%.

Modelled annual peaks are compared to the observed annual peaks in figure 4.3. Focus is put towards areas Omos4, Sauer1, Unsaar, Umos3, as those are the most downstream calibrated catchments for the previously defined South, West, East and North parts of the Moselle.

Figure 4.3 visualizes the underprediction and overprediction of annual maximum daily discharges with respect to the measured peak values. When a simulated event is located directly on the black line in the graph, it indicates that the simulated value is equal to the observed discharge. The dotted lines indicate a +/- 20% interval.

The two highest discharges were deliberately chosen to be part of the validation set, since the scope of this research involves the prediction of even more extreme discharges than measured in the past. As can be seen in Umos 3, visualized in figure 4.3 below, the two highest simulated flood magnitudes correlate well with the measured annual maximum discharge of the corresponding years. It should be noted that the measured values of these extreme discharges are uncertain.

One particular outlier is visible in the Omos4 graph. This is due to missing data from that year during the simulated peak. For this year, the highest observed discharge occurred on a different date, for which the data was available. As multiple subcatchments have these missing data points (figure 3.4), this may not be the only occurrence of this issue.

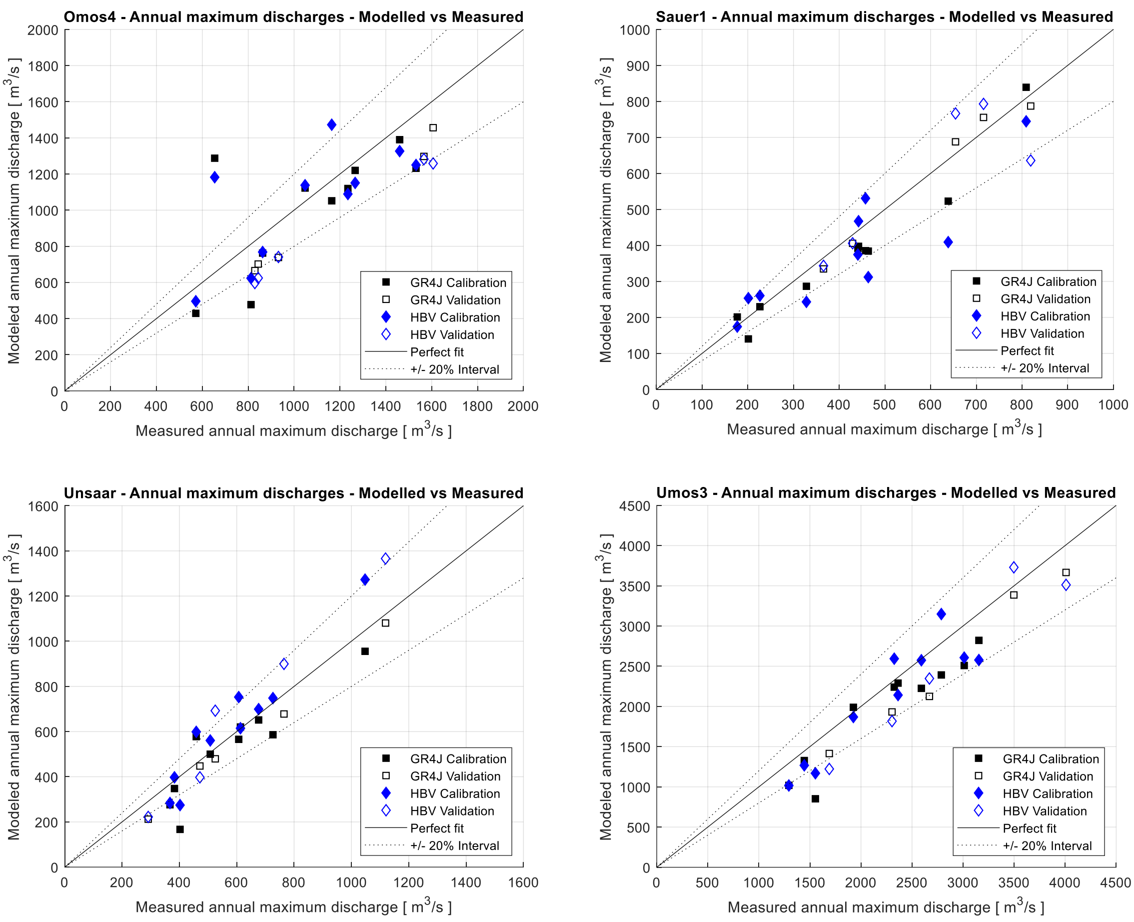


FIGURE 4.3 - ANNUAL MAXIMUM DISCHARGES DISPLAYED AGAINST MODELED FOR AREAS OMOS4, SAUER1, UNSAAR, UMOS3

4.2 VALIDATION RESULTS - SYNTHETIC CLIMATE DATA

This paragraph covers the validation of model outcomes using the synthetic 50.000 year series, with Precipitation and Temperature generated by a weather generator.

Both GR4J and HBV received these same inputs, and their synthetic modelled outputs are presented here on Gumbel reduced variate plots, alongside observed and modelled datapoints covered earlier in paragraph 4.2 as a reference. Gumbel nonlinear plots allow to display numerical data over a veryu wide range of values in a compact manner. Omos4, Sauer1, Unsaar, Umos3 are shown here, with the full set of graphs for all catchments available in appendix E.

For all catchments, the 15 years of actual observed data as well as the simulated 15 years based on observed climatic input are located above the synthetic 50.000 years series in the plots. This is partly due to the fact that the 15 years selected for this research contain two of the highest peaks observed in the 20th century. The weather generator uses a larger time series as input for resampling than the 15 years used in the calibration/validation set.

