Deltares

Enabling Delta Life

Generalised Likelihood Uncertainty Estimation for the daily HBV model in the Rhine Basin, Part A: Germany

Generalised Likelihood Uncertainty Estimation for the daily HBV model in the Rhine Basin

Part A: Germany

Hessel Winsemius
Willem van Verseveld
Albrecht Weerts
Mark Hegnauer

1207771-003

Deltares

Title
Generalised Likelihood Uncertainty Estimation for the daily HBV model in the Rhine Basin

Client	Project	Reference	Pages
Rijkswaterstaat (WVL)	1207771-003	1207771-003-ZWS-0018	81

Keywords
GRADE, GLUE analysis, parameter uncertainty estimation

Summary
This report describes the derivation of a set of parameter sets for the HBV models for the German part of the Rhine basin covering the catchment area between Basel and Lobith, including the uncertainty in these parameter sets. These parameter sets are required for the project "Generator of Rainfall And Discharge Extremes (GRADE)". GRADE aims to establish a new approach to define the design discharges flowing into the Netherlands from the Meuse and Rhine basins. The design discharge return periods are very high and GRADE establishes these by performing a long simulation using synthetic weather inputs. An additional aim of GRADE is to estimate the uncertainty of the resulting design discharges. One of the contributions to this uncertainty is the model parameter uncertainty, which is why the derivation of parameter uncertainty is required.

Parameter sets, which represent the uncertainty, were derived using a Generalized Likelihood Uncertainty Estimation (GLUE), which conditions a prior parameter distribution by Monte Carlo sampling of parameter sets and conditioning on a modelled v.s. observed flow in selected flow stations. This analysis has been performed for aggregated sub-catchments (see also Figure 2.1[1]) separately using the HYRAS 2.0 rainfall dataset and E-OBS v4 temperature dataset as input and a discharge dataset from the German Federal States, combined with the HYMOG dataset as flow observations. To ensure that the conditioned parameter sets are suitable for the high flow domain, additional performance measures were introduced which reflect the behaviour of a parameter set in the high flow domain. It was assumed that precipitation corrections were not required. Furthermore the modelled flow contributions from the "Zwischeneinzugsgebieten" (intermediate basins between the larger tributaries and the main stem of the Rhine) were set on zero. This enabled investigation into the significance of the proportion of water, coming from these areas with respect to the total river flow.

The GLUE analysis showed good results over most of the aggregated sub-catchments. Areas which showed a lower performance were the Erft, parts of the Main and the upper Rhine basins. The differences in the Erft can be explained by the fact that most of the Erft discharges are affected by lignite mining industry. Differences in the other basins can be explained by the fact that the hydrology in these basins is likely to be dominated by processes that occur on a smaller time scale than the (daily) model time step.

After the GLUE analysis, a small selection of parameter sets, representative for the distribution of the parameter sets, was selected from the conditioned parameter sets of each of the aggregated sub-cathcments and combined into 5 representative parameter sets for the whole area. A simulation over the complete Rhine basin shows that the uncertainty, encapsulated by the 5 parameter sets encapsulates the observations quite well.

[1] Note that the figure also includes the two aggregated sub-catchments in Switzerland. These aggregated sub-catchments are not part of this study, but are presented in an other report (Verseveld (2013)).

Deltares

Title
Generalised Likelihood Uncertainty Estimation for the daily HBV model in the Rhine Basin

Client	Project	Reference	Pages
Rijkswaterstaat (WVL)	1207771-003	1207771-003-ZWS-0018	81

The contribution of the smaller intermediate catchment areas (Zwischeneinzugsgebieten, ZWEs) in between the aggregated sub-catchments considered in the GLUE analysis, and the Rhine river itself, was demonstrated to be small. At Lobith, high flows are even slightly overestimated by the HBV model. This may be caused by the fact that flow peak attenuation due to retention areas or floods in upstream areas are not considered in HBV. In GRADE, a SOBEK model, which includes such retention and flooding effects will be used.

From this study it is recommended that the effect of SOBEK on peak flow simulations of GRADE is investigated in a further study, that the unaccounted flow from the ZWEs is in the short term accounted for through a correction factor or a simple groundwater outflow model and that model uncertainties in the Swiss part of the HBV model are analysed using the GLUE method as well. Finally, it was recommended that the behaviour of the ZWEs during extreme flows is analysed in more detail in the long term.

Version	Date	Author	Initials	Review	Initials	Approval	Initials
	dec. 2013	Hessel Winsemius		Micha Werner		Gerard Blom	
		Willem van Verseveld					
		Albrecht Weerts					
		Mark Hegnauer					

State
final

Contents

Appendices

1 Introduction

Within the framework of the “Generator of Rainfall And Discharge Extremes, ample effort is being put into the derivation of uncertainty estimates of the Rhine and Meuse design discharges. GRADE aims to establish a new approach to define the design discharges flowing into the Netherlands from the Meuse and Rhine basins. The design discharge return periods are very high and GRADE establishes these by performing a long simulations with a hydrological model cascade using synthetic weather inputs. For an overview of the GRADE methodology, we refer to De Wit and Buishand (2007). A recent review of the current state of GRADE has been performed by Ogink (2011). Naturally, the synthetic weather inputs as well as the model cascade (consisting of a daily hydrological model in the HBV software and a hydraulic model in the SOBEK software) are subject to uncertainties.

In this study, we derived parameter uncertainty for the hydrological model HBV used in GRADE for the German part of the Rhine basin (between Basel and Lobith). The Swiss part of the basin is considered separately, first of all because the discharge data became available later for this part of the basin to perform a parameter uncertainty analysis, and second because the lakes and reservoirs in the Swiss part were not explicitly accounted for yet in the daily HBV model. An uncertainty analysis for the Swiss part of the Rhine basin, including the Swiss lakes is described in Verseveld (2013). HBV is a hydrological model software by SMHI (Lindström et al., 1997) which is here run on a daily basis. For GRADE use is made of the original HBV-96 software.

To estimate (as part of the full uncertainty) the uncertainty as a result of the hydrological model parameter choice, a Generalized Likelihood Uncertainty Assessment (GLUE) has been recommended during a previous review (Weerts and Van der Klis, 2006). A GLUE analysis is used to asses and reflect the uncertainty, contained in the selection of hydrological model parameters. In GRADE such an analysis may be used to assess the effect of parameter uncertainty on the design discharge for the Rhine. By performing a GLUE analysis, one accepts the presence of multiple acceptable parameter sets, instead of a single optimal parameter set. For GRADE, this means that ones multiple parameter sets are considered, not a single value for the design discharge, but a range of discharges with different peak values and different shapes of the flood wave as a result of parameter uncertainty, may be provided.

For the HBV model of the Meuse, a GLUE analysis was already performed (Kramer and Schroevers, 2008; Kramer et al., 2008). This analysis resulted in the selection of 5 behavioural parameter sets for each of the 15 HBV subcatchments of the Meuse, which provide a range of extreme value distributions for discharges at Borgharen. These parameter sets were conditioned on a long time series of measured precipitation, temperature and potential evaporation as inputs, and discharges throughout the Meuse basin as outputs. Parameter sets were marked as behavioural if they showed a good resemblance with the full hydrograph, as well as good performance in reproducing peak discharges.

The result of the analysis of the Rhine basin is (similar to the Meuse case), a set (5) of behavioural parameter sets of the HBV hydrological model. A GLUE analysis on the Rhine is far more complex than a GLUE analysis on the Meuse, because of the large amounts of sub-basins. This report describes the GLUE experiment for the sub-basins that contribute to the Rhine between Basel and Lobith. The method used to deal with the large size of the basin is explained in Chapter 2, with a summary of results of the GLUE analysis given in Chapter 3. The approach to select a representative sample to use in GRADE is outlined in Chapter 4. Chapter 5 describes the water balance of the flow along main stem of the Rhine using the newly derived parameter sets. Chapter 6 describes which assumptions and limitations the analysis has and what the effect is on the results. In Chapter 7, we conclude on the uncertainty analysis using GLUE and recommend on potential improvements of the model.

2 Approach

2.1 Generalized Likelihood Uncertainty Estimation (GLUE) analysis

2.1.1 GLUE in general

Working with (complex) models with many parameters introduces the problem of equifinality. This is the effect that multiple parameter sets give approximately the same results. The question is therefore whether one should look for the “best” parameter set. The philosophy of the Generalized Likelihood Uncertainty Estimation (GLUE) is that instead of finding one optimal parameter set, multiple behavioural parameter sets are accepted as a possible realisation of the hydrology in a catchment. By selecting one or multiple likelihood measures (e.g. Nash-Sutcliff, or Relative Volume error), the parameter sets are analysed on their performance. Only the parameter sets that meet the constraints of the Likelihood measure are selected as “behavioural sets”.

The steps of a GLUE analysis are generally as follows:

1) Define the parameters that are to be evaluated (i.e. which are assumed to be unknown a priori).
2) Select a performance measure.
3) Perform a Monte-Carlo simulation on the selected unknown parameters with a sufficient amount of samples. For every run, a set of parameters is randomly selected from a pre-defined uniform distribution of each parameter (all dots in Figure 2.1).
4) Analyse the performance of all selected parameter sets for the selected performance measure.
5) Select ‘behavioural’ parameter sets. These are parameter sets which give a performance above a user-defined threshold. This is one of the subjective steps in the GLUE analysis
6) Rescale the performance measure of each behavioural parameter set into a likelihood (zero likelihood where the parameter is equal to the performance measure threshold value) so that the sum of all likelihood values equals one.

By applying the GLUE analysis an estimate for the model parameter uncertainty is given. The number of approved parameter sets can then be seen as a value for the uncertainty. The more approved parameter sets there are, the lower the uncertainty is.

2.1.2 GLUE for the Rhine

Although GLUE is relatively straightforward, applying it for a large catchment such as the Rhine does pose some challenges. In the section below we describe these challenges and how we have dealt with them:

1. The daily Rhine HBV model consists of 148 subbasins. For GLUE, a Monte-Carlo sampling must be performed. This means that the model should be run a large number of times. With 148 subbasins, this will result in large computational cost. Therefore the Rhine HBV model has been divided into a number of large sub-catchments (e.g. Main, Neckar) and GLUE has been performed for each particular large sub-catchment. We used the same subdivision as used by SMHI in their

calibration report (Berglov et al., 2009). An overview of the large sub-catchments of the Rhine is given in Figure 2.1.