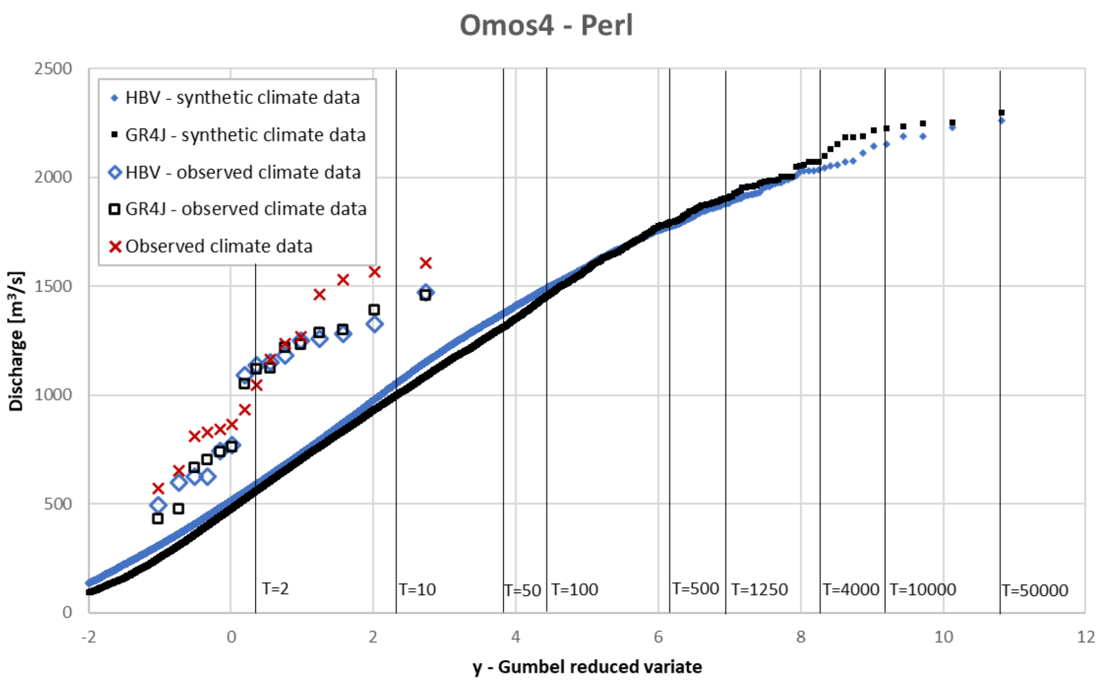


FIGURE 4.4A - FLOOD FREQUENCY CURVE OMOS4

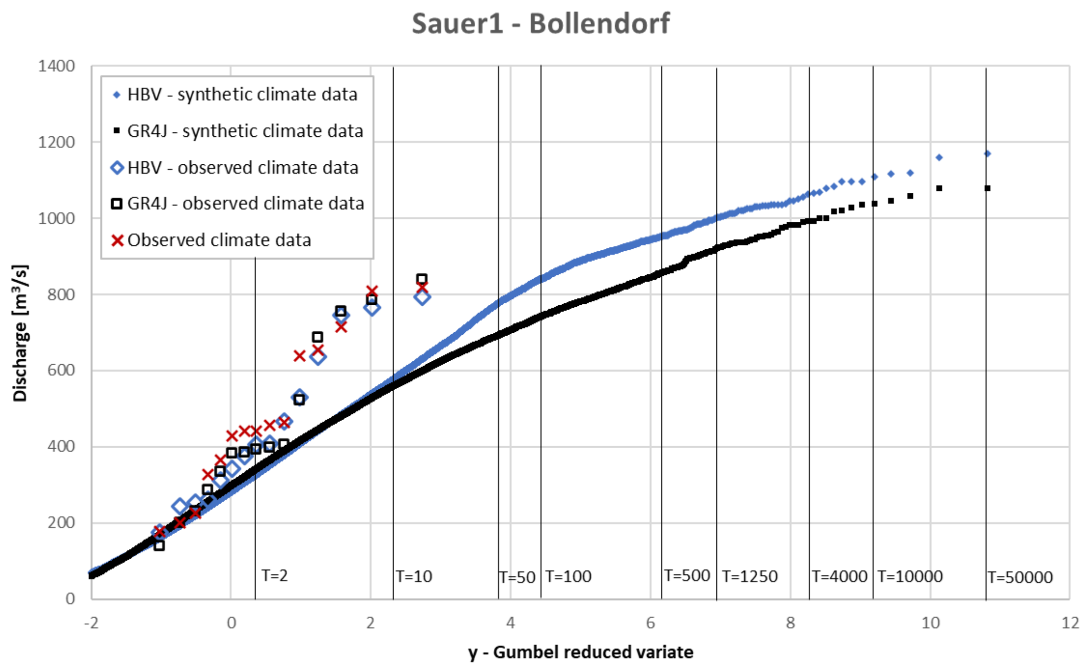


FIGURE 4.4B - FLOOD FREQUENCY CURVE SAUER1

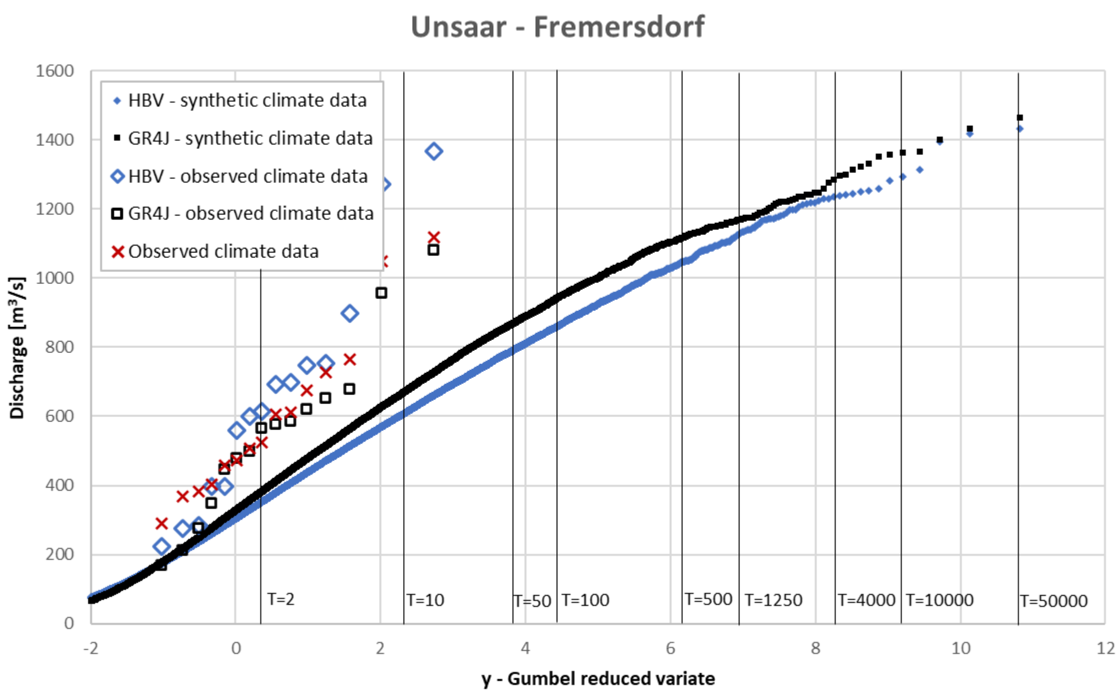


FIGURE 4.4C - FLOOD FREQUENCY CURVE UNSAAR

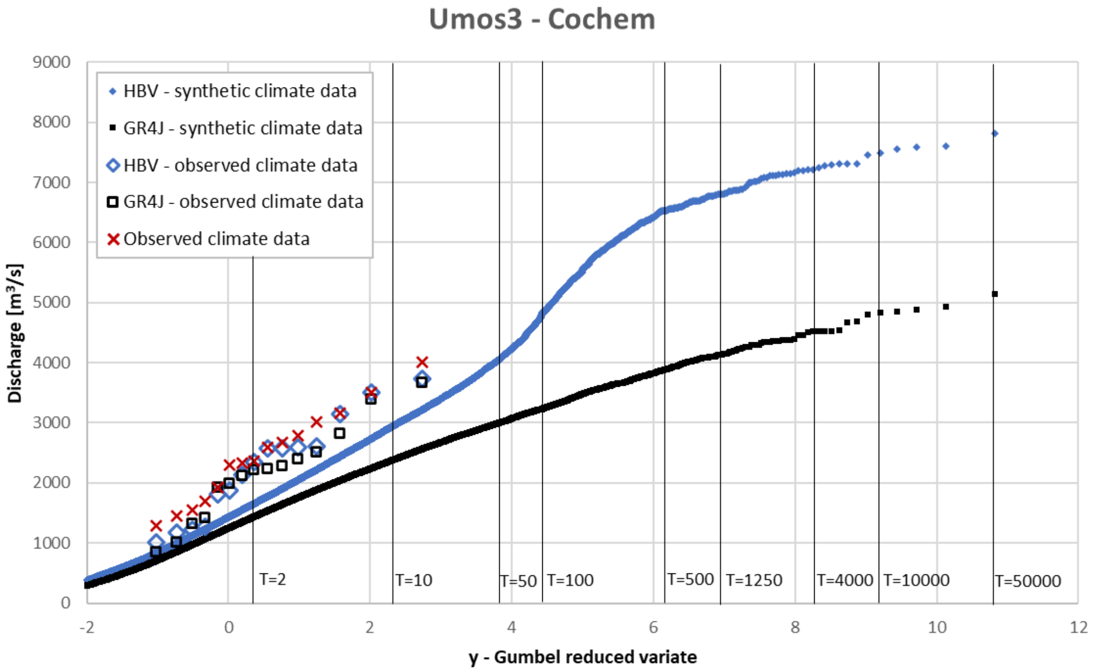


FIGURE 4.4D - FLOOD FREQUENCY CURVE UMOS3

When analysing figures 4.4a through 4.4d, Omos4 and Unsaar show no significant difference between GR4J and HBV. Saurer1 shows a minor drift upward for HBV at the end of the scale, a pattern which is obvious in Umos3.

Areas in which a drift occurs have one factor in common: an uncalibrated upstream subcatchment. For Sauer1 this is Sure, and for Umos3 this is Umos2, Rest1, and Sauer2, as can be seen in the flowchart in figure 3.3

In earlier runs, this effect is even more extreme. As an uncalibrated upstream catchment was identified as the source, another approach was taken in selecting substitute parameters for uncalibrated catchments, diminishing this effect. This is further elaborated upon in Appendix D.

In paragraph 4.1 the downstream effect of uncalibrated catchments was already discussed. SCEM found a parameter set for catchment Umos3 that scored well in both calibration and validation, however with the more extreme events in the synthetic series a divergence is found between GR4J and HBV. This could be an overcompensation by the model.

A method to verify if the models are overcompensating after an uncalibrated or badly calibrated sub-catchment, is to check the ratio between precipitation and discharge. For all discharge series (regarding 15 and 50000 years of input), for the whole Moselle catchment, this ratio $\sum P/\sum Q$ is equal to 2.5.

Table 4.4 shows these ratios per sub-catchment. Conditional formatting is applied for values deviating from the median of all four values of the respective sub-catchment. As can be seen in the flowchart (figure 3.3) Umos1 has two directly upstream uncalibrated areas. Table 4.4 shows compensation by the GR4J model for those uncalibrated areas. Umos1 retains water, the discharge/precipitation ratio shows an extremely low value. Downstream of Umos1, Umos2 is uncalibrated. Umos3 is then overcompensating in the other direction. GR4J as well as HBV show an unrealistic ratio in table 4.4 for Umos3, because of compensating for all of the 'mistakes' made upstream.

TABLE 4.4 - DISCHARGE/PRECIPITATION RATIO PER SUBCATCHMENT PER MODEL PER SIMULATION

	SOUTH							*WEST*						
	Omos1	**Omos2**	**Seille**	**Omos3**	**Orne**	**Omos4**		**Alzette**	**Our**	**Sure**	**Sauer1**	**Pruem**	**Nims**	**Sauer2**
GR4J	0.48	0.41	0.32	0.27	0.33	0.44		0.33	0.45	0.40	0.65	0.43	0.37	0.36
HBV	0.49	0.42	0.32	0.12	0.33	0.60		0.34	0.45	0.40	0.62	0.43	0.37	0.37
synthetic GR4J	0.36	0.26	0.36	0.55	0.36	0.44		0.42	0.37	0.44	0.65	0.45	0.26	0.38
synthetic HBV	0.39	0.31	0.35	0.24	0.35	0.62		0.41	0.38	0.43	0.63	0.45	0.31	0.39
	EAST							*NORTH*						
	Blies	**Obsa**	**Nied**	**Prims**	**Unsaar**	**Rest1**		**Umos1**	**Ruwer**	**Kyll**	**Lieser**	**Umos2**	**Umos3**	**Umos4**
GR4J	0.34	0.31	0.36	0.44	0.56	0.43		0.03	0.51	0.40	0.35	0.36	1.01	0.25
HBV	0.34	0.31	0.36	0.44	0.57	0.44		0.22	0.51	0.40	0.35	0.39	0.86	0.33
synthetic GR4J	0.38	0.29	0.39	0.43	0.52	0.52		0.04	0.48	0.24	0.28	0.40	1.02	0.32
synthetic HBV	0.36	0.29	0.38	0.42	0.54	0.49		0.54	0.48	0.25	0.28	0.42	0.96	0.37

CHAPTER 5

DISCUSSION

This study compared hydrological models GR4J and HBV, first in a fifteen year observed climate data set, followed by a synthetic 50.000 year resampled data set using a weather generator. In the first part of this study the GR4J model showed better objective function scores during calibration and validation of the observed climate data. When applied to the synthetic 50.000 year data series, it is HBV showing larger runoffs at the final point of the Moselle, flowing into the Rhine.

INTERPRETATIONS

With an outperformer in the calibration/validation phase of the research, some deviation in outputs in the synthetic series were to be expected. Found differences in outputs were traced back to the uncalibrated sub-catchments in the area. Only for the HBV model, extreme flood magnitudes, for higher return periods, appear to be sensitive for selection of method of dealing with these uncalibrated subcatchments.

Results show a flatter flood frequency curve for both HBV and GR4J using the synthetic data as compared to the observed climate data. The observed data series is only a subset of the climate data used for resampling for the weather generator, with two of the most extreme events occurring in this subset of fifteen years.

IMPLICATIONS

This research shows a sensitivity for selection of hydrological model in the Generator of Rainfall And Discharge Extremes instrument. This would imply that the original research (Hegnauer et al., 2014) could be amended by rerunning the instrument with other hydrological models.

Differences in expected flood magnitudes for a certain return period could on one hand indicate that the Netherlands is less safe as previously simulated. On the other hand it could be safer if flood magnitudes would actually be less extreme, hence flood protection investments could be reduced.

LIMITATIONS

In this research GR4J scored better during calibration, and showed more desirable results: lower flood magnitudes. This can not be generalized to a point where GR4J would always be the model of choice over HBV. SCEM convergence graphs as shown in appendix D, visualize that HBV did not converge fully in all subcatchments, especially in subcatchments after uncalibrated subcatchments.

GENERALIZATION

No studies have been found applying multiple hydrological models to a synthetic data set. Parallel to this research, another master thesis (Brink, 2018) took place on a different basin, the Meuse, which shows no discernible difference between GR4J and HBV, contrary to this study. That study however did not include multiple uncalibrated subcatchments.

This research, which includes various uncalibrated subcatchments, shows that hydrological model choice could impact the flood frequency curve, hence other studies, especially with uncalibrated subcatchments, should consider applying multiple hydrological models.

During this research, the calibration procedure was fully automated, limiting the required time to implement additional hydrological models, or make small adjustments. If this automation process can be further generalized, it should no longer be a practical limitation to only use one model in a study. Especially for studies on topics as important as our national flood defence.

CHAPTER 6

CONCLUSION AND RECOMMENDATIONS

This chapter covers the conclusion and recommendations following findings in previous chapters and answers the research questions.

6.1 CONCLUSION

The two research questions formulated for this research were as follows:

1. *Which hydrological model shows the best performance after calibration/validation, regarding observed climate data?*

   Both GR4J and HBV perform well regarding observed climate data, but GR4J outperforms HBV. On total modelled flow (RVE) scores are comparable, but the measure which focuses on peak discharges (NS), GR4J scores significantly better on all but one sub-catchments.

2. *What influence do different hydrological models combined with the weather generator have on the flood magnitudes?*

   When combined with the weather generator's synthetic series, the HBV model shows a sensitivity for uncalibrated upstream areas. This effect is not observed with GR4J. HBV predicts more extreme flood magnitudes for high return periods regarding the annual maxima than GR4J.

With the two research questions answered, the goal of this research is met: to assess whether the choice of hydrological model within the discharge generator affects the annual maximum discharges of the Moselle basin, and therefore predicting different extreme discharges at certain return periods.

For the Moselle basin used in this research, model choice does influence annual maximum discharge, with HBV predicting lower extreme discharges than GR4J. HBV appears to be sensitive for missing data, subcatchments with uncalibrated upstream neighbours have extreme parameter values and produces nonlinearity in the flood frequency curve when combined with the synthetic data series.

Differences in expected flood magnitudes for a certain return period could on one hand indicate that the Netherlands is less safe as previously simulated. On the other hand it could be safer if flood magnitudes would actually be less extreme, hence flood protection investments could be reduced

6.2 Recommendations

Stakeholder recommendations

The GRADE instrument, used as a basis for this research, could also benefit from including multiple hydrological models, especially in areas with scarce data. Based on this research, HBV appears to be more susceptible for incomplete data than GR4J, but further research should confirm.

This research shows the value of complete data. Investing in further data acquisition, retention and accessibility could pay off in terms of quality of research outcomes.

Further Research Recommendations

The automated framework developed for this study could be extended with different hydrological models improving the insight of the influences of these models, combined with the weather generator, on the flood frequency curves.

For some catchments, to ensure full convergence during calibration, the number of iterations should be increased. However, this is not necessary for the majority of the sub-catchments. Increasing the number of iterations for all sub-catchments increases the calculation time significantly, therefore an adaptive iteration function could be implemented to optimize both.

The impact of uncalibrated catchments could be further investigated. One could for instance look into better substitute parameters or merge uncalibrated catchments downstream. As stated in Appendix D, during the research two methods have been investigated, showing differences for the high return periods for HBV output.

The current objective function could be re-evaluated to include another term that focusses on extreme peak discharge performance.

The current lag function that couples the different catchments uses an average 2m/s for the whole Moselle Catchment. This could be further refined, for instance with different values per subcatchments based on e.g. slope and roughness of the bedding.

With scarce data as in this research, warmup of the model could be done using a duplicate year instead of the first year of the available data set. This ensures more available data to calibrate/validate.

Other data sets than used in this research exist. Combining into a richer data set could mitigate issues around data quality, both a longer timeframe and more complete measurement data would be beneficial.

REFERENCES

Akhtar, M., Ahmad, N. & Booij, M. J. (2009) Use of regional climate model simulations as input for hydrological models for the Hindukush-Karakorum-Himalaya region. *Hydrol. Earth System Sci.* 13, 1075–1089.

Brink, C. J. R. (2018). *Influence of hydrological model structures on extreme high flow simulations in the Meuse basin* (Master's thesis, University of Twente).