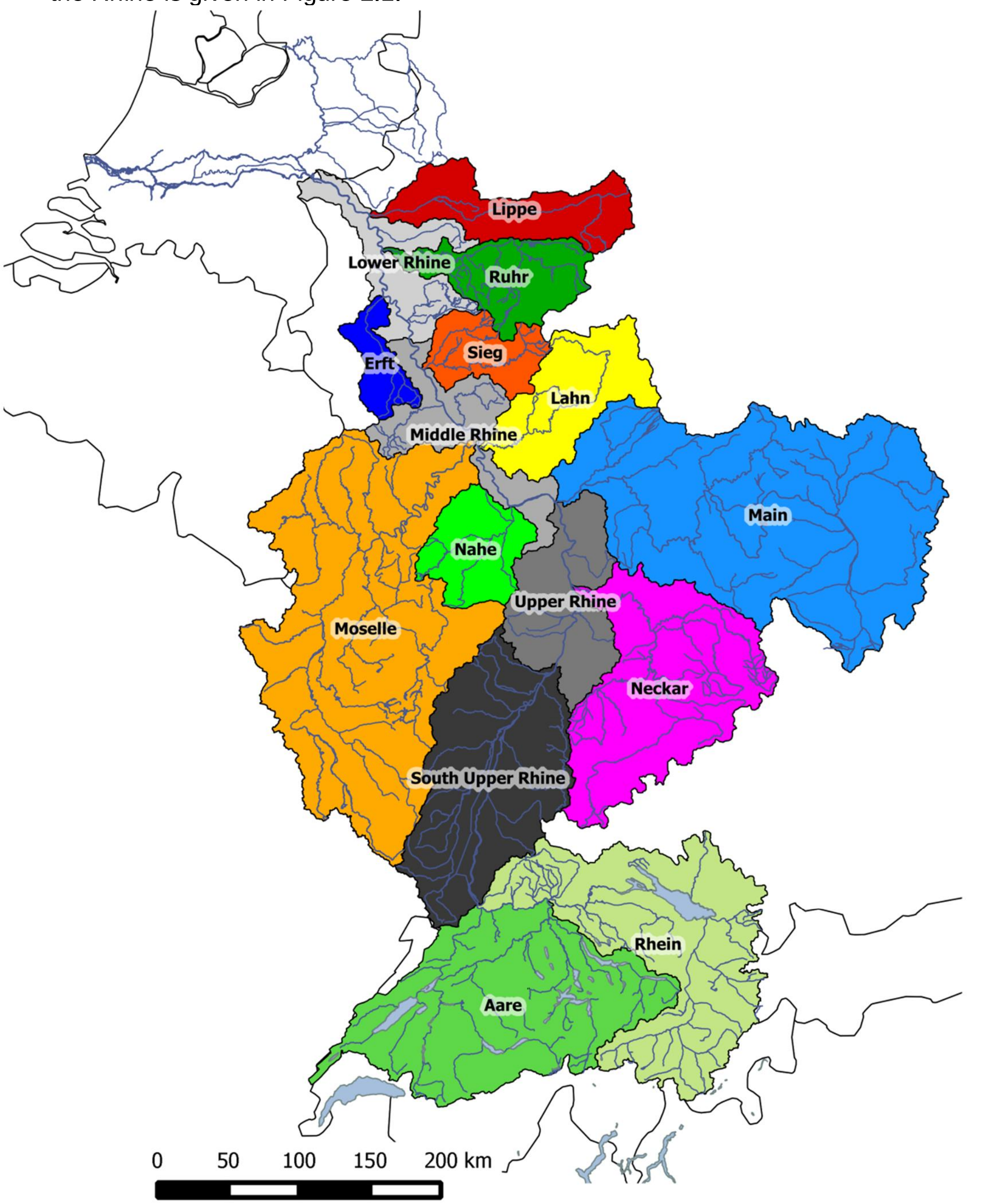


Figure 2.1 Overview of the large sub-catchments of the Rhine basin

2. Instead of constraining the parameter uncertainty on one gauge only, we wish to constrain the model in many places in the basin, wherever we have reliable discharge series available. Therefore the GLUE analysis is done first for the most upstream HBV units at any place where a discharge measurement series is available. To constrain in the more downstream basins, a random selection from the behavioural sets in the upper basins has been performed, and GLUE applied to the intermediate basin. In this way, only the behavioural (approved) parameter sets from the upper

basins were passed to the downstream basin. This process has been performed until the most downstream located gauged basin. The process has been schematized in Figure 2.2.

3. The original daily HBV Rhine model contains many correction factors for precipitation and corrections which translate outflows from the Rhine's subcatchments into inflows into the Rhine River itself. The precipitation corrections are likely to be the effect of measurement uncertainty and/or undersampling in calibration data, used to construct the daily HBV model. We use a new rainfall database in this analysis. Therefore we have removed all correction factors at the beginning of the experiment, assuming that precipitation correction is not required. In any case, the precipitation correction factors which were part of the HBV parameters in the original parameter set are based on previous calibration studies and not valid for the new forcing datasets.
4. Instead of one performance measure, we have used multiple performance measures. This has been done to ensure that not only the overall hydrograph shape is simulated satisfactorily, but also the extreme values. This compromises the classical GLUE approach in that an unambiguous scaling of the performance measure into a likelihood cannot be done. Therefore we have assumed that each behavioural parameter set is equally likely.
5. The Rhine basin contains ungauged areas in between sub-catchments and the Rhine River itself. These intermediate basins largely schematize the Rhine valley. There is insufficient information available to constrain parameter sets of the associated HBV units here. We have therefore set the lateral flow from the valley on zero. This can be altered if too large volume errors are experienced when running the model over the full basin. This check is demonstrated in Chapter 5.

The result is a set of behavioural parameter sets for each sub-basin.

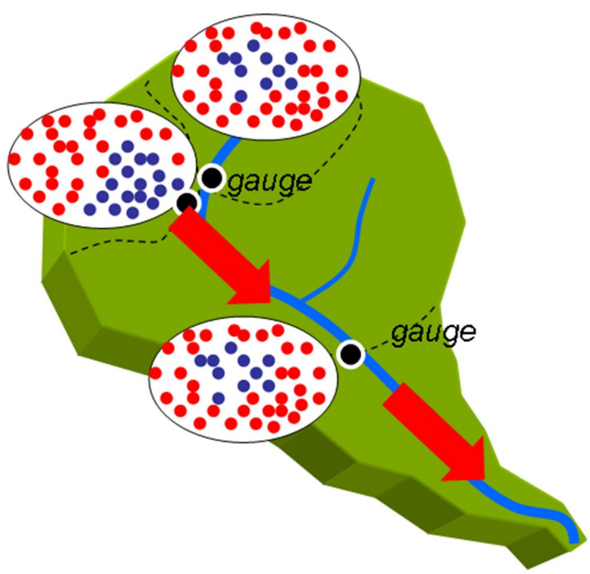


Figure 2.2 Schematic showing the method of the GLUE analysis for a series of basins. The red dots within the circle represent the samples taken from the prior uniform parameter distribution. The blue dots are the selected behavioural parameter sets, given the likelihood measures, derived from the gauged location. The blue sets are passed on to the neighbouring downstream area, which is consequently constrained on the more downstream located gauge

2.2 Input/output data

2.2.1 Meteorological forcing

As input data, we used the HYRAS 2.0 rainfall (Rauthe et al., 2012). HYRAS 2.0 is a gridded dataset (0.25 degree resolution) of rainfall over Germany, containing a large set of observations for the period 1955-2006. This dataset has been generated and quality checked by the German Weather Service and is therefore treated *as is* in this study. For temperature, we have relied on KNMI's 0.25 degree gridded E-OBS version 4.0 (Haylock et al., 2008) , containing a large set of observations for the period 1955-2006 This dataset has also been applied *as is*.

2.2.2 Discharge measurements

We used a collection of discharge "measurements" (discharge data mostly derived from measured waterlevel and calculated using Q-H-relations) for our GLUE analysis. The collection is a merge of the corrected data series from the HYMOG dataset and data, collected by the BfG from the German Federal States to be used for the re-calibration of the HBV-model by the SMHI in 2009 (see also Berglov et al., 2009). Prior to consideration for calibration, a rigorous data screening of the latter set has been performed. Where overlap occurred between HYMOG records and BfG records, the HYMOG records were prioritized. For discharge stations where there were no HYMOG data available, the BfG data were used. This resulted in a dataset containing data for all stations needed for the calibration (if available) for the period 1989-2006.

Data was screened by plotting station records from upstream to downstream and rigorously checking whether the amount of water from upstream to downstream was accumulating properly. Sometimes zeros were found instead of missing numbers or strange periods with offsets. These were all removed. Without going into detail, the cleaned records were used to select appropriate groups of HBV subcatchments to calibrate together based on a certain station. The groups of sub catchments considered per station and per tributary, and the cleaning which was applied to the datasets of some of the catchments, is given in Appendix 8A.

The grouping shows that in some parts, the GLUE analysis can be performed in a lot of detail, because many stations are available in a certain sub-catchment, while in other regions, the detail is quite low and many basins are calibrated together with only one station. This inconsistency in detail is unavoidable given the data available.

2.3 Parameter treatment and range

HBV uses many parameters to simulate discharge in response to rainfall. Mathematically, each parameter gives an additional degree of freedom and therefore also more risk of equifinality. Many of the parameters are such that they can be expected to be strongly correlated. For instance, parameters which represent a time scale (in particular the routing parameters of the fast (HQ, KHQ, alpha) and slow (K4, perc) responding reservoir) may easily compensate for each other, meaning that the effect of a wrong value for one of them, can be compensated for by another wrong value for another parameter. To prevent unnecessary correlation problems, a number of parameters have been fixed, following the procedures, outlined below. Other parameters, for which discharge is particularly sensitive, have been sampled which is also outlined below:

K4: slow recession:
The recession (unforced groundwater outflow) of a catchment is often schematised as a 'linear reservoir'. In HBV, this is equivalent with the outflow from the slow reservoir. This outflow is modelled in HBV as:

$$Q_s(t) = K_4 S_s(t) \quad (1.1)$$

Where Q_s [L/T] is the flow from the slow reservoir, K_4 [1/T] is the linear outflow coefficient (reciprocal of residence time) of the reservoir and S_s [L] is the storage in the slow reservoir. In periods with no rainfall, K_4 can be read from the recession curve section of the hydrograph by plotting on log-scale and estimating the slope. The slope is equal to K_4. An expected correlated parameter is *perc*, which conceptualises percolation to the deeper reservoir of HBV (reservoir, assumed to be correlated with the groundwater table). By fixing the parameter K_4, the parameter *perc* can be estimated more accurately in the GLUE setup.

HQ: fast flow related parameter:
HQ, KHQ and alpha are all together determining the outflow from the fast reservoir of HBV. HQ is a somewhat strange parameter in that it mathematically correlates very strongly to KHQ. Therefore HQ was fixed, assuming it was equal to the 90% percentile of flow probability, expressed in units of mm/day. A similar approach to the fixing of HQ has been presented in HBV manual (SMHI, version 4.5).

After fixing the above parameters, we have selected a limited number of sensitive parameters to include in the GLUE analysis. There are many more parameters, such as for instance snow related parameters. These have not been considered in the analysis and were fixed on default values instead. There is too few information content in discharge alone to estimate the uncertainty of these parameters from a GLUE analysis. Note that in the Swiss part of the basin this may be different. Here, for instance snow is much more important and therefore should be included in the sampled parameters. Table 2.1 shows the considered parameters and ranges. The prior ranges are based on the Meuse GLUE analysis (Kramer and Schroevers, 2008; Kramer et al., 2008).

Table 2.1 Standard parameter ranges for GLUE parameters

Parameter	Unit	Minimum	Maximum
fc	mm	100	500
lp	-	0.3	1.0
Beta	-	1.0	3.0
alpha	-	0.2	1.2
KHQ	1/day	0.05	0.2
perc	mm/day	1	4

Fc:	Maximum value of the soil moisture storage	(mm)
Lp:	Limit for potential evaporation	(-)
Beta:	Control for the increase in soil moisture for every mm of precipitation	(-)
Alpha:	Parameter for the non-linear behaviour in the response function	(-)
KHQ:	Recession parameter at HQ (high flow parameter)	(1/day)
Perc:	Percolation	(mm/day)

Table 2.2 Values for fixed parameters

Parameter	Unit	Value
K4	1/day	Varies per basin
Tcalt	°C/hm	0.60
Cfmax		3.50
Tt	°C	0.00
Tti		2.00
Cfr		0.05
Whc		0.10
Fosfcf	-	0.80
Focfmax	-	0.60
Etf		0.05
Cevpfo	-	1.20
Icfo		4.00
Icfi		1.50
Cevpl	-	1.10

For catchments with limited or no behavioural parameter sets an analysis is done on the parameter ranges. If there is an indication that the majority of the behavioural sets is not within the original range, the range is extended somewhat for specific basins. This extension has been restricted to the fc, lp and the KHQ parameters, as these have the strongest relation with the physical characteristics of the basin. A sensitive parameter, which was not included in the analysis is MAXBAS. MAXBAS is a routing parameter and simulates the lag and attenuation occurring throughout the HBV unit considered. We have kept the MAXBAS parameter values of the original daily model and have only adapted in some cases MAXBAS where the results showed that there is a clear timing discrepancy between modelled and observed flows. Wherever MAXBAS has been adapted, this is described in the results.