Chen, J., Brissette, F., & Leconte, R. (2010). A daily stochastic weather generator for preserving low-frequency of climate variability. *Journal of Hydrology,* 388(3-4), 480-490.

Dawdy, D. R., & O'Donnell, T. (1965). Mathematical models of catchment behavior. *Journal of the Hydraulics Division*, *91*(4), 123-137.

Demirel, M. C., Booij, M. J., & Hoekstra, A. Y. (2013). Effect of different uncertainty sources on the skill of 10 day ensemble low flow forecasts for two hydrological models. *Water resources research*, *49*(7), 4035-4053.

Demirel, M. C., Booij, M., & Hoekstra, A. (2015). The skill of seasonal ensemble low-flow forecasts in the Moselle River for three different hydrological models. *Hydrology and earth system sciences*.

Demirel, M. C., Özen, A., Orta, S., Toker, E., Demir, H. K., Ekmekcioğlu, Ö., ... & Tuncer, E. (2019). Additional value of using satellite-based soil moisture and two sources of groundwater data for hydrological model calibration. *Water*, *11*(10), 2083.

Duan, Q., Sorooshian, S., & Gupta, V. (1992). Effective and efficient global optimization for conceptual rainfall-runoff models. *Water Resources Research,* 28(4), 1015-1031.

Falter, D., Schröter, K., Dung, N., Vorogushyn, S., Kreibich, H., Hundecha, Y., . . . Merz, B. (2015). Spatially coherent flood risk assessment based on long-term continuous simulation with a coupled model chain. *Journal of Hydrology,* 524, 182-193.

Hegnauer, M., Beersma, J., & van den Boogaard, H. (2014b). Toekomstige ontwikkelingen GRADE. Delft: Deltares.

Hegnauer, M., van den Boogaard, H., Buishand, T., & Passchier, R. (2014). Generator of Rainfall and Discharge Extremes (GRADE) for the Rhine and Meuse basins - Final report of GRADE 2.0. Delft: Deltares.

Khalili, M., Brissette, F., & Leconte, R. (2011). Effectiveness of multi-site weather generator for hydrological modelling. *Journal of the American Water Resources Association*, 47(2), 209-431

Linström, G., Johansson, B., Persson, M., Gardelin, M., & Bergström, S. (1997). Development and test of the distributed HBV-96 hydrolocial model. 201, 272-288.

Nash, J., & Sutcliffe, J. (1970). River flow forecasting through conceptual models part 1 - A discussion of principles. Journal of Hydrology, 10, 282-290.

Pagano, T., Hapuarachchi, P., & Wang, Q. J. (2010). Continuous rainfall-runoff model comparison and short-term daily streamflow forecast skill evaluation.

Pechlivanidis, I. G., Jackson, B. M., McIntyre, N. R. & Wheater, H. S. Catchment scale hydrological modelling: a review of model types, calibration approaches and uncertainty analysis methods in the context of recent developments in technology and applications. *Global NEST Journal* 13,193–214 (2011).

Perrin, C., Michel, C., & Andréassian, V. (2003). Improvement of a parsimonious model for streamflow simulation. *Journal of hydrology*, *279*(1-4), 275-289.

Singh, V. P., & Woolhiser, D. A. (2002). Mathematical modeling of watershed hydrology. *Journal of hydrologic engineering*, *7*(4), 270-292.

Steinrücke, J. F. (2012). *HYMOG – Hydrologische Modellierungsgrundlagen im Rheingebiet*. Lelystad, The Netherlands: Internationale Kommission für die Hydrologie des Rheingebietes (KHR/CHR).

Van Den Tillaart, S. P., Booij, M. J., & Krol, M. S. (2013). Impact of uncertainties in discharge determination on the parameter estimation and performance of a hydrological model. *Hydrology research*, *44*(3), 454-466.

Van Esse, W., Perrin, C., Booij, M., Augustijn, D., Fenicia, F., Kavetski, D., & Lobligeois, F. (2013). The influence of conceptual model structure on model performance: a comparative study for 237 French catchments. *Hydrology and Earth System Sciences*, 2013; 17(10):4227-4239

Vrugt, J., Gupta, H., Bastidas, L., Bouten, W., & Sorooshian, S. (2003). Effective and efficient algortihm for multiobjective optimization of hydrologic models. Water Resources Research, 39(8).

Winsemius, H., van Verseveld, W., Weerts, A., & Hegnauer, M. (2013). *Generalised Likelihood Uncertainty Estimation for the daily HBV model in the Rhine Basin, Part A: Germany*. Delft: Deltares.

Zhang, X., Booij, M., & Xu, Y. (2015). Improved simulation for peak flows under climate change: postprocessing or composite objective calibration? Journal of Hydrometeorology.

APPENDIX A

DISCHARGE GENERATOR - GRADE

The following appendix is a based upon literature study preceding this research, a detailed description is provided there. The Generator of Rainfall and Discharge Extremes (GRADE) is a tool used for water policy analysis. This tool provides a more physically based method, compared to earlier methods, for the estimation of the design discharge and is used to determine the new flood standards of 2017 in the Netherlands: WTI2017 “Wettelijk Toetsinstrumentarium 2017”(Hegnauer et al., 2014). The basins in scope are the Rhine basin and the Meuse basin. GRADE comprises three components: the weather generator, a hydrological model, and a hydrodynamic model (figure A.1).

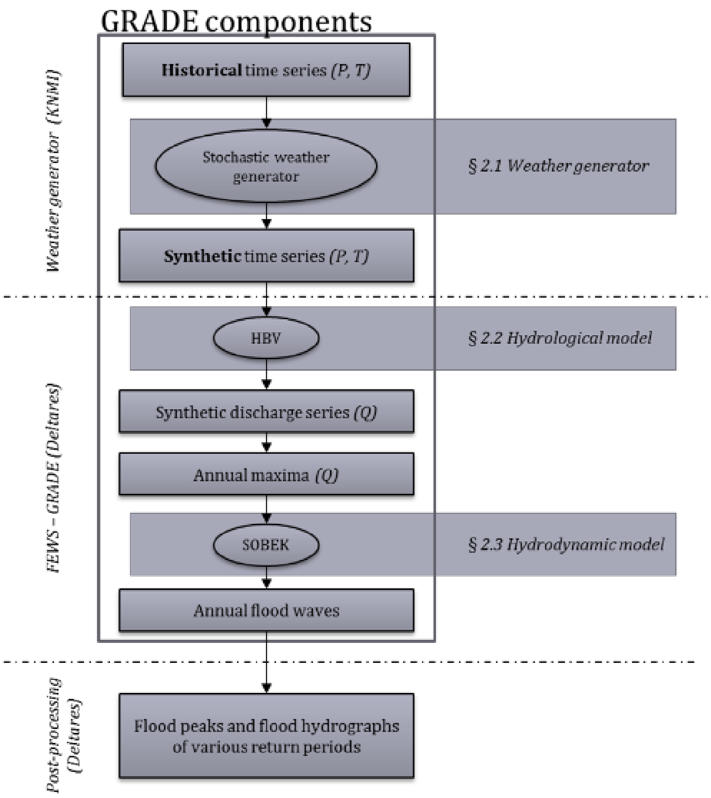


FIGURE A.1 GRADE COMPONENTS

APPENDIX B

AUTOMATED CALIBRATION FOR WHOLE CATCHMENT

The calibration procedure was fully automated, with one simulation within matlab all 26 catchments are calibrated and validated. A flowchart was added to assure all 26 subcatchments are calibrated in the correct order, using discharges from previous subcatchments with a simple lag-function.

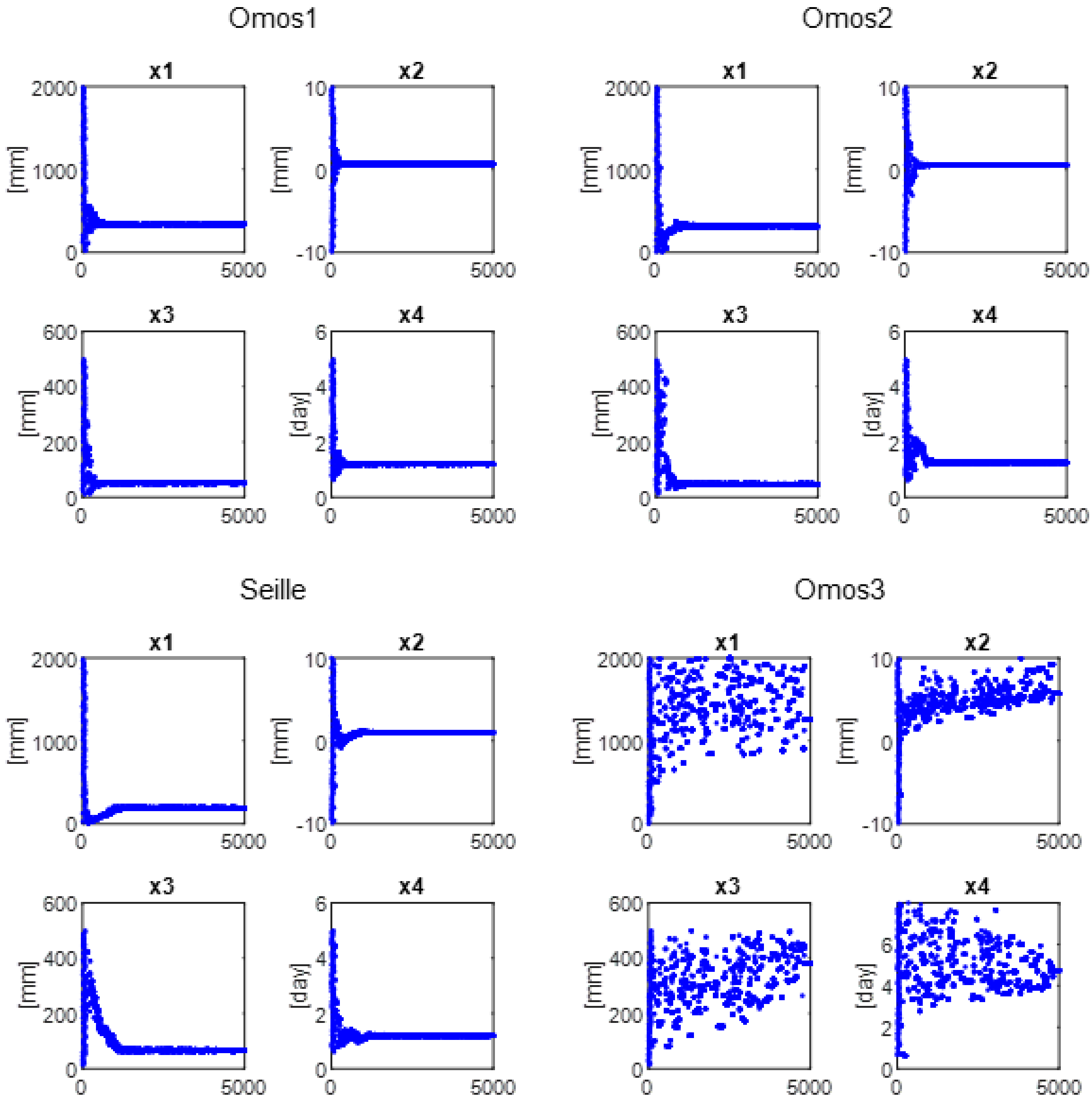
Simulation time on four core laptop processor:

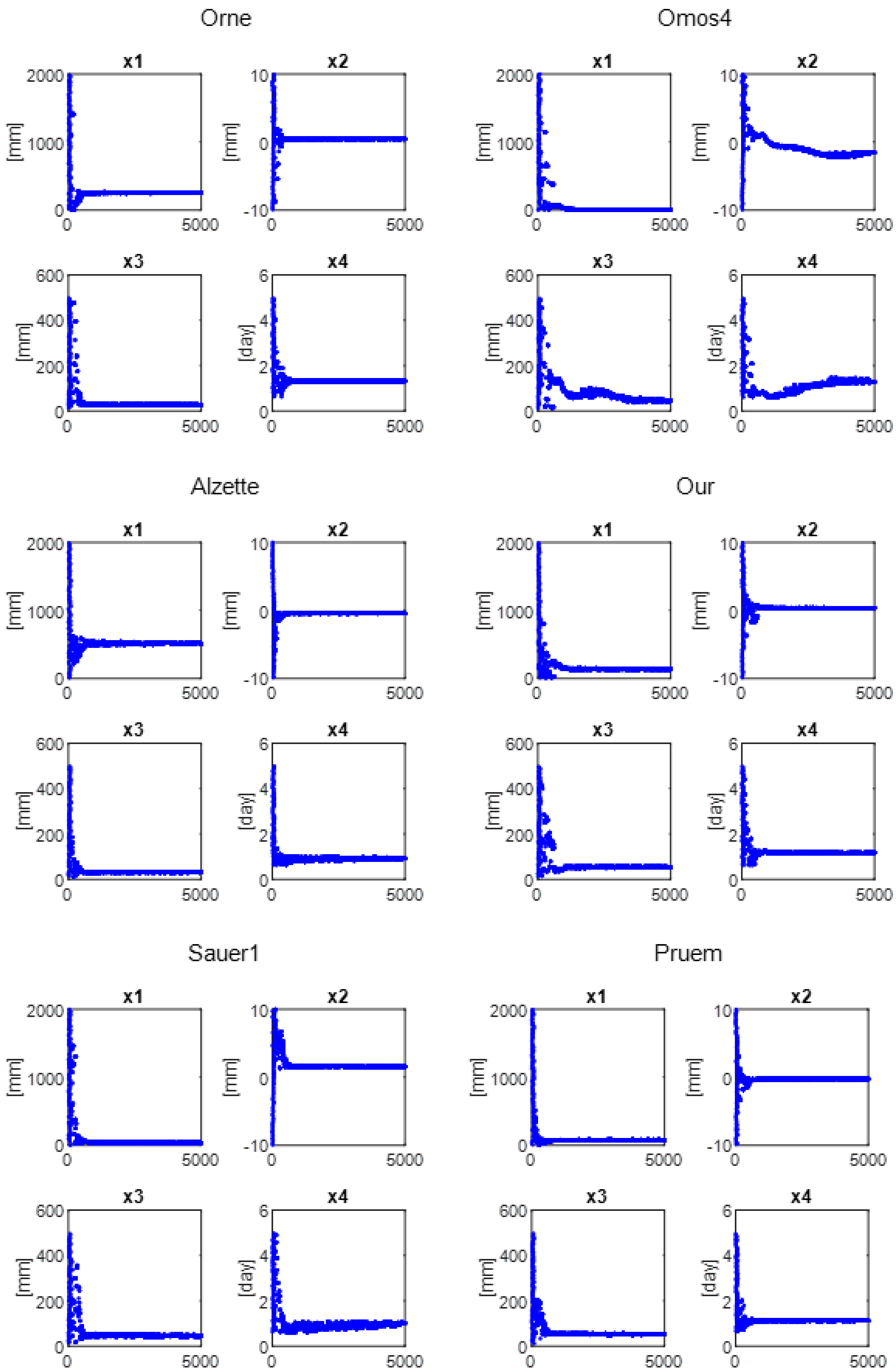
GR4J	
Calibration/Validation - observed climate data	45 min
Simulation - synthetic climate data	60 min
HBV	
Calibration/Validation - observed climate data	60 min
Simulation - synthetic climate data	180 min