In Table 2.3 the ranges used for each catchment are listed. The bold numbers indicate that a value different from the default range was used.

Table 2.3 Adjusted ranges for different sub-basins

	fc		lp		alpha		Beta		KHQ		perc	
Catchment	*min*	*max*	*min*	*max*	*min*	*max*	*min*	*max*	*min*	*max*	*min*	*max*
Neckar	**50**	500	**0.25**	**0.90**	0.2	1.2	1	3	0.05	**0.30**	1	4
Moselle	**50**	500	**0.25**	**0.90**	0.2	1.2	1	3	0.05	**0.25**	1	4
Lahn	100	500	0.30	1.00	0.2	1.2	1	3	0.05	0.20	1	4
Ruhr	100	500	**0.25**	**0.90**	0.2	1.2	1	3	0.05	0.20	1	4
Lippe	100	500	**0.25**	**0.90**	0.2	1.2	1	3	0.05	0.20	1	4
UpperRhine	**50**	**600**	**0.10**	**0.90**	0.2	1.2	1	3	**0.01**	0.20	1	4
MidRhine	**50**	**600**	0.30	1.00	0.2	1.2	1	3	0.05	**0.30**	1	4
Nahe	**50**	500	0.30	1.00	0.2	1.2	1	3	0.05	**0.30**	1	4
S.UpRhine	100	500	0.30	1.00	0.2	1.2	1	3	0.05	0.20	1	4
Erft	100	500	**0.25**	**0.90**	0.2	1.2	1	3	0.05	0.20	1	4
Sieg	100	500	0.30	1.00	0.2	1.2	1	3	0.05	**0.30**	1	4
Main	**50**	500	**0.25**	1.00	0.2	1.2	1	3	0.05	**0.30**	1	4
LowerRhine	100	500	0.30	1.00	0.2	1.2	1	3	0.05	0.20	1	4

2.4 Performance measures

We used the following performance measures to distinguish behavioural from non-behavioural parameter sets:

- Nash and Sutcliffe efficiency. This performance measure normalises the squared residuals of the observed minus simulated time series and is a measure for the overall performance, with an emphasis on errors at high flows. A score of 1 means a perfect fit with the observations, while a value of zero means that the average of the observed is an equally good predictor of discharge as the modelled series. Nash and Sutcliffe efficiency is computed as follows:

$$L_{nse} = 1 - \left[\frac{\sum_t \left[Q_s(t) - Q(t) \right]^2}{\sum_t \left[Q(t) - \overline{Q} \right]^2} \right], \tag{1.2}$$

  where L_{nse} [-] is the Nash and Sutcliffe efficiency, and Q_s and Q are simulated and observed discharge respectively [L^3 T^{-1}]. t represent the time step. Parameter sets should have a Nash and Sutcliffe efficiency value, equal to at least 90% of the highest value obtained during the Monte-Carlo simulation.
- Relative volume error. This score evaluates the long-term volumetric error. This is computed as:

$$L_{rev} = \frac{\sum_t \left[Q_s(t) - Q(t) \right]}{\sum_t Q(t)} \tag{1.3}$$

  where L_{rev} [-] is the relative volume error. Behavioural parameter sets should have a relative volume error, smaller than 0.1 (i.e. 10%).

- Relative Extreme Value Error. This error measures the deviation of observed and simulated extreme values. This is computed as follows:

$$L_{reve} = \frac{Q_s(T) - Q(T)}{Q(T)} \qquad (1.4)$$

where L_{reve} [-] is the relative extreme volume error, $Q_s(T)$ is the simulated extreme value of discharge for a return period T, based on an extreme value distribution, fitted through the simulated series Q_s, and $Q(T)$ is the same, but for the observed series. We applied Eq. (1.4) on two return periods being 5 years and 20 years, as well as using two extreme value distribution functions (Gumbel and GEV). We did not select higher return periods in order to prevent putting too strong confidence in the fitted extreme value distributions (in fact, GRADE is meant to provide the high return period discharges, as a replacement of such overfitting procedures). We selected parameter sets with a L_{reve} smaller than 0.1 as behavioural. The use of this performance measure ensures that the selected parameter sets have a good performance during extreme discharges.

In chapter 3), the results of the GLUE analysis are presented. In the tables, values for the performance measures are included. The standard performance measures are presented in Table 2.4. In the results tables in chapter 3), the performance measures that deviate from the standard values are printed in bolt.

Table 2.4 Performance measures standard values and how they are read from the results tables in chapter 3)

Performance measure	**In results table**
Nash and Sutcliff efficiency > 90%	Thres_R2 > 0.9
Relative Volume Error < 5%	Thres_REV < 0.05
Relative Extreme Value Error < 10%	Thres_T5 < 0.1
	Thres_T10 < 0.1

Note that the chosen acceptance thresholds are rather subjective. This subjectivity is inherent in the GLUE methodology. In theory, a statistical test could be performed to judge whether the choice of the chosen threshold was adequate, by evaluating the number of observation points, that remain within the uncertainty bounds. If e.g. a 90% uncertainty bound is expected, then 90% of the observations should be bracketed by the uncertainty bounds, created by the selected behavioural parameter sets. However, because the GLUE parameter sets are to be used for GRADE, the particular interest is on high flow periods. This test should then be performed on high flow periods only. The accurate estimation of flows during high flow conditions is particularly sensitive to the quality of the rainfall and therefore such a test over only short high flow periods may render unreliable. We therefore have chosen to subjectively judge the uncertainty bounds generated by the above criteria, and have loosened the criteria in case the uncertainty generated by them, was deemed too small.

2.5 Establishing a GLUE experiment in OpenDA

To perform the GLUE analysis, the OpenDA framework was used. OpenDA is an open software which allows a user to perform conditioning of model states or parameter sets based on observations. This can be done in a historic mode (i.e. calibration) or real-time mode (i.e. data assimilation). More details about OpenDA can be found on http://www.openda.org. A monte carlo framework has been added to OpenDA for this project to allow random sampling from uniform parameter ranges, as well as predefined parameter sets. The assessment and selection of parameter sets based on the performance measures has been done in Matlab.

Each tributary to the Rhine has a number of gauging stations in different subcatchments that could be used. The screened and cleaned data was used to make OpenDA GLUE setups along the schematics given in Appendix A. The setup of a OpenDA setup is not trivial. Therefore, to ensure that this process can be repeated for other basins in a later stage, this procedure has been extensively described in Appendix B. This appendix can be used as a manual for deriving an OpenDA setup for an HBV model.

3 Results per subcatchment

The results of the GLUE analysis are presented below. For each sub-basin, any particularities about the GLUE experiment setup are mentioned. Then the results are tabelised and results discussed.

3.1 Neckar

The Neckar has been analysed in three parts. The first part treats all the independent HBV units in the upstream part of the basin. The second part treats the HBV units that receive inflows from the HBV units analysed in the first part. The third part treats the HBV units furthest downstream, which depend on all the aforementioned units. The treatment of all HBV units is summarised in Table 3.1. HBV units that were calibrated together to one station are given in one box. Enz1 and Enz2 were calibrated together to one station since only one station was available at the outlet of Enz2 (Enz1 flows into Enz2). Rems and Murr, two neighbouring catchments with similar characteristics were calibrated together to one station. Only data at the outlet of Rems (Neustadt) were considered reliable. Elsenz and Neckar5 represent the drainage areas downstream of Rockenau. The discharge from these HBV subcatchments is set on zero. This has been implemented by setting ‘pcorr’ (precipitation correction factor) to zero.

Table 3.1 Calibration setup for the Neckar

Calibration experiment	HBV Units included	BfG station calibration
Neckar1	Enz1 Enz2	Pforzheim
	Fils	Plochingen/Fils
	Neckar1	Horb
	Rems Murr	Neustadt
	Jagst	Untergriesheim
	Kocher	Stein
Neckar2	Neckar2	Plochingen/Neckar
Neckar3	Neckar3	Rockenau
	Neckar4	
	Elsenz, Neckar5	**Not calibrated**

In Table 3.2 the criteria thresholds are summarized. The thresholds are selected in a way that there are at least 15 – 20 behavioural parameter sets. This means that in some cases the threshold value has to be increased.

Table 3.2 Selection criteria for the Neckar

	Thres_R2	Thres_REV	Thres_T5	Thresh_T20	Nr. of sets
Enz1, Enz2	0.9	0.05	**5**	**5**	**15**
Fils	0.9	0.05	**0.2**	**0.2**	**19**
Neckar1	0.9	**0.1**	0.1	0.1	61
Rems, Murr	0.9	0.05	0.1	0.1	**15**
Jagst	0.9	0.05	0.1	0.1	113
Kocher	0.9	**0.1**	**0.2**	**0.2**	81
Neckar2	0.9	0.05	**0.2**	**0.2**	23
Neckar3, Neckar4	0.9	0.05	0.1	0.1	122

The observations add Pforzheim for the Enz1 and Enz2 basins are not representative. It would appear that discharges above 100 m^3/s are not measured correctly and the signal becomes very irregular. This is shown in Figure 3.1. Note that the incorrect values were removed from the measurement series during the screening process, before using the series for conditioning of parameter sets in the GLUE analysis.

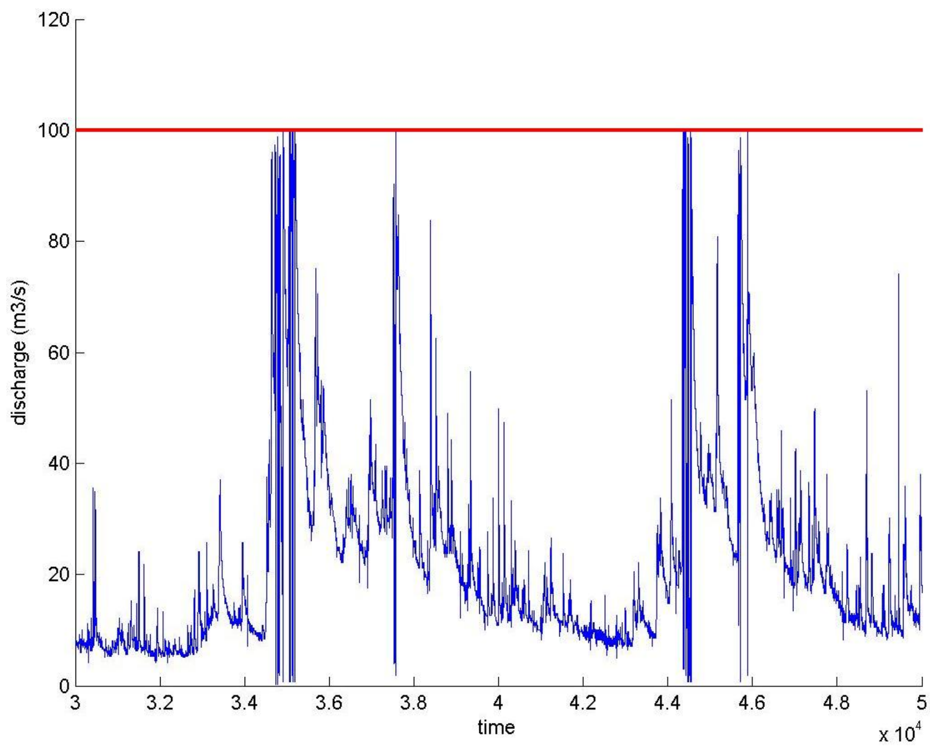


Figure 3.1 Plot representing the observed discharge at Pforzheim. Values above 100 (m3/s) are incorrectly given as near zero values

The analysis for Enz1 and Enz2 was done using the volume error and the Nash and Sutcliffe criterion only, because the measured data are lacking the peak discharges.

The Neckar2 experiment is done using the behavioural parameters sets from the first experiment and the Neckar3 experiment is done using all behavioural sets from the first two experiments.

The results for the Neckar3 experiment are good. The optimum Nash-Sutcliff value is 0.89. In Figure 3.2, the modelled hydrographs for the 1993 and 1995 events are plotted with the observations.

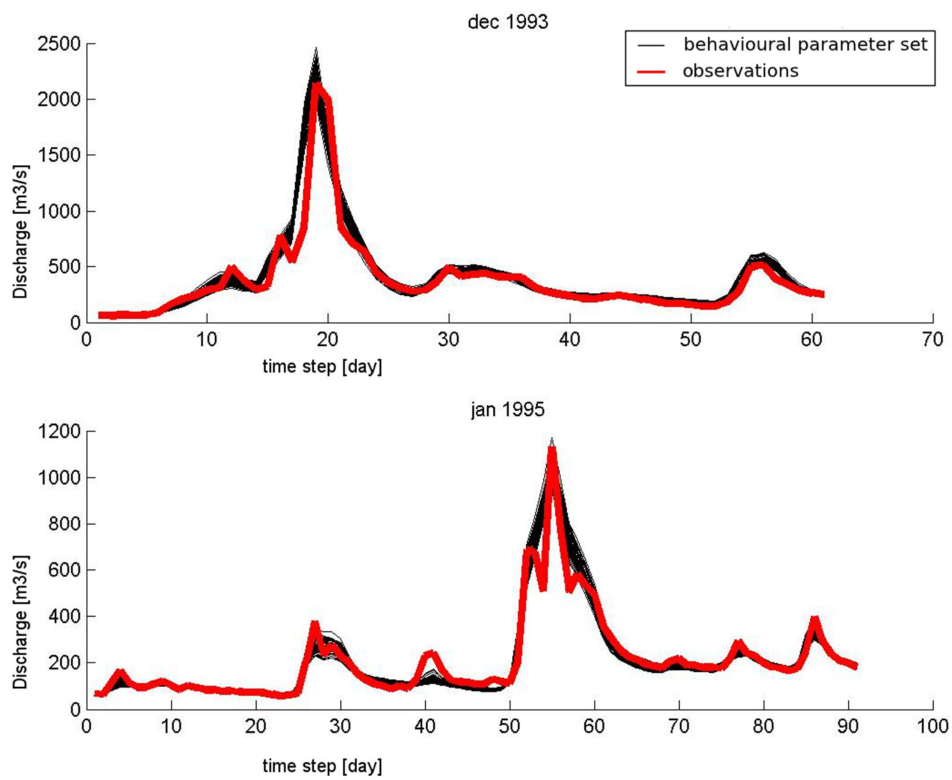


Figure 3.2 The modelled discharges from all behavioural parameter sets (black lines) and the observed discharges (red line) for the Neckar

3.2 Main

The Main has been analysed in three parts. The first part treats all the independent HBV units in the upstream part of the basin. The second part treats the HBV units, which receive inflows from the HBV units, analysed in the first part and the third part treats the most downstream HBV units (mostly the Main river itself), which depend on all the aforementioned units. The treatment of all HBV units is summarised in Table 3.2. HBV units that were calibrated together on one station are given in one box. Pegnitz and Rednitz are neighbouring catchments and have similar characteristics, but only data at the outlet of the Pegnitz (Nuernberg) were deemed reliable. Therefore, Rednitz has been given the same parameter sets as Pegnitz.

Table 3.3 Calibration setup for the Main

Calibration experiment	HBV Units included	BfG station calibration
Main1	Pegnitz Rednitz	Pruemzurlay
	Aisch	Laufermuehle
	Main1	SChwuerbitz
	FrSaale	Wolfsmünster
	Tauber	Tauberbischofsheim
	Kinzig	Hanau
	Nidda	Bad Vilbel
Main2	Main2	Kemmern
Main3	Regnitz Main3 Main4 Main5 Main6 Main7 Main8	Raunheim

Table 3.4 Selection criteria for the Main

	HBV Units Included	Thres_R2	Thres_REV	Thres_T5	Thresh_T20	Nr. of sets
Main1	Pegnitz, Rednitz	**0.8**	**0.5**	**0.5**	**0.5**	23
	Aisch	0.9	**0.1**	**5**	**5**	**10**
	Main1	0.9	0.05	0.1	0.1	110
	FrSaale	0.9	0.05	0.1	0.1	**14**
	Tauber	0.9	0.05	**0.3**	**0.3**	42
	Kinzig	0.9	0.05	0.1	0.1	36
	Nidda	0.9	0.05	**0.2**	**0.2**	**10**
Main2	Main2	0.9	0.05	0.1	0.1	962
Main3	Regnitz, Main3-8	0.9	0.05	0.1	0.1	**15**

In the Aisch basin, problems occur with the peak flows. The problems are probably caused by the fact that the geology of the Aisch basin mainly consist of Karst (discussed with BfG, Dennis Meissner, personal communication), which is not included in the HBV model structure. Additionally, the hydrological threshold processes related to karst occur at a much smaller time scale than the daily time step of the model presented here. Karst systems have the characteristics to react very fast and the timescale of these processes is in the order of hours. In Figure 3.3 it is visible that the model cannot reproduce the majority of the peaks.

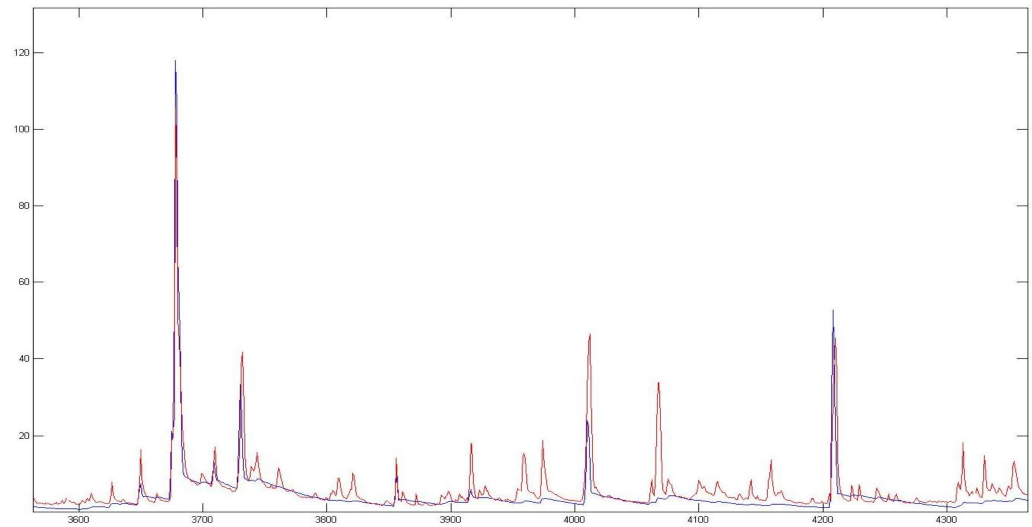


Figure 3.3 Graph showing the measured (red) and modelled (blue) flows from the Aisch basin

Calibration for the Rednitz and Pegnitz basins gives bad results, probably due to anthropogenic influences in these basins. These include interbasin transfers to neighbouring catchments (Meissner, personal communication).

Despite these issues in three of the sub-basins, the performance of the Main basin at its outflow point is good. Apparently the impact of the three relatively small basins on overall performance is small. Despite the fact that only a limited amount of data was available for calibration, the optimal Nash-Sutcliffe for the most downstream basins (Main3 experiment) is 0.92.

3.3 Nahe

The Nahe consists of 3 HBV units. Each unit has been calibrated on a station as given in Table 3.5. The HBV unit Nahe1 flows into Nahe2, and Nahe2 flows into Nahe3.

Table 3.5 Calibration setup for the Nahe

Calibration experiment	HBV Units included	BfG station calibration
Nahe1	Nahe1	MartinStein
Nahe2	Nahe2	Boos
Nahe3	Nahe3	Grolsheim

The Nahe experiments are analyzed using the criteria thresholds as given in Table 3.6. These are the default threshold values. The Nash-Sutcliffe value for the Nahe3 experiment is 0.93, which is good.

Table 3.6 Calibration setup for the Nahe

Basin	Thres_R2	Thres_REV	Thres_T5	Thresh_T20	Nr. of sets
Nahe1	0.9	0.05	0.1	0.1	109
Nahe2	0.9	0.05	0.1	0.1	404
Nahe3	0.9	0.05	0.1	0.1	37

In Figure 3.4 the results are shown for the 1993 and 1995 events at the Nahe outlet at Grolsheim. The modelled hydrographs are in agreement with the observed discharges. Furthermore, the modelled peak discharges are relatively close to the observed discharges. There is not a real under- or overestimation of the peak discharges.

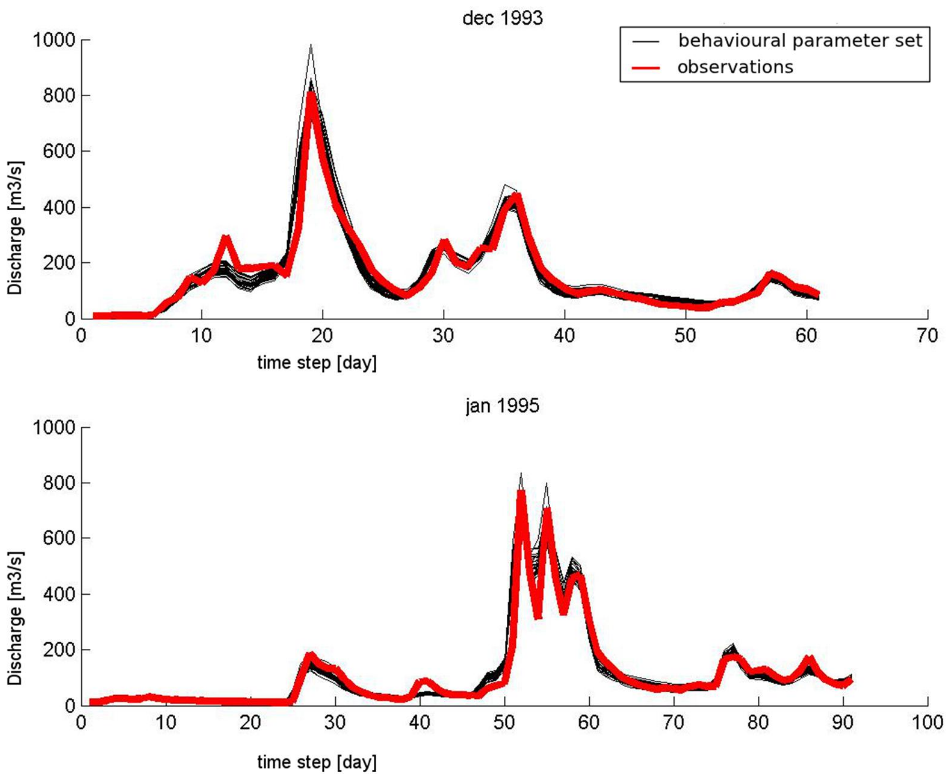


Figure 3.4 Results for the Nahe3 experiment (Grolsheim). Plotted are the behavioural parameter sets (black), versus the measured flows

3.4 Lahn

The Lahn has been analysed in three parts. The first part treats the two independent HBV units in the upstream part of the basin. The second part treats the HBV units that receive inflows from the HBV units analysed in the first part and the third part treats the most downstream HBV units, which depend on all the aforementioned units. The treatment of all HBV units is summarised in Table 3.7. Lahn5 represents the drainage area in between Kalkofen and the confluence with the Rhine. For the moment, the discharge from this unit is assumed to be zero. This has been implemented by setting ‘pcorr’ (precipitation correction factor) to zero.