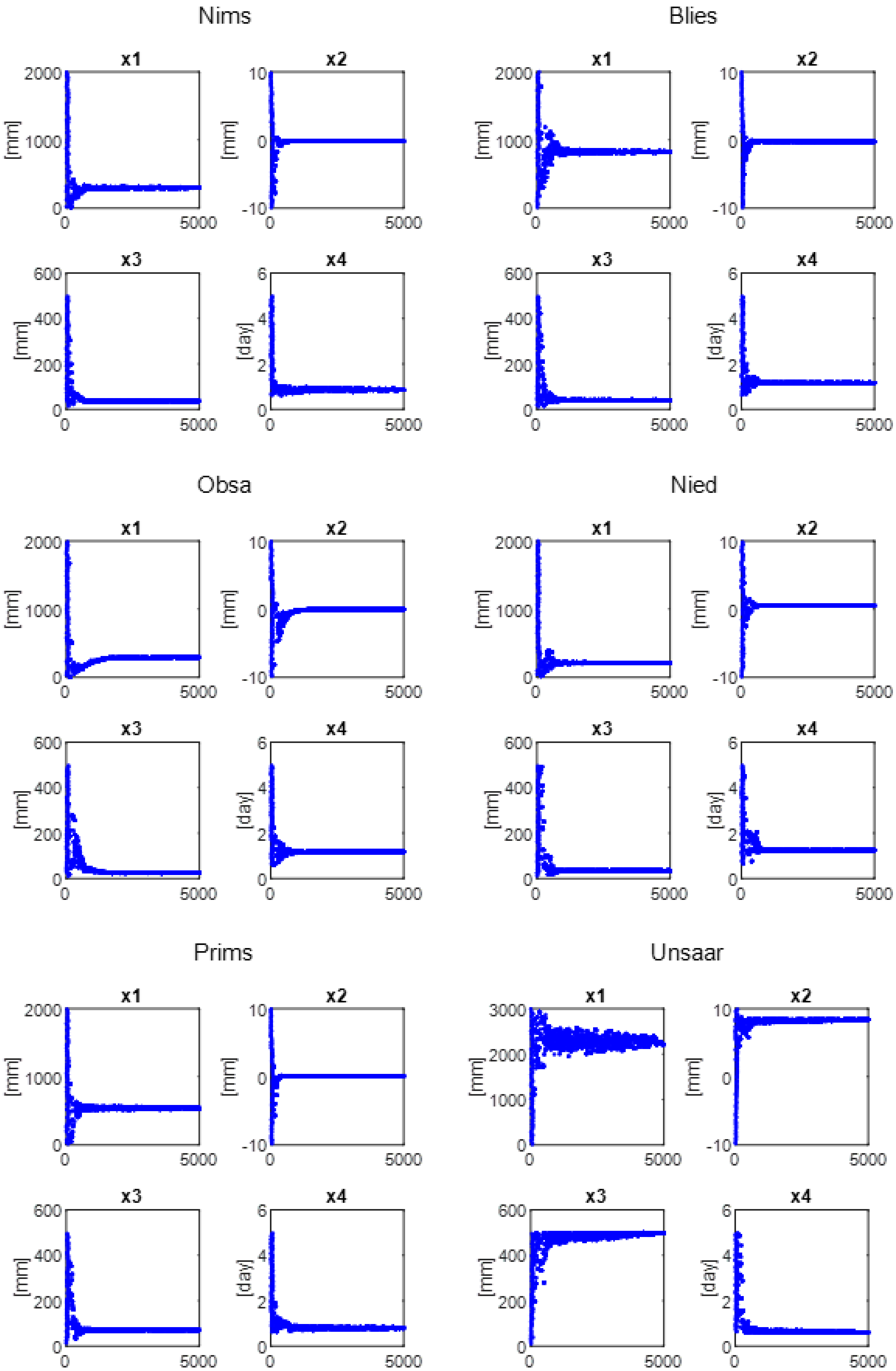
APPENDIX C

SCEM CALIBRATION - PARAMETER CONVERGENCE

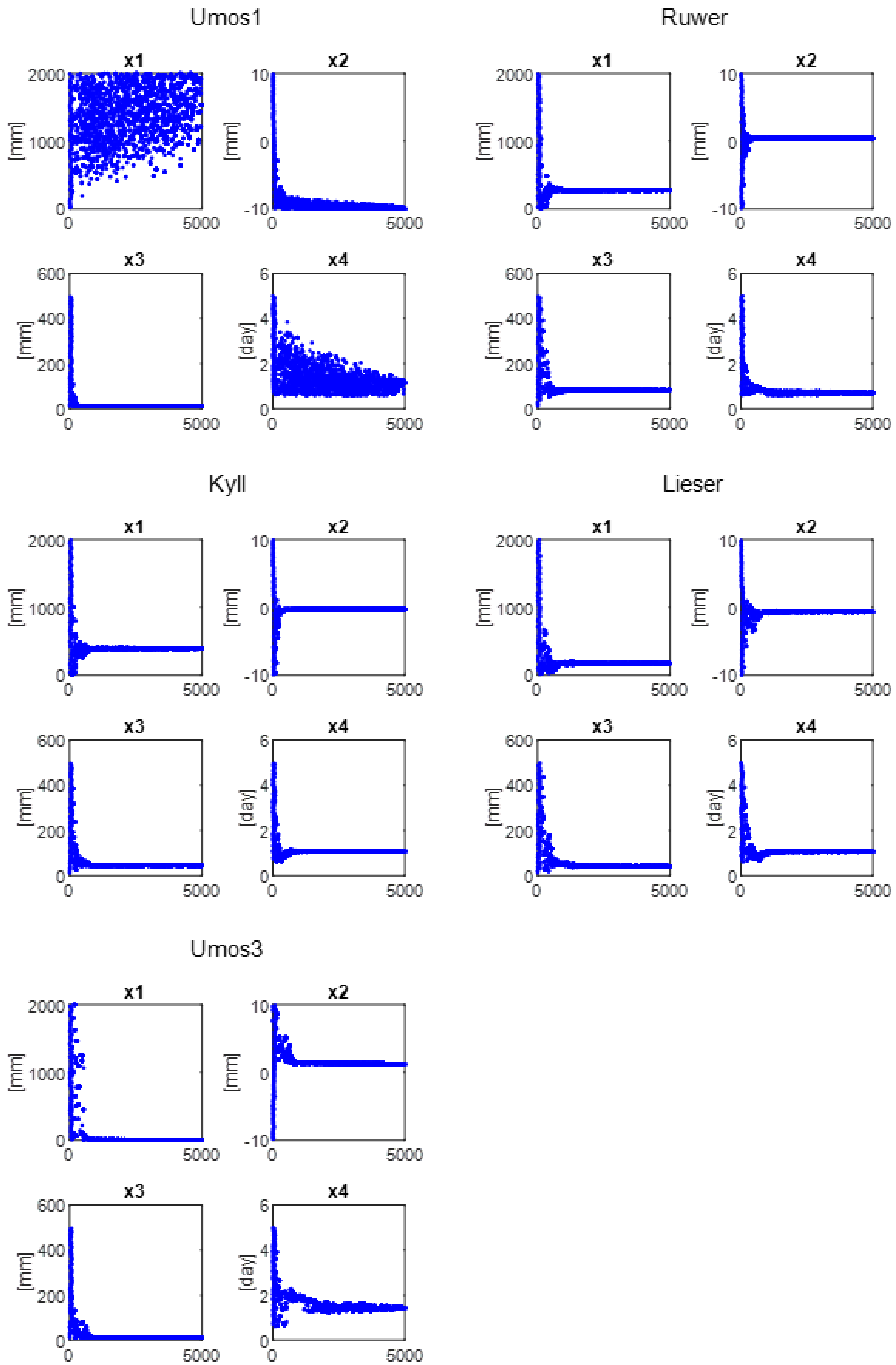
GR4J



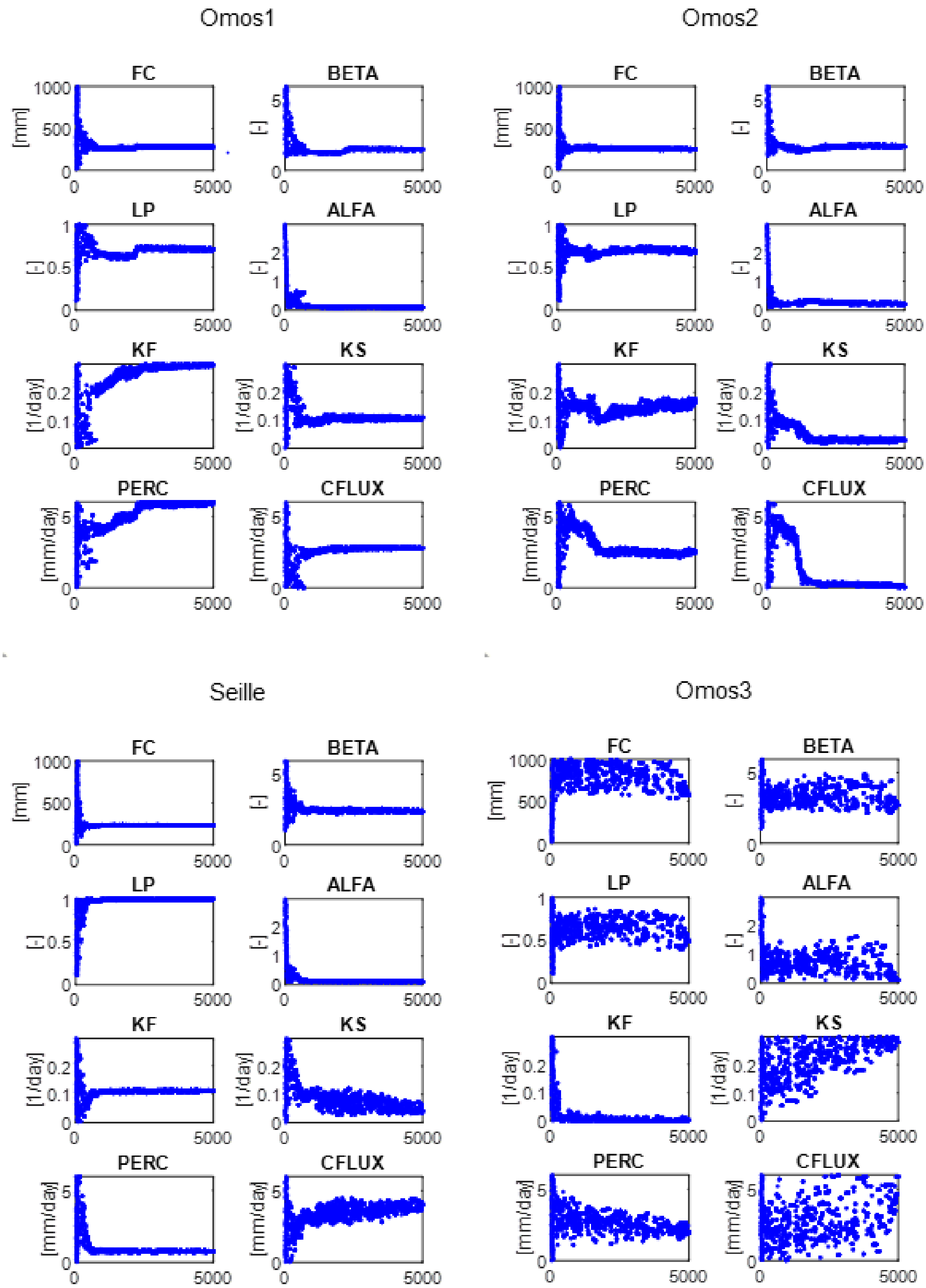

Orne
Omos4
Alzette
Our
Sauer1
Pruem
x1
x2
x3
x4
[mm]
[day]
2000
1000
0
10
-10
600
400
200
6
4
2
5000

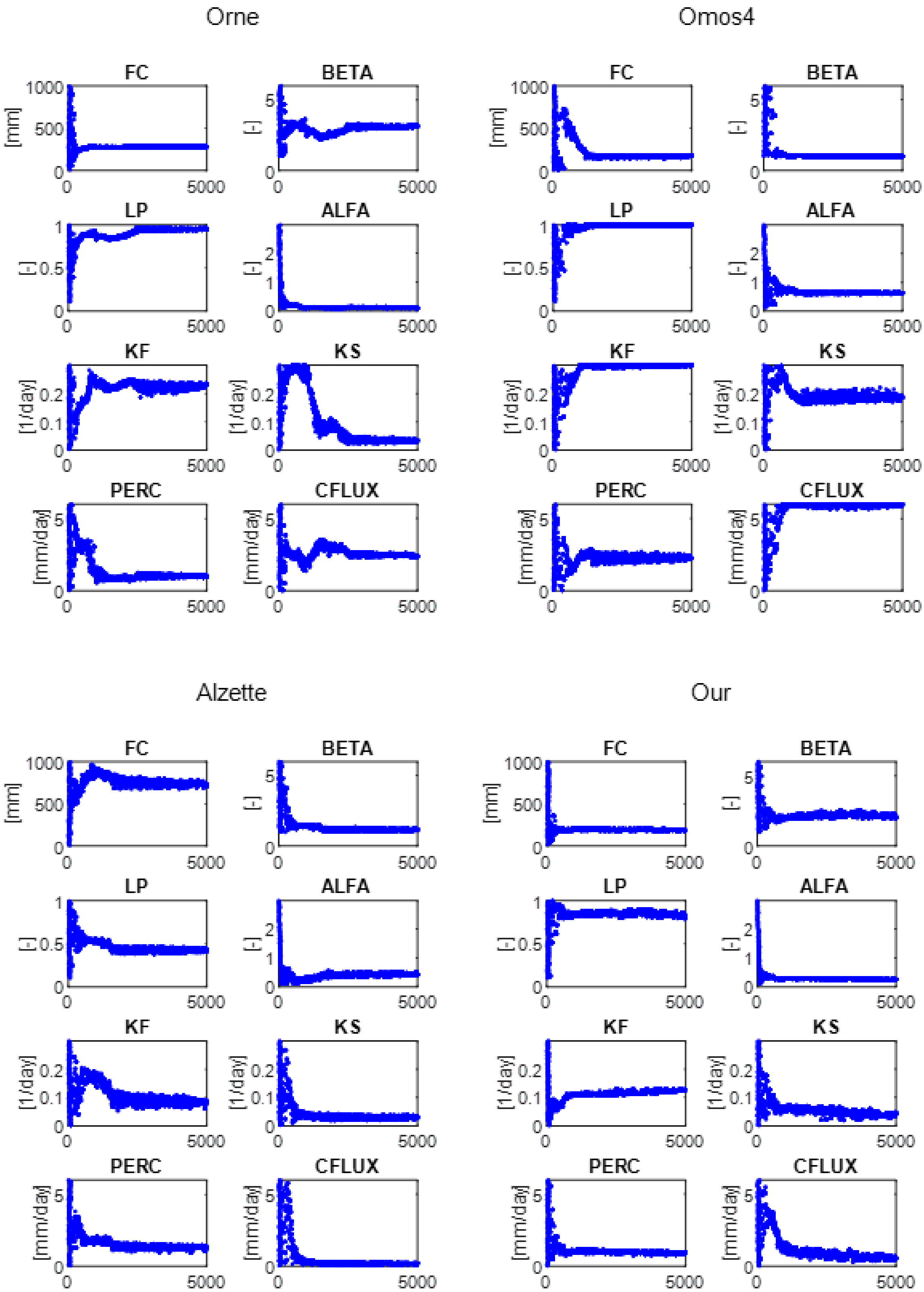


Nims
Blies
Obsa
Nied
Prims
Unsaar
x1
x2
x3
x4
[mm]
[day]
2000
1000
0
10
-10
600
400
200
6
4
2
3000
5000

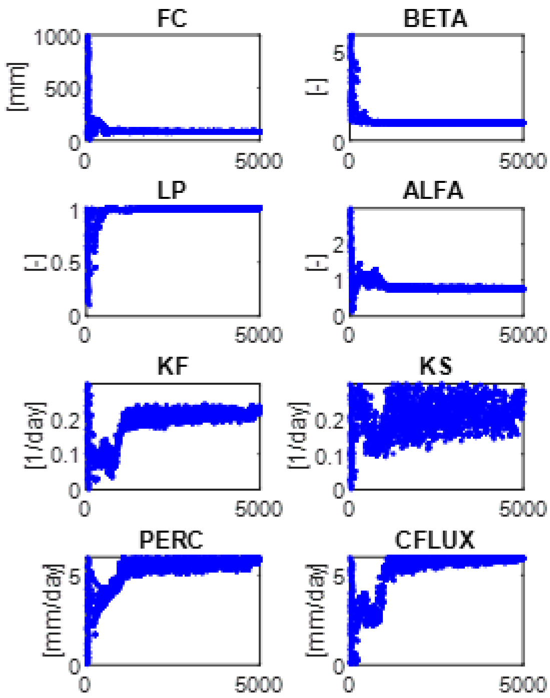

Umos1
Ruwer
Kyll
Lieser
Umos3
x1
x2
x3
x4
[mm]
[day]

HBV

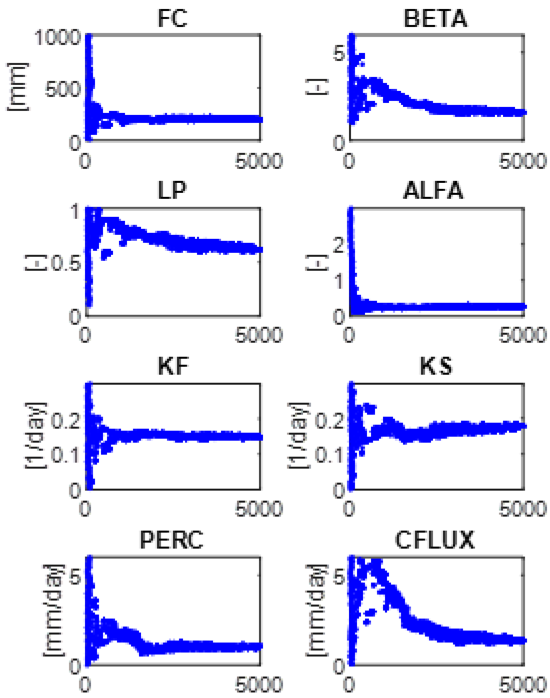

Omos1
Omos2
Seille
Omos3
FC
BETA
LP
ALFA
KF
KS
PERC
CFLUX
[mm]
[-]
[1/day]
[mm/day]
1000
500
0
5000
0.5
0.2
0.1


Orne
Omos4
Alzette
Our
FC
BETA
LP
ALFA
KF
KS
PERC
CFLUX
[mm]
[-]
[1/day]
[mm/day]
0
5000