Table 3.7 Calibration setup for the Lahn

Calibration experiment	HBV Units included	BfG station calibration
Lahn1	Lahn1	Marburg
	Dill	Asslar
Lahn2	Lah2	Leun
Lahn3	Lahn4	Kalkofen
	Lahn5	**Not calibrated**

The Lahn experiments are analyzed using the criteria from Table 3.8.

Table 3.8 Selection criteria for the Lahn

Catchment	HBV Units included	R2	REV	T5	T20	Nr. of sets
Lahn1	Dill	0.9	0.05	0.1	0.1	71
	Lahn1	0.9	0.05	0.1	0.1	20
Lahn2	Lahn2	0.9	0.05	**0.3**	**0.3**	**19**
Lahn3	Lahn4	0.9	0.05	0.1	0.1	516

In the original setup, the model predictions for the Lahn3 experiment were always in advance of the observed flows. This has been adjusted by setting the MAXBAS parameter for the Lahn2 basin to 2.0 by trial and error.

Good results are obtained during the Lahn1 experiment. The results for the Lahn2 experiment are reasonably good but do not bracket the observations enough. To obtain a better bracketing of observations the T5 and T20 criteria are slightly widened. Apparently there is more uncertainty in the modelled flows. This uncertainty has therefore been accounted for by this widening of the criteria.

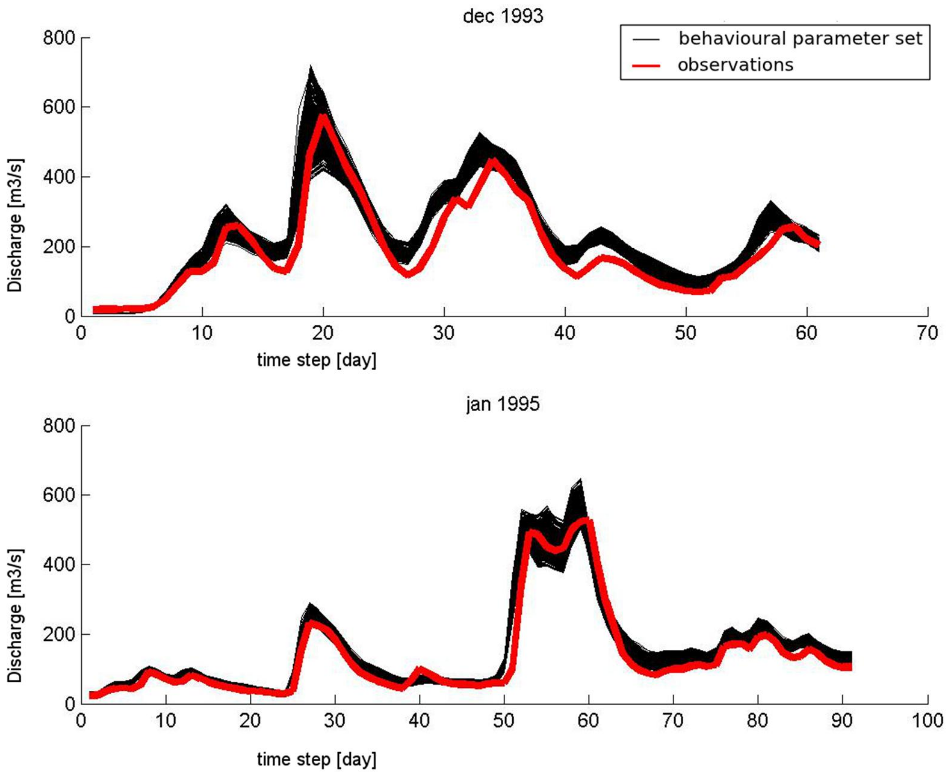


Figure 3.5 Results for the Lahn3 experiment (Kalkofen). Plotted are the behavioural parameter sets (black), versus the measured flows

The Lahn3 experiment generates good results, although the model slightly overestimates the discharge during the 1993 event. This is shown in Figure 3.5.

3.5 Moselle

The Moselle has been analysed in three parts. The first part treats all the independent HBV units in the upstream part of the basin. The second part treats the HBV units, which receive inflows from the HBV units, analysed in the first part and the third part treats the most downstream HBV units, which depend on all the aforementioned units. The treatment of all HBV units is summarised in Table 3.9 HBV units that were calibrated together on one station are given in one box. Umos4 represents the drainage area in between Cochem and the confluence with the Rhine. For the moment, the discharge from this unit is assumed to be zero. This has been implemented by putting ‘pcorr’ (precipitation correction factor) on zero.

Table 3.9 Calibration setup for the Moselle

Calibration experiment	HBV Units included	BfG station calibration
Moselle1	Alzette Sure Sauer	Pruemzurlay
	Our	Gemuend
	Pruem	Pruemzurlay
	Nims	Alsdorf
	Blies_1	Reinheim
	Nied_1	Niedaltdorf
	Prims_1	Nalbach
	Obsa	Wittringen
	Omos2	Malzevillier
	Omos1	Toul
	Seille	Metz
	Kyll	Kordel
	Lieser	Plein
	Orne	Rosselange
Moselle2	Unsaar Rest1 Sauer2 Omos3 Omos4 Umos1	Trier
Moselle3	Ruwer Umos2 Umos3	Cochem
	Umos4	**Not calibrated**

In general, results for the Moselle river are very good. The Nash-Sutcliffe for all sub-basins varies between 0.84 and 0.94. Only in a few sub-basins, the selection criteria needed to be widened in order to find enough behavioural parameter sets.

In Table 3.10 the used criteria threshold values for the different sub-basins are summarized. In figure Figure 3.6 the resulting hydrograph for the Moselle2 experiment is shown.

Table 3.10 Selection criteria for the Moselle

	HBV Units included	R2	REV	T5	T20	Nr. of sets
Moselle1	Alzette,Sure, Sauer1/2	0.9	0.05	0.1	0.1	26
	Our	0.9	0.05	0.1	0.1	**17**
	Pruem	0.9	0.05	**0.3**	**0.3**	200
	Nims	0.9	0.05	0.1	0.1	23
	Blies_1	0.9	0.05	0.1	0.1	106
	Nied_1	0.9	0.05	**0.2**	**0.2**	52
	Prims_1	0.9	0.05	0.1	0.1	23
	Obsa	0.9	0.05	0.1	0.1	26
	Omos2	0.9	**0.1**	**0.2**	**0.2**	**14**
	Omos1	0.9	0.05	0.1	0.1	35
	Seille, Orne, Omos3	0.9	0.05	0.1	0.1	31
	Kyll	0.9	0.05	0.1	0.1	194
	Lieser	0.9	0.05	**0.2**	**0.2**	62
Moselle2	Unsaar,Omos4,Rest1, Umos1	0.9	0.05	0.1	0.1	3604

The parameters from the Moselle2 experiment are used in the Moselle3 basins and therefore the Moselle3 experiment is not calibrated. The reason for this is that the observed discharge at Cochem is believed to be underestimated (discussed with BfG, based on conclusions from HYMOG). This is likely to be caused by a rating curve problem.

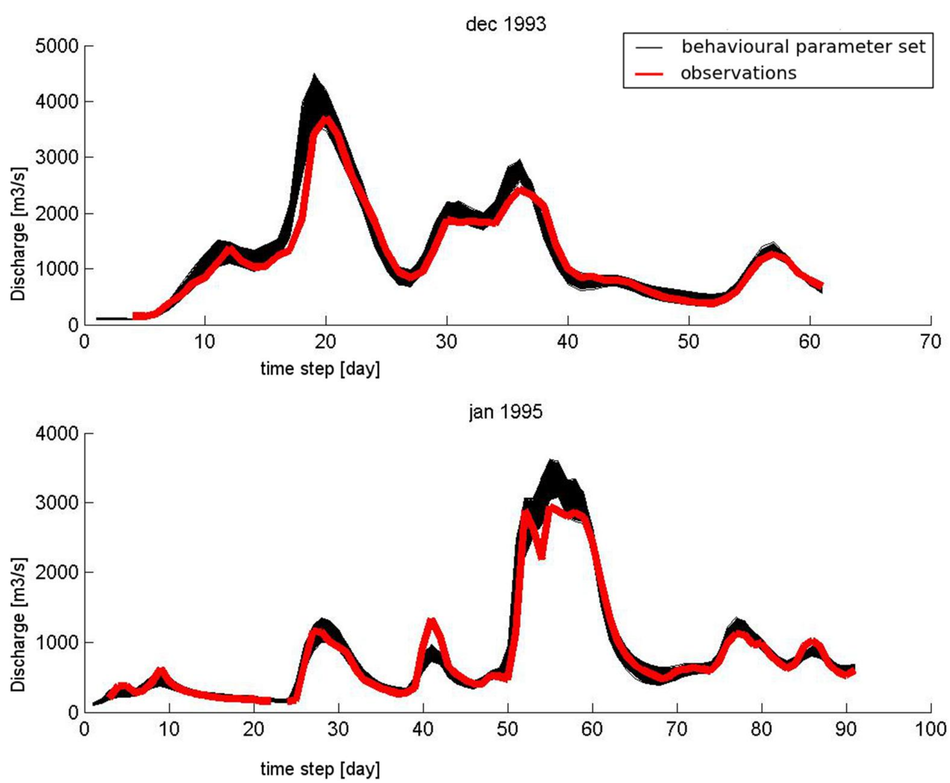


Figure 3.6 Results for the Moselle2 experiment (Trier). Plotted are the behavioural parameter sets (black), versus the measured flows

3.6 Sieg

The Sieg has been analysed in three parts. The first part treats the two independent HBV units in the upstream part of the basin. The second part treats the HBV units, which receive inflows from the HBV units, analysed in the first part and the third part treats the most downstream HBV units, which depend on all the aforementioned units. The treatment of all HBV units is summarised in Table 3.11.