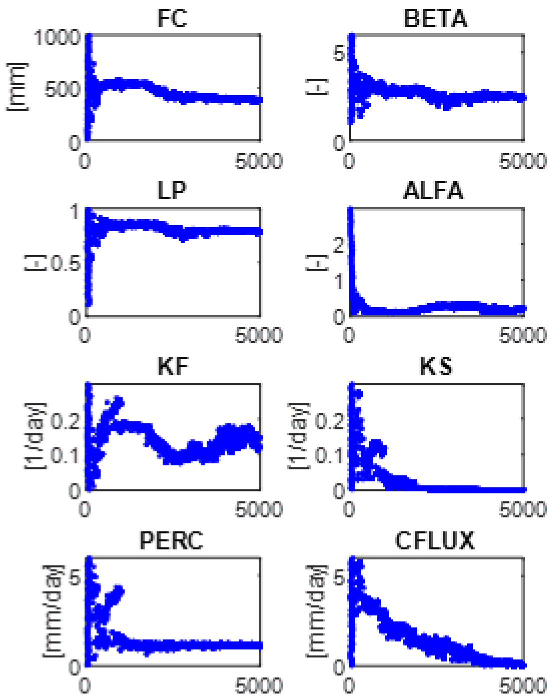
Sauer1



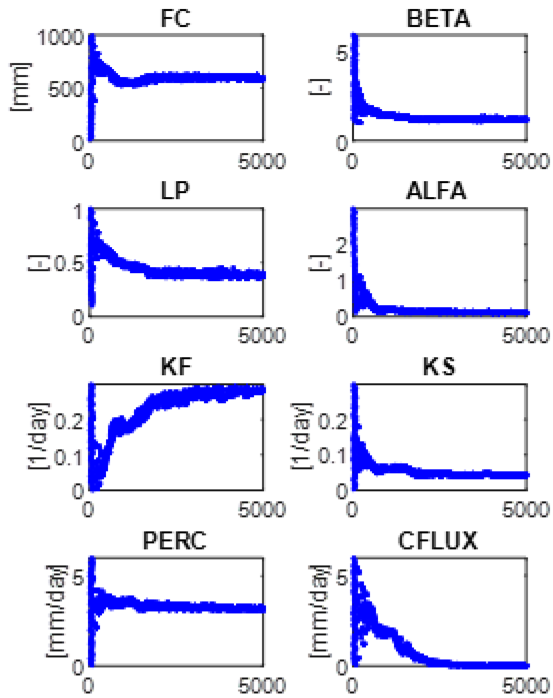
Pruem

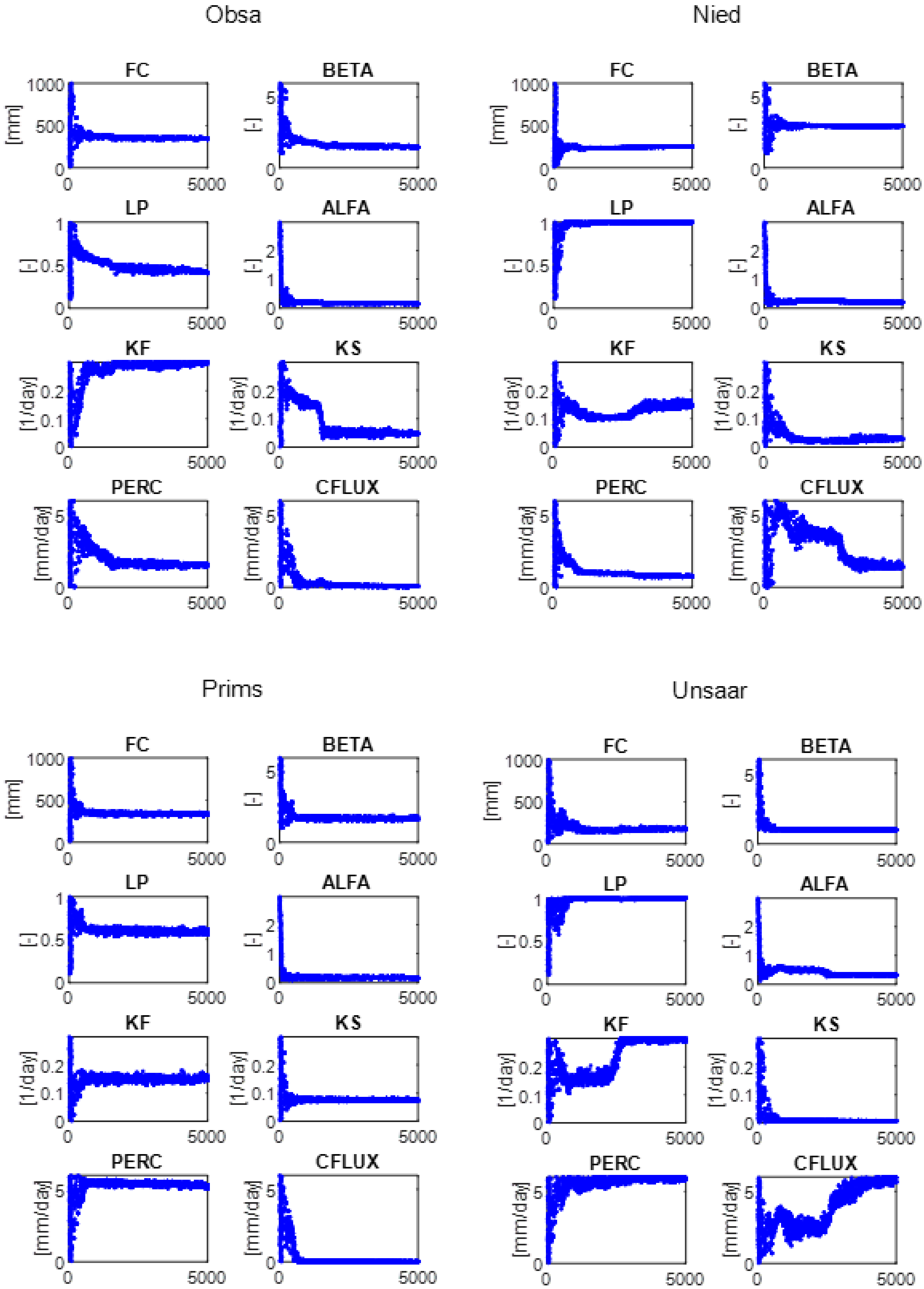


Nims



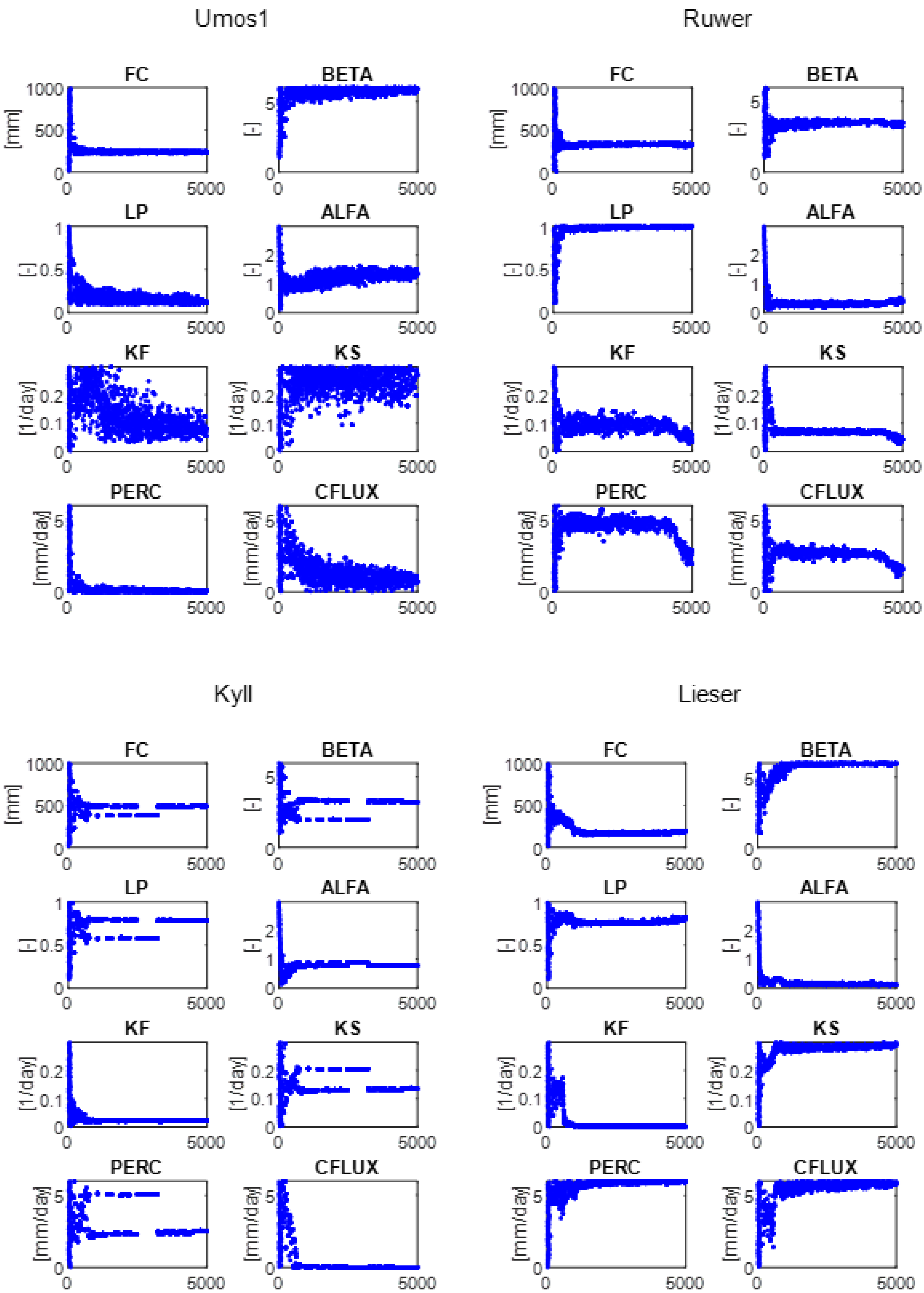
Blies

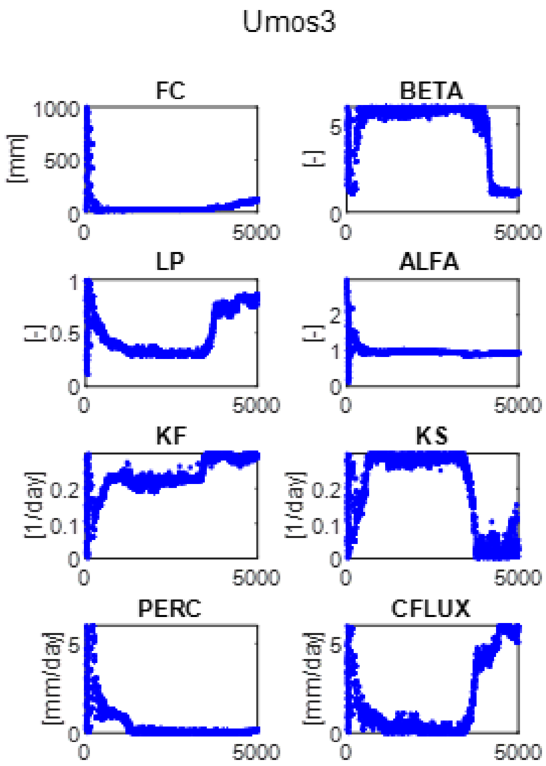




Obsa
Nied
Prims
Unsaar
FC
BETA
LP
ALFA
KF
KS
PERC
CFLUX
[mm]
[-]
[1/day]
[mm/day]
1000
500
0
5000

X


Umos1
Ruwer
Kyll
Lieser
FC
BETA
LP
ALFA
KF
KS
PERC
CFLUX
[mm]
[-]
[1/day]
[mm/day]
1000
500
0
5000
5
2
1
0.5
0.2
0.1

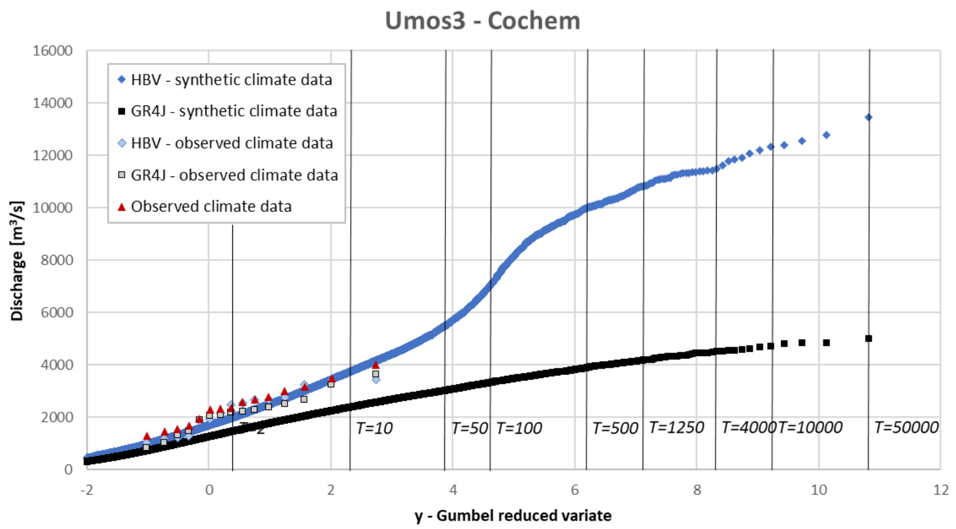

Umos3
FC
[mm]
1000
500
0
0
5000
BETA
[-]
5
0
0
5000
LP
[-]
1
0.5
0
0
5000
ALFA
[-]
2
1
0
0
5000
KF
[1/day]
0.2
0.1
0
0
5000
KS
[1/day]
0.2
0.1
0
0
5000
PERC
[mm/day]
5
0
0
5000
CFLUX
[mm/day]
5
0
0
5000

APPENDIX D

DIFFERENCE PARAMETER SET FOR UNCALIBRATED AREAS

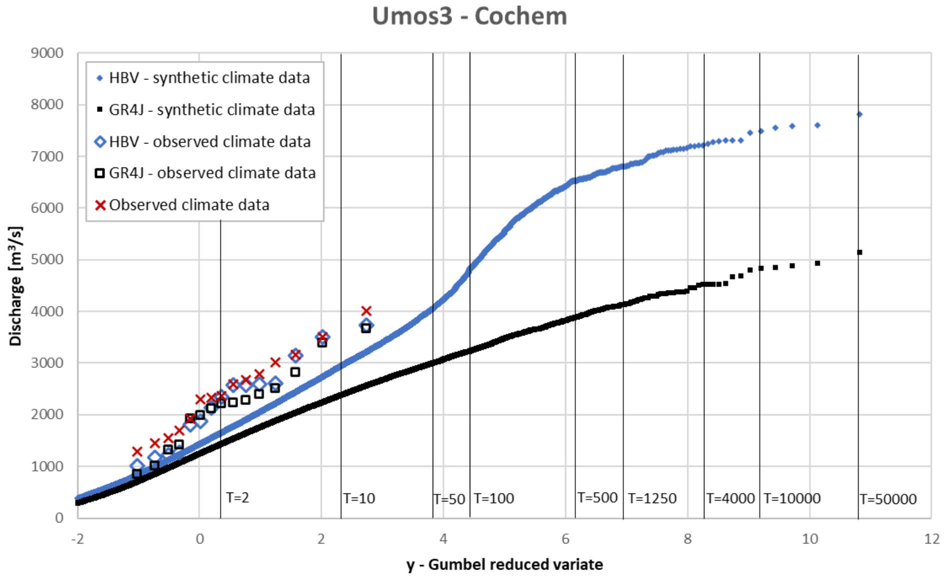
UNCALIBRATED AREAS WITH COUPLED CATCHMENT

Not for all catchments used in this research, data was available to calibrate and validate the models on. As a proxy, for uncalibrated areas parameters of a coupled neighbouring catchment were 'borrowed'. For some areas this led to acceptable results, but not for all. Below an example of catchment Umos3, in which particularly the HBV model shows a large outlier in this catchment with the 'borrowed' parameters.