Table 3.11 Calibration setup for the Sieg

Calibration experiment	HBV Units included	BfG station calibration
Sieg1	Obsi	Betzdorf
	Agger	Lohmar
Sieg2	Misi	Eitorf
Sieg3	Unsi	Menden

Table 3.12 Selection criteria for the Sieg

Experiment	HBV Units included	R2	REV	T5	T20	Nr. of sets
Sieg1	Obsi	0.9	0.05	**5**	**5**	63
	Agger	0.9	0.05	0.1	0.1	959
Sieg2	Sieg2	0.9	0.05	**5**	**5**	2265
Sieg3	Sieg3	0.9	0.05	**0.2**	**0.2**	522

The results for the Sieg are good (see Figure 3.7), although for the Obsi basin and for the Lahn2 experiment the T5 and T20 criteria were left out. The reason was mainly that the observed/simulated Gumbel distribution and GEV distribution were poorly overlapping (see Figure 3.8), so no behavioural sets common to both extreme value distribution functions could be found. The reason for this has to be investigated.

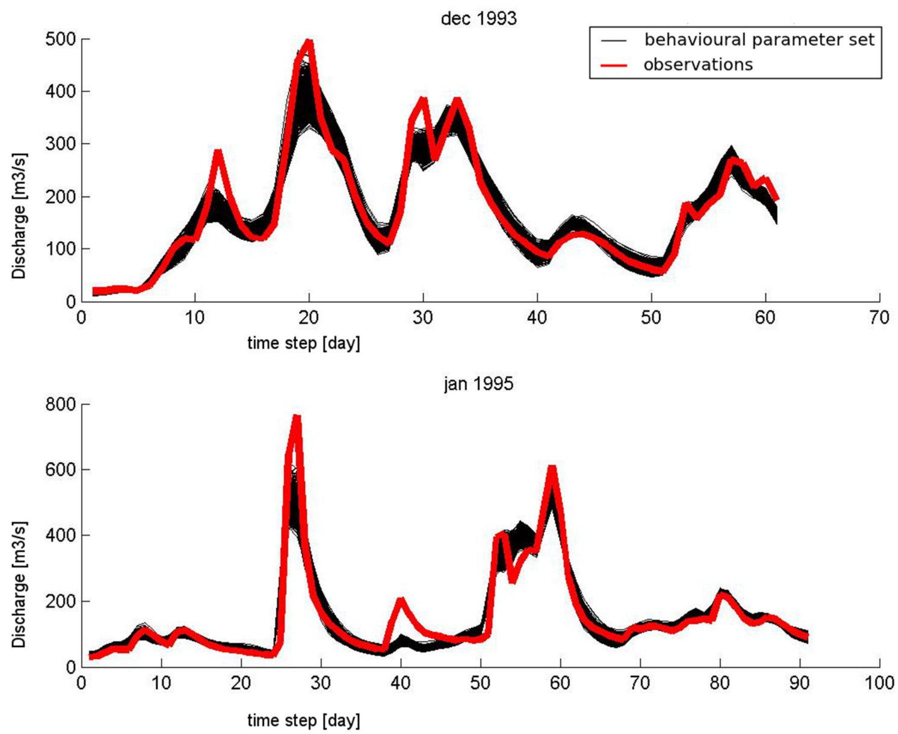


Figure 3.7 Plot showing the calibration results for the Sieg3 experiment (Menden)

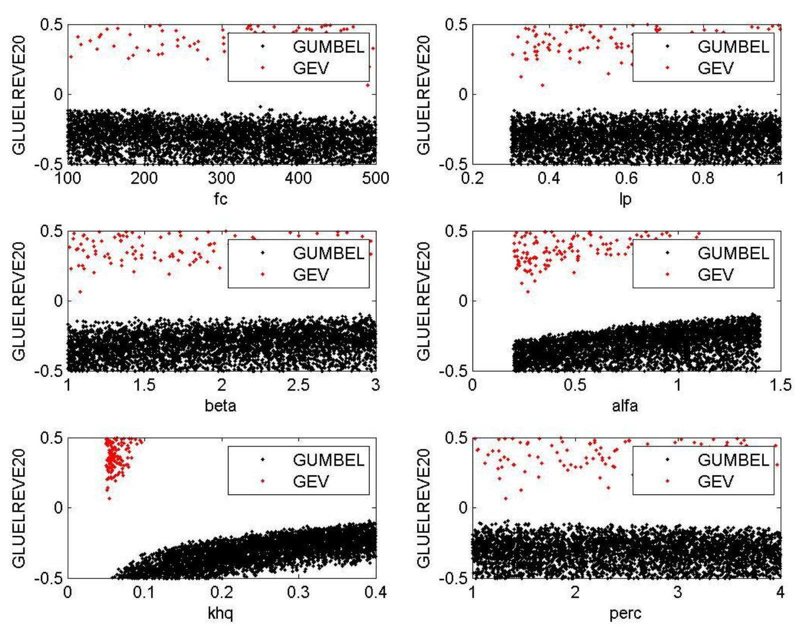


Figure 3.8 Plots showing the relative extreme value error based on Gumbel (black) and GEV (red) values for each model run. The y-axis GLUELREVE20 is the relative volume error of the extreme discharge value with a return period of 20 years

3.7 Erft

The Erft consists of 3 HBV units. The most downstream unit was assumed to provide zero discharge. This has been implemented by setting 'pcorr' (precipitation correction factor) to zero. The 2 upper units were calibrated as listed in Table 3.13.

Table 3.13 Calibration setup for the Erft

Calibration experiment	HBV Units included	BfG station calibration
Erft1	Erft1	Bliesheim
Erft2	Erft2	Neubrück
	Erft3	**Not calibrated**

Table 3.14 Selection criteria for the Erft

Experiment	HBV Units included	R2	REV	T5	T20	Nr. of sets
Erft1	Erft1	0.9	0.05	**5**	**5**	37
Erft2	Erft2	-	-	-	-	-

The results for the Erft1 experiment are reasonable good. The optimum Nash-Sutcliff value is 0.76. The reason for the low value is probably that this basin is too small to be modelled on daily basis. The important processes in this basin then have a characteristic time scale of less than a day.

The results for the Erft2 experiment are not good. In fact, no behavioural sets were found for Erft2. The reason is that there are anthropogenic processes such as lignite mining that influence the hydrological processes significantly. These are not included in the model. In Erft2, water is pumped from the (deep) groundwater to ensure open cast lignite mining in the area. This mining is done via open cast methods, until a depth of 300-500 meter. An example of such a sight is shown in Figure 3.9.

It was decided to use the same parameter sets in Erft2 as in Erft1, so as to ensure that natural conditions are simulated as much as possible. This is justified by the fact that a) the required return period of 1/1250 years is beyond a time scale at which the mining activity takes place; b) peak discharges are likely to be less impacted by mining activities; and c) the Erft2 contribution to total flow at Lobith is small.

Figure 3.9 Photograph of the open cast Lignite mining near Hambach (source: wikipedia)

3.8 Ruhr

The Ruhr consists of 4 HBV units. The 3 most upstream units are all calibrated on a station as given in Table 3.15. Ruhr4, drains the area in between Hattingen (outlet of Ruhr3) and the Rhine confluence. For the moment, the discharge from this unit is assumed to be zero. This has been implemented by putting 'pcorr' (precipitation correction factor) on zero.

Table 3.15 Calibration setup for the Ruhr

Calibration experiment	HBV Units included	BfG station calibration
Ruhr1	Ruhr1	Villigst
	Ruhr2	Hagen-Hohenlimburg
Ruhr2	Ruhr3	Hattingen
	Ruhr4	**Not calibrated**

The Ruhr 1 is analyzed with a band on the criteria following Table 3.16.

Table 3.16 Selection criteria for the Ruhr

Experiment	HBV Units included	R2	REV	T5	T20	Nr. of sets
Ruhr1	Ruhr1	0.9	0.05	0.1	0.1	181
	Ruhr2	0.9	**0.15**	0.1	0.1	64
Ruhr2	Ruhr3	0.9	0.05	0.1	0.1	64

The resulting hydrograph for the Ruhr2 experiment is shown in Figure 3.10. The results are good (N-S of 0.93), although it seems that in the December 1993 case a considerate part of the rainfall was not accounted for. This causes a too low peak in this particular period within the complete time series. The dynamics, as well as the remainder of the time series are in good correlation with the observations.

It is important to note that the Ruhr is a highly regulated river. This is not accounted for in HBV and therefore the expectation would be that the results of the conditioning process would lead to high uncertainty in the parameter sets. The results however, reveal that the parameter uncertainty is not very high. Apparently the regulation of the Ruhr does not have significant impact on the flow regime.

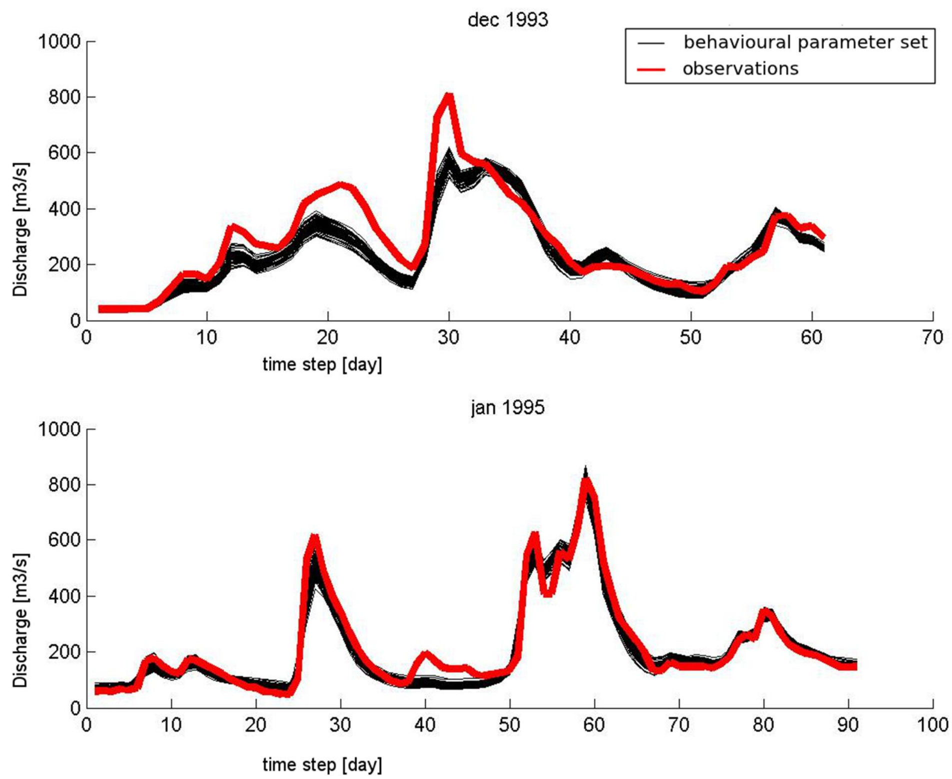


Figure 3.10 Results for the Ruhr2 experiment (Hattingen). Plotted are the behavioural parameter sets (black), versus the measured flows

3.9 Lippe

The Lippe consists of 3 HBV units. All three have a discharge station to use for calibration (see Table 3.17).

Table 3.17 Calibration setup for the Lippe

Calibration experiment	HBV Units included	BfG station calibration
Lippe1	Lippe 1	Lippstadt
Lippe 2	Lippe 2	Haltern
Lippe 3	Lippe 3	Schermbeck

The Lippe experiments are analyzed using the bounds on the criteria as described in Table 3.18. The results for the Lippe1 and Lippe2 experiments provide just enough behavioural parameter sets to pass on to the Lippe3 experiment.

The Lippe3 experiment gives good results (Figure 3.11), but the model overestimates the peak discharge during the 1995 event. This overestimation is not the case for every peak flow.