UNCALIBRATED AREAS WITH MEDIAN VALUES

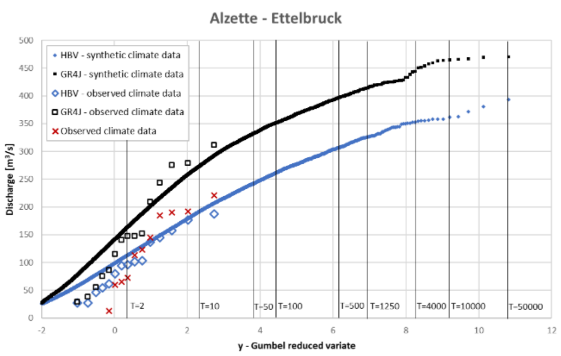
To circumvent these outliers, another approach was eventually chosen to handle catchments without data. Uncalibrated areas now get a median value for their parameter set, obtained from previous calibrated subcatchments within the catchment. As visible in the next graph, again for catchment Umos3 the discharge outlier in the HBV model is greatly reduced, from 13000m3/s, down to 7500m3/s. In this particular area there is still a difference between the HBV and GR4J models visible. Further study could focus specifically on handling data-scarce areas.

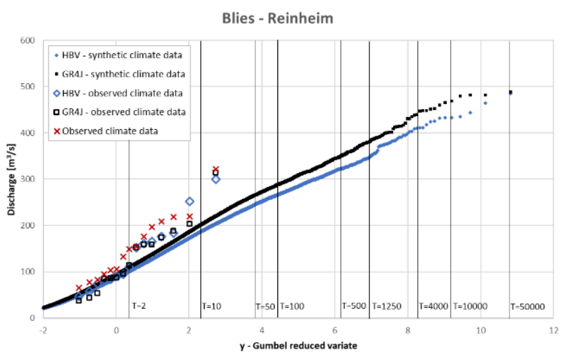


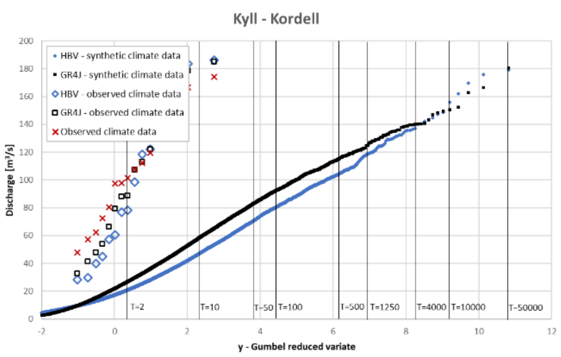
Umos3 - Cochem
HBV - synthetic climate data
GR4J - synthetic climate data
HBV - observed climate data
GR4J - observed climate data
Observed climate data
Discharge [m³/s]
y - Gumbel reduced variate
T=2
T=10
T=50
T=100
T=500
T=1250
T=4000
T=10000
T=50000
0
1000
2000
3000
4000
5000
6000
7000
8000
9000
-2
0
2
4
6
8
10
12

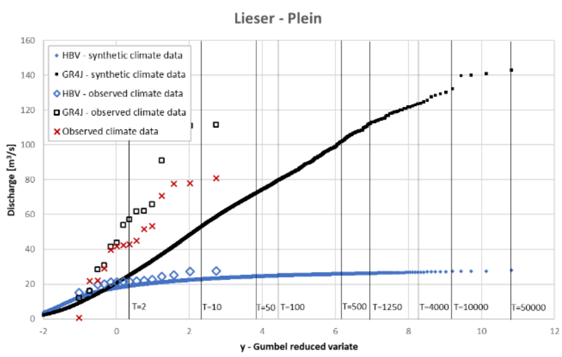
Appendix E

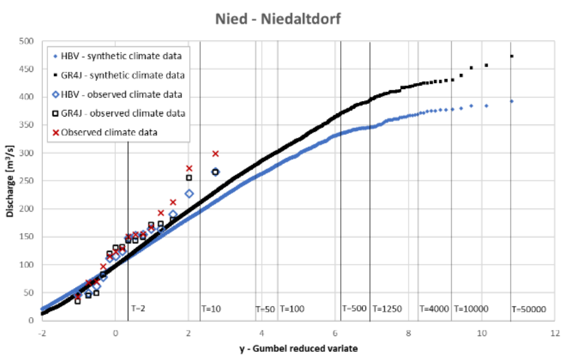
Flood frequency curves for all catchments

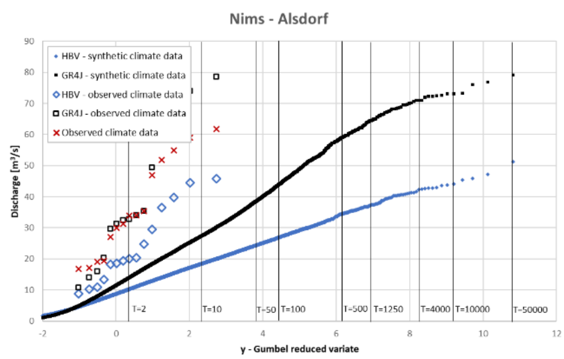


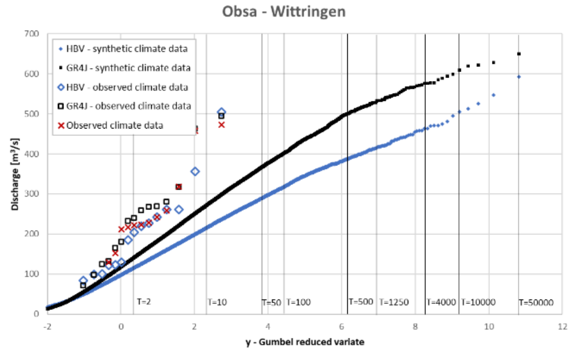


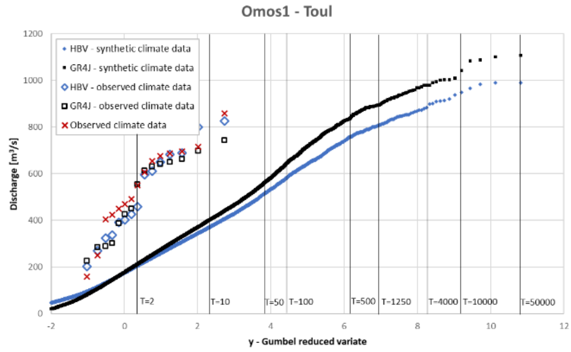


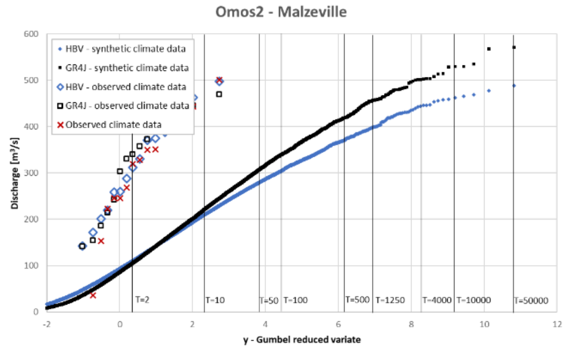


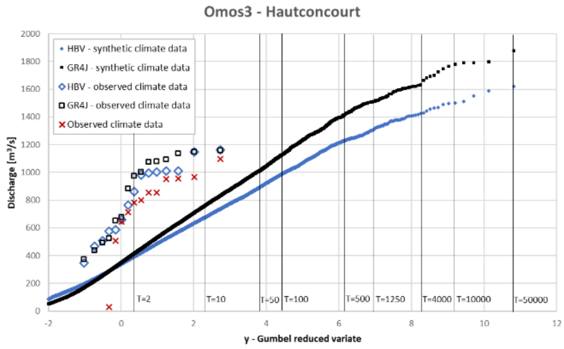


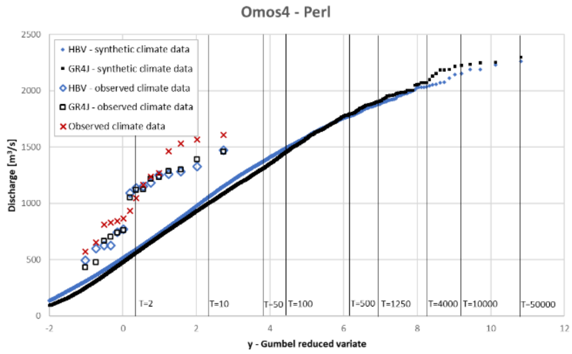


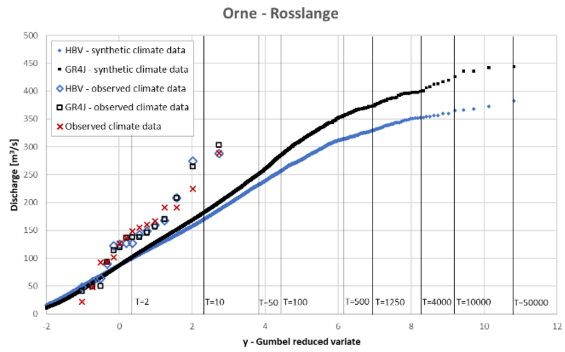

Obsa - Wittringen
HBV - synthetic climate data
GR4J - synthetic climate data
HBV - observed climate data
GR4J - observed climate data
Observed climate data
T=2
T=10
T=50
T=100
T=500
T=1250
T=4000
T=10000
T=50000
Discharge [m³/s]
y - Gumbel reduced variate

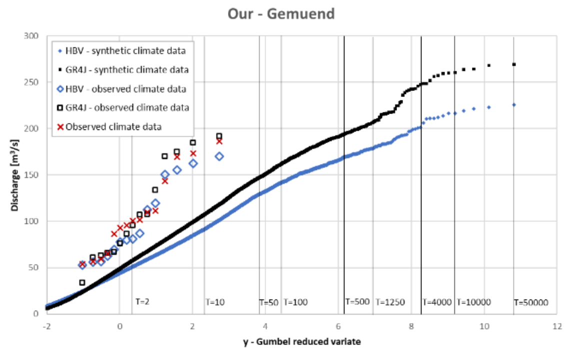

Omos1 - Toul
HBV - synthetic climate data
GR4J - synthetic climate data
HBV - observed climate data
GR4J - observed climate data
Observed climate data
T=2
T=10
T=50
T=100
T=500
T=1250
T=4000
T=10000
T=50000
Discharge [m³/s]
y - Gumbel reduced variate

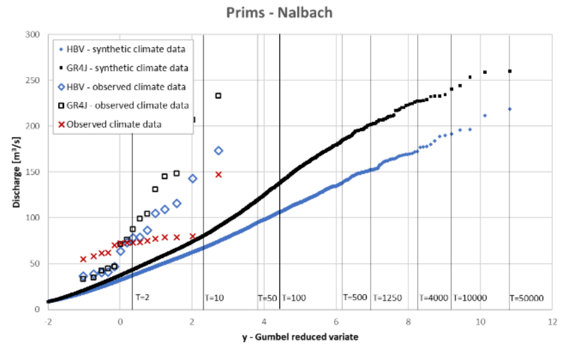

Omos2 - Malzeville
HBV - synthetic climate data
GR4J - synthetic climate data
HBV - observed climate data
GR4J - observed climate data
Observed climate data
T=2
T=10
T=50
T=100
T=500
T=1250
T=4000
T=10000
T=50000
Discharge [m³/s]
y - Gumbel reduced variate

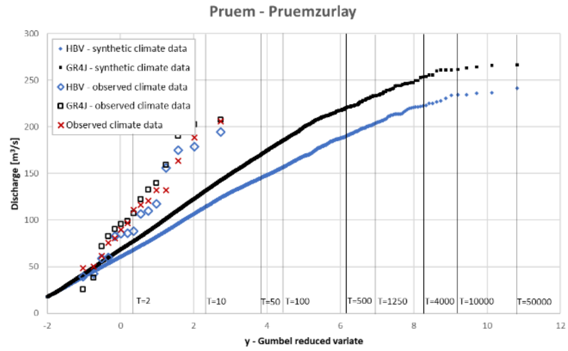

Omos3 - Hautconcourt
HBV - synthetic climate data
GR4J - synthetic climate data
HBV - observed climate data
GR4J - observed climate data
Observed climate data
T=2
T=10
T=50
T=100
T=500
T=1250
T=4000
T=10000
T=50000
Discharge [m³/s]
y - Gumbel reduced variate