Table 3.18 Selection criteria for the Lippe

Experiment	HBV Units included	R2	REV	T5	T20	Nr. of sets
Lippe1	Lippe1	0.9	**0.15**	0.1	0.1	**15**
Lippe2	Lippe2	0.9	0.05	0.1	0.1	**17**
Lippe3	Lippe3	0.9	0.05	0.1	0.1	298

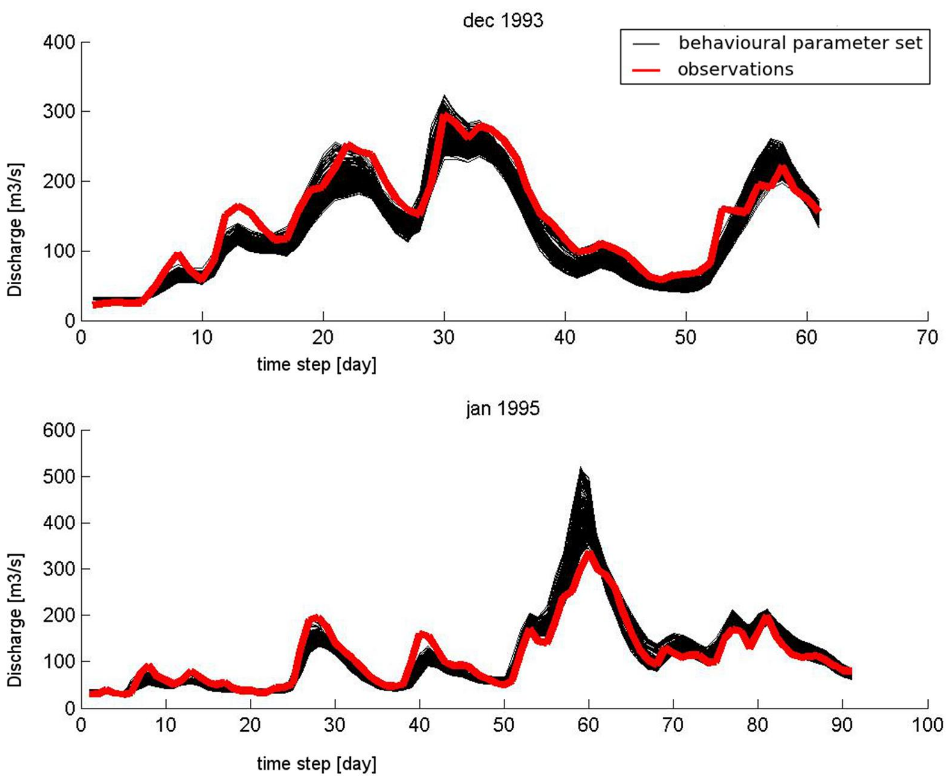


Figure 3.11 Results for the Lippe3 experiment (Schermbeck). Plotted are the behavioural parameter sets (black), versus the measured flows

3.10 Southern Upper Rhine

In the Southern Upper Rhine, a number of HBV units consisted of two or three smaller streams with discharge stations at their outlets. These station data are summed to yield effective discharge from the specified unit. In the table below, this is indicated by '+' signs.

Table 3.19 Calibration setup for the Upper Rhine

Calibration experiment	HBV Units included	BfG station calibration
SouthUpRhine1	Ill1 Ill2 Ill3	Colmar-Ladhoff
	Fecht Bruche	Ostheim
	Elzdreis1 Elzdreis2	Ebnet + Gutach
	KinzigUp UpRh2_2	Schwaibach + Lahr
	SauWies Moder Zorn	Salmbach + Niederrödern
	Murgrenz	Ramsbach + Kappelrodeck + Rotenfels
	UpRh2_1 **Kanal** **UpRh2_3**	**Not calibrated** **Not calibrated** **Not calibrated**

In Table 3.20 the criteria ranges are listed for the different basins. The results for the South Upper Rhine are not very good, resulting is more uncertainty in the simulated discharges. This is reflected in wider ranges for the selection criteria.

The reason for the rather bad behaviour in the South Upper Rhine is found in the fact that the South Upper Rhine basins' hydrological processes have a typical time window of less than a day. This is probably caused by the topography (steep slopes) of the area.
Figure 3.12 and Figure 3.13 show the results for two basins for 1993 and 1995 event.

Table 3.20 Selection criteria for the Upper Rhine

Experiment	HBV Units included	R2	REV	T5	T20	Nr. of sets
SouthUpRhine	Ill1, Ill2, Ill3	0.9	**0.20**	0.1	0.1	26
	Fecht, Bruche	0.9	0.05	0.1	0.1	169
	Elzdreis1, Elzdreis2	**0.8**	0.05	**0.4**	**0.4**	59
	KinzigUp, UpRh2_2	**0.8**	0.05	**0.2**	**0.2**	56
	Sauwies, Modder, Zorn	**0.8**	**0.2**	**1.0**	**1.0**	**4**
	Murgrenz	**0.8**	0.05	**0.2**	**0.2**	**13**

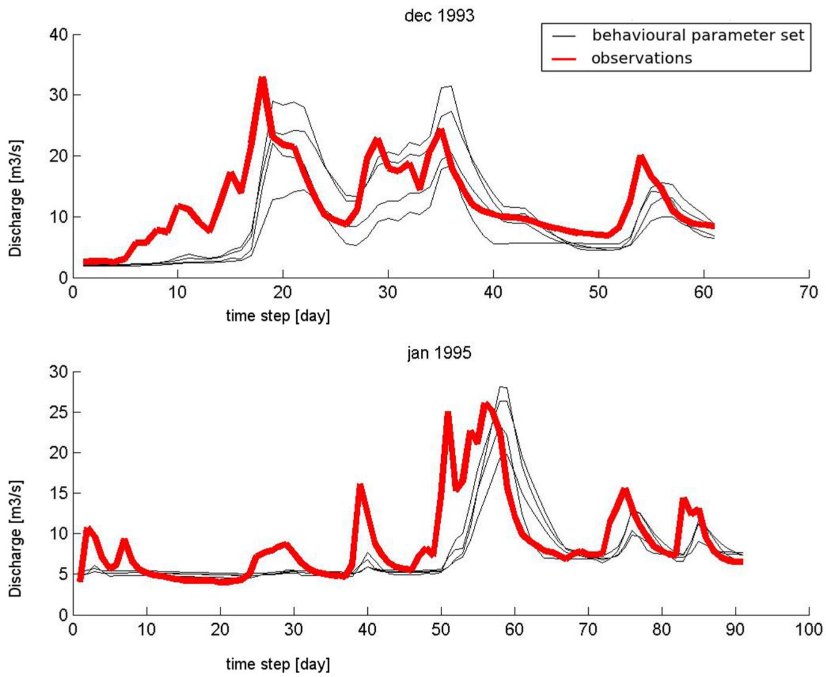


Figure 3.12 Results for the Sauwies, Moder and Zorn basins

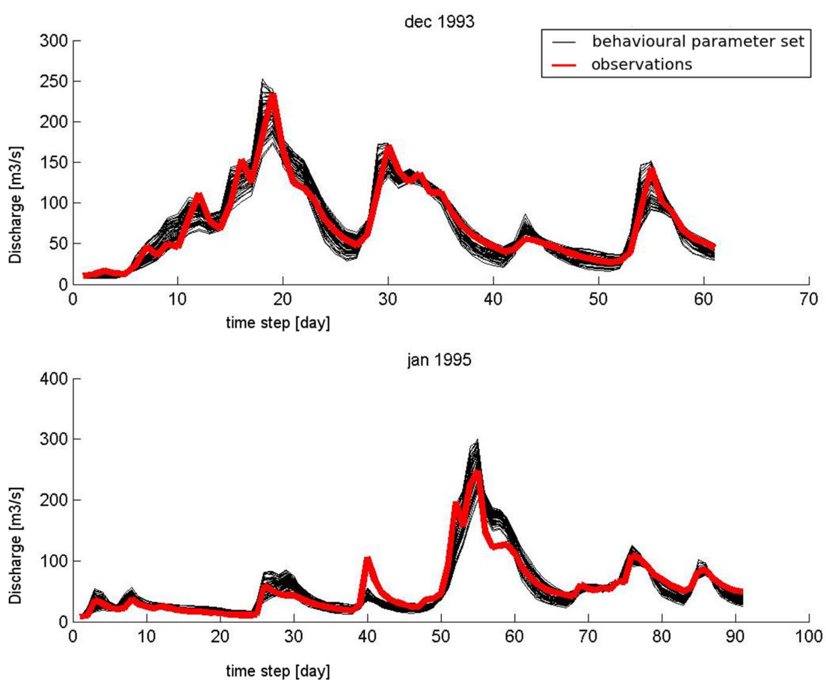


Figure 3.13 Results for the Elzdreis1 and Elzdreis2 basins

3.11 Upper Rhine

For the Upper Rhine the independent upstream HBV units have been analysed (Table 3.21). The UpRhine HBV units drain the areas between the independent HBV units and the Rhine. For the moment, the discharges from these units are assumed to be zero. This has been implemented by setting 'pcorr' (precipitation correction factor) to zero and by excluding inflows from the upstream part the Rhine (UpRh2_3), Neckar, Main and Worms.

Table 3.21 Calibration setup for the Upper Rhine

Calibration experiment	HBV Units included	BfG station calibration
UpperRhine	QueichSpeyerbach	Neustadt/Wst + Siebeldingen
	AlbPfinz	Berghausen + Ettlingen
	Nette	Nettegut
	Wied	Frierichstal
	WeschnitzModau	Lorsch + Eberstadt
	UpRhine1	**Not calibrated**
	UpRhine2	**Not calibrated**
	UpRhine3	**Not calibrated**
	UpRhine4	**Not calibrated**

The UpperRhine is analyzed using the criteria thresholds from Table 3.22.

Table 3.22 Calibration setup for the Upper Rhine

Experiment	HBV Units included	R2	REV	T5	T20	Nr. of sets
UpperRhine	Albpfinz	0.9	0.05	0.1	0.1	67
	QueichSpeyerbach	0.9	0.05	0.1	0.1	**16**
	WeschnitzModau	0.9	**0.1**	**0.2**	**0.2**	**9**

The results for the UpperRhine are reasonable. In Figure 3.14 and Figure 3.15 the hydrographs for two basins are shown. The flows in both cases are mainly overestimated. A source for this error is found in the way these basins are calibrated. The sub-basins are calibrated on the sum of two discharge stations. Although the total volume could be accounted for correctly, the peaks in the measurements can be damped or can be increased because of a shift in time.