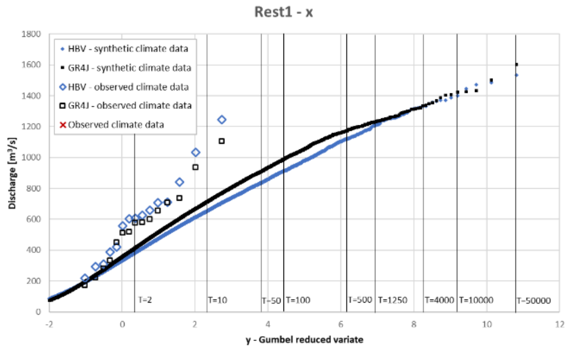

Omos4 - Perl
HBV - synthetic climate data
GR4J - synthetic climate data
HBV - observed climate data
GR4J - observed climate data
Observed climate data
T=2
T=10
T=50
T=100
T=500
T=1250
T=4000
T=10000
T=50000
Discharge [m³/s]
y - Gumbel reduced variate

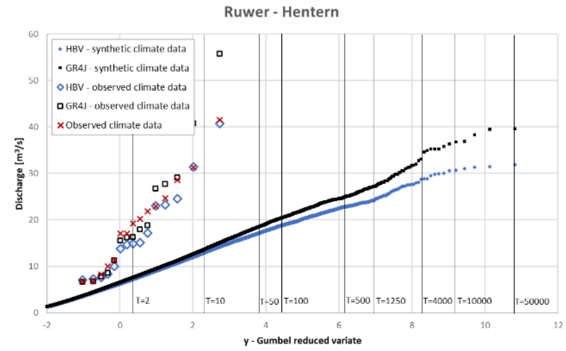

Orne - Rosslange
HBV - synthetic climate data
GR4J - synthetic climate data
HBV - observed climate data
GR4J - observed climate data
Observed climate data
T=2
T=10
T=50
T=100
T=500
T=1250
T=4000
T=10000
T=50000
Discharge [m³/s]
y - Gumbel reduced variate

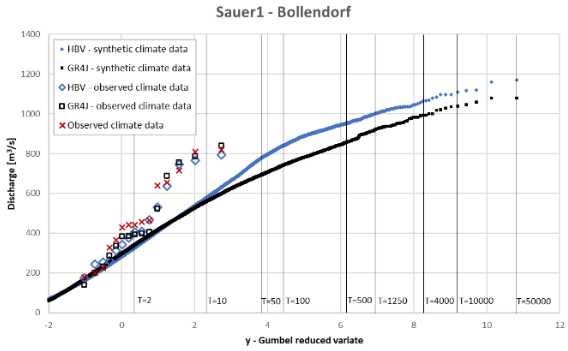

Our - Gemuend
HBV - synthetic climate data
GR4J - synthetic climate data
HBV - observed climate data
GR4J - observed climate data
Observed climate data
T=2
T=10
T=50
T=100
T=500
T=1250
T=4000
T=10000
T=50000
Discharge [m³/s]
y - Gumbel reduced variate

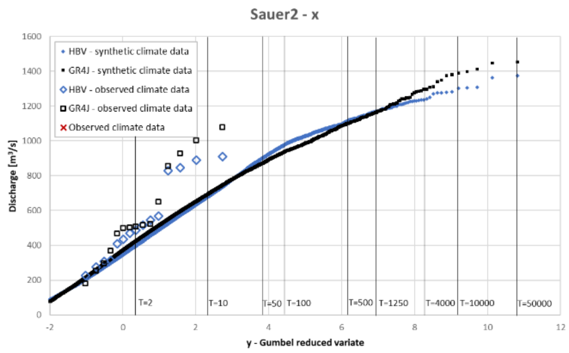

Prims - Nalbach
HBV - synthetic climate data
GR4J - synthetic climate data
HBV - observed climate data
GR4J - observed climate data
Observed climate data
T=2
T=10
T=50
T=100
T=500
T=1250
T=4000
T=10000
T=50000
Discharge [m³/s]
y - Gumbel reduced variate

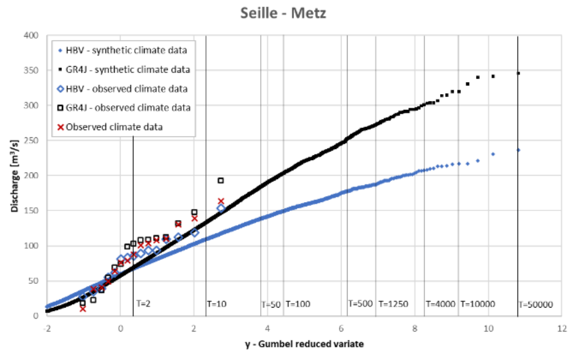

Pruem - Pruemzurlay
HBV - synthetic climate data
GR4J - synthetic climate data
HBV - observed climate data
GR4J - observed climate data
Observed climate data
Discharge [m³/s]
T=2
T=10
T=50
T=100
T=500
T=1250
T=4000
T=10000
T=50000
y - Gumbel reduced variate

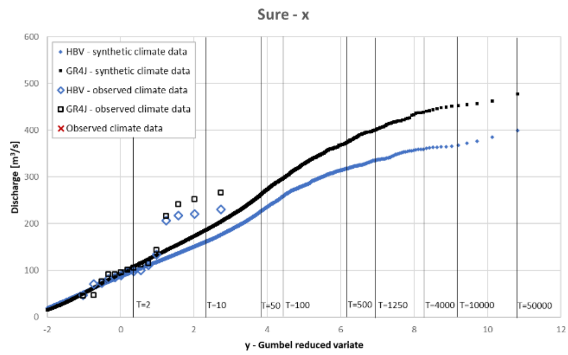

Rest1 - x
HBV - synthetic climate data
GR4J - synthetic climate data
HBV - observed climate data
GR4J - observed climate data
Observed climate data
Discharge [m³/s]
T=2
T=10
T=50
T=100
T=500
T=1250
T=4000
T=10000
T=50000
y - Gumbel reduced variate

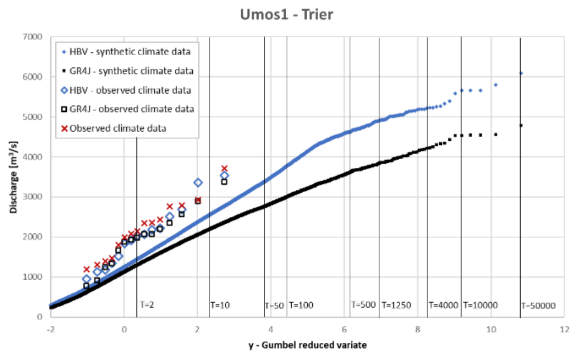

Ruwer - Hentern
HBV - synthetic climate data
GR4J - synthetic climate data
HBV - observed climate data
GR4J - observed climate data
Observed climate data
Discharge [m³/s]
T=2
T=10
T=50
T=100
T=500
T=1250
T=4000
T=10000
T=50000
y - Gumbel reduced variate

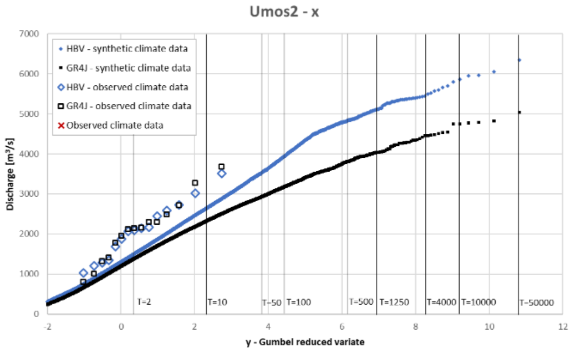

Sauer1 - Bollendorf
HBV - synthetic climate data
GR4J - synthetic climate data
HBV - observed climate data
GR4J - observed climate data
Observed climate data
Discharge [m³/s]
T=2
T=10
T=50
T=100
T=500
T=1250
T=4000
T=10000
T=50000
y - Gumbel reduced variate

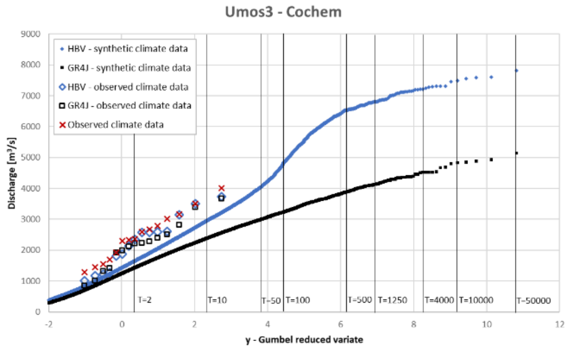

Sauer2 - x
HBV - synthetic climate data
GR4J - synthetic climate data
HBV - observed climate data
GR4J - observed climate data
Observed climate data
Discharge [m³/s]
T=2
T=10
T=50
T=100
T=500
T=1250
T=4000
T=10000
T=50000
y - Gumbel reduced variate

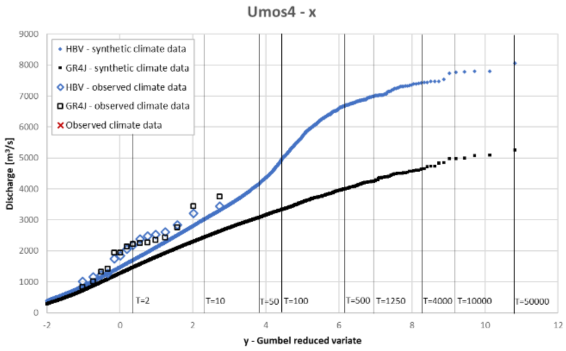

Seille - Metz
HBV - synthetic climate data
GR4J - synthetic climate data
HBV - observed climate data
GR4J - observed climate data
Observed climate data
Discharge [m³/s]
T=2
T=10
T=50
T=100
T=500
T=1250
T=4000
T=10000
T=50000
y - Gumbel reduced variate

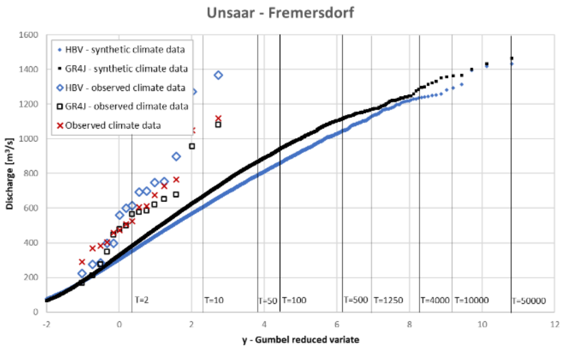

Sure - x
HBV - synthetic climate data
GR4J - synthetic climate data
HBV - observed climate data
GR4J - observed climate data
Observed climate data
Discharge [m³/s]
T=2
T=10
T=50
T=100
T=500
T=1250
T=4000
T=10000
T=50000
y - Gumbel reduced variate


Umos1 - Trier
HBV - synthetic climate data
GR4J - synthetic climate data
HBV - observed climate data
GR4J - observed climate data
Observed climate data
Discharge [m³/s]
T=2
T=10
T=50
T=100
T=500
T=1250
T=4000
T=10000
T=50000
y - Gumbel reduced variate


Umos2 - x
HBV - synthetic climate data
GR4J - synthetic climate data
HBV - observed climate data
GR4J - observed climate data
Observed climate data
Discharge [m³/s]
y - Gumbel reduced variate
T=2
T=10
T=50
T=100
T=500
T=1250
T=4000
T=10000
T=50000


Umos3 - Cochem
HBV - synthetic climate data
GR4J - synthetic climate data
HBV - observed climate data
GR4J - observed climate data
Observed climate data
Discharge [m³/s]
y - Gumbel reduced variate
T=2
T=10
T=50
T=100
T=500
T=1250
T=4000
T=10000
T=50000


Umos4 - x
HBV - synthetic climate data
GR4J - synthetic climate data
HBV - observed climate data
GR4J - observed climate data
Observed climate data
Discharge [m³/s]
y - Gumbel reduced variate
T=2
T=10
T=50
T=100
T=500
T=1250
T=4000
T=10000
T=50000


Unsaar - Fremersdorf
HBV - synthetic climate data
GR4J - synthetic climate data
HBV - observed climate data
GR4J - observed climate data
Observed climate data
Discharge [m³/s]
y - Gumbel reduced variate
T=2
T=10
T=50
T=100
T=500
T=1250
T=4000
T=10000
T=50000