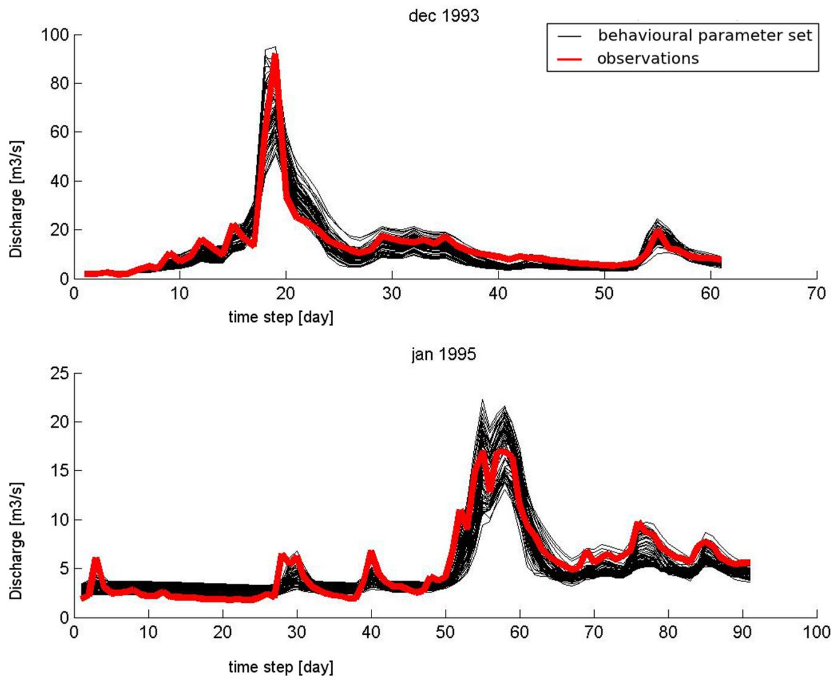


Figure 3.14 Results for the AlbPfinz experiment (Berghausen + Ettlingen)

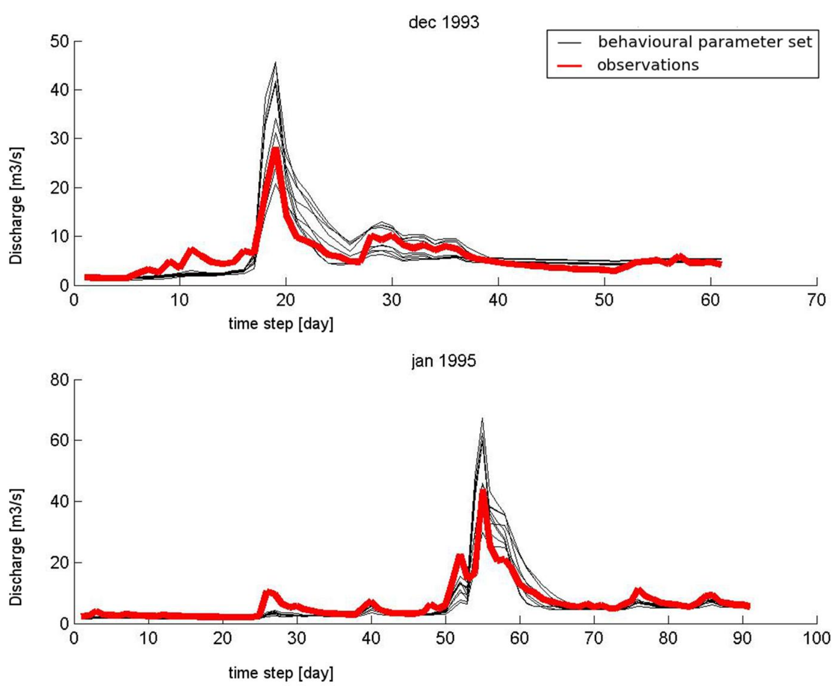


Figure 3.15 Results for the WeschnitzModau experiment (Lorsch + Eberstadt

3.12 Middle Rhine

For the Middle Rhine the independent upstream HBV units have been analysed (Table 3.23). The HBV units that were not calibrated drain the areas between the independent HBV units and the Rhine. For the moment, the discharges from these units are assumed to be zero. This has been implemented by setting 'pcorr' (precipitation correction factor) to zero and by excluding inflows from the upstream part the Rhine (UpRhine4), Nahe, Lahn, Moselle and Sieg.

Table 3.23 Calibration setup for the Middle Rhine

Catchment	HBV Units included	BfG station calibration
MiddleRhine	Selz	Oberingelheim
	Wisper	Pfaffental
	Nette	Nettegut
	Wied	Frierichstal
	Ahr	Altenahr
	MidRhine1	**Not calibrated**
	MidRhine2	**Not calibrated**
	Saynbach	**Not calibrated**
	MidRhine3	**Not calibrated**

Table 3.24 Selection criteria for the Middle Rhine

Catchment	HBV Units included	R2	REV	T5	T20	Nr. of sets
MiddleRhine	Selz	0.9	0.1	**5**	**5**	**2**
	Wisper	0.9	0.05	0.1	0.1	21
	Nette	0.9	0.05	0.1	0.1	51
	Wied	0.9	0.05	0.1	0.1	70
	Ahr	0.9	0.05	0.1	0.1	30

The results for the MiddleRhine experiment are good, with the exception of the Selz basin. This is also mentioned in SMHI report (2009). The reason could be an error in the discharge data.

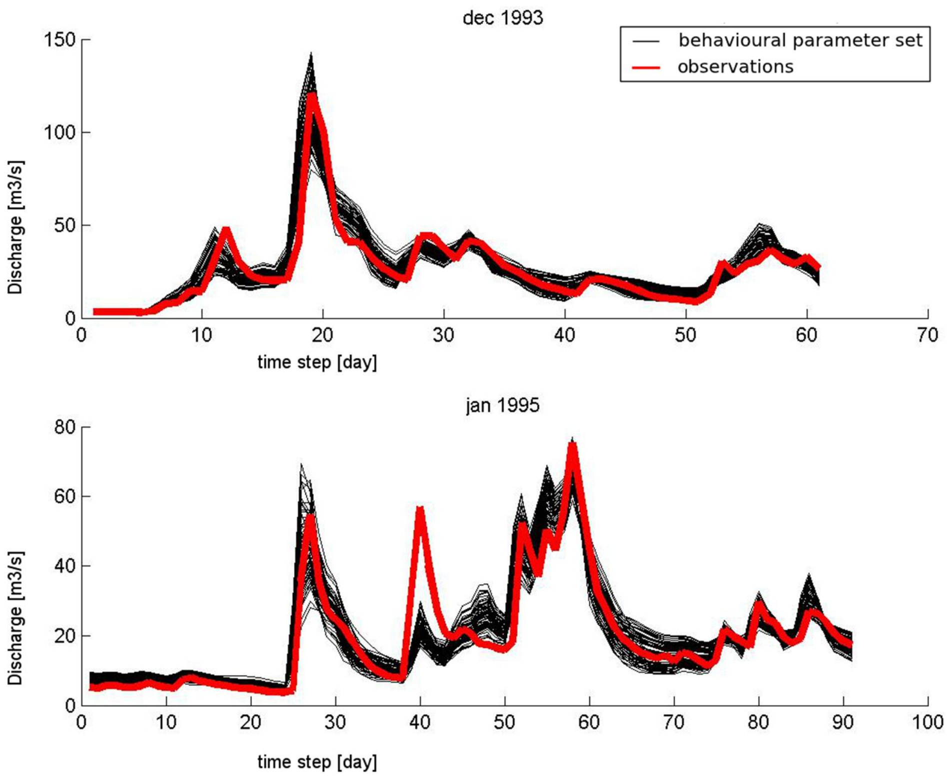


Figure 3.16 Results for Wied basin(Frierichstal)

3.13 Lower Rhine

Only two HBV units in the lower Rhine were included for calibration. Wupper1 and Emscher. Wupper 2 was assumed to have the same parameter set as Wupper1. The remainder of the units are Zwischeneinzugsgebiete and were not included in calibration. For the moment, their contribution has been set on zero by setting the parameter 'pcorr' to zero.

Table 3.25 Calibration setup for the Lower Rhine

Calibration experiment	HBV Units included	BfG station calibration
LowerRhine1	Wupper1 Wupper2	Opladen
	Emscher	Konigstrasse
	Other HBV units	**not calibrated**

Table 3.26 Selection criteria for the Lower Rhine

Experiment	HBV Units included	R2	REV	T5	T20	Nr. of sets
LowerRhine	Wupper1, Wupper2	0.9	**0.1**	**5**	**5**	**13**
	Emscher	0.9	0.05	0.1	0.1	**12**

The results for the LowerRhine are reasonably good. The main problem occurs in the Emscher basin. In the Emscher basin there are reservoirs and a high degree of urbanization, impacting on the hydrological processes. Furthermore, the Emscher receives water from the Ruhr in an interbasin transfer. All these processes are not accounted for in HBV. The Nash-Sutcliffe for the Emscher is therefore only 0.7. Although the Wupper also contains reservoirs, the Nash-Sutcliffe values for the Wupper basins are 0.88.

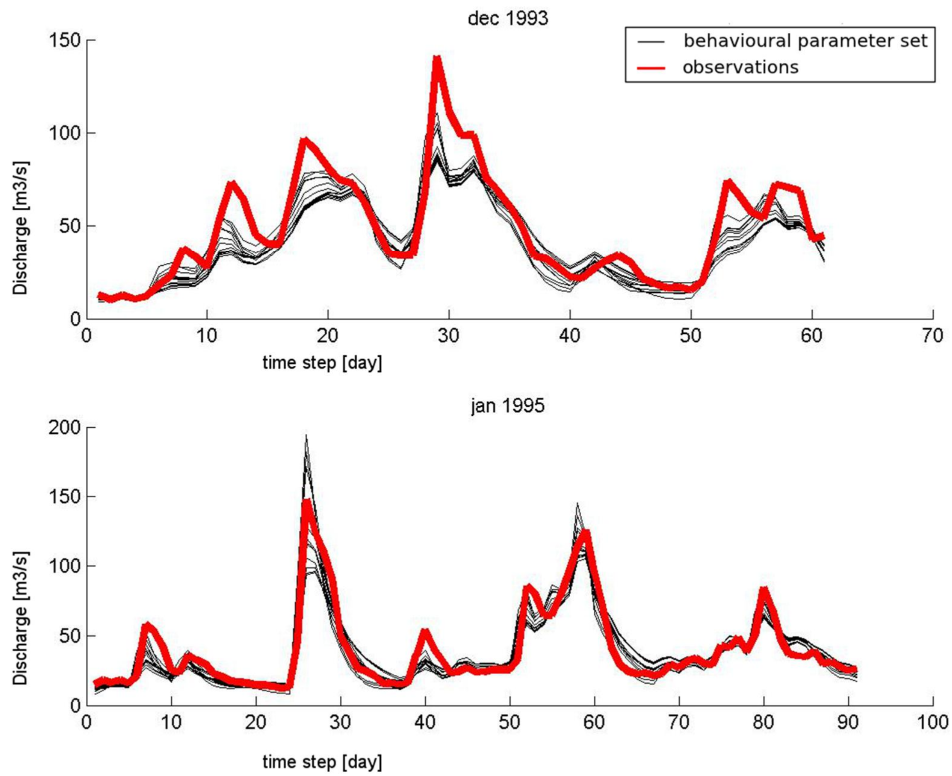


Figure 3.17 Results for Wupper1 and Wupper2 basin (Opladen)

3.14 Overall performance in the Rhine, are precipitation corrections still required?

An overview has been made of the general performance over the complete Rhine basin. Of each HBV unit considered during the GLUE analysis, we have computed the best performing Nash-Sutcliffe Efficiency value, according to the flow station, used for constraining in GLUE. In Figure 3.18 the resulting optimum Nash-Sutcliff values for each sub-basin are given. Although these N-S values do not correspond to the selected parameter sets, this value gives an impression of the performance during the GLUE analysis. The basins that were calibrated together (on the same discharge station) were given the same N-S value.

Figure 3.19 gives the precipitation correction factors, as they were present in the original HBV model, as calibrated by SMHI. The figures show clearly that where high precipitation correction factors were established in previous calibration studies, we now have good performances, without any precipitation corrections. Therefore, we have decided that precipitation factors can be removed from the configuration.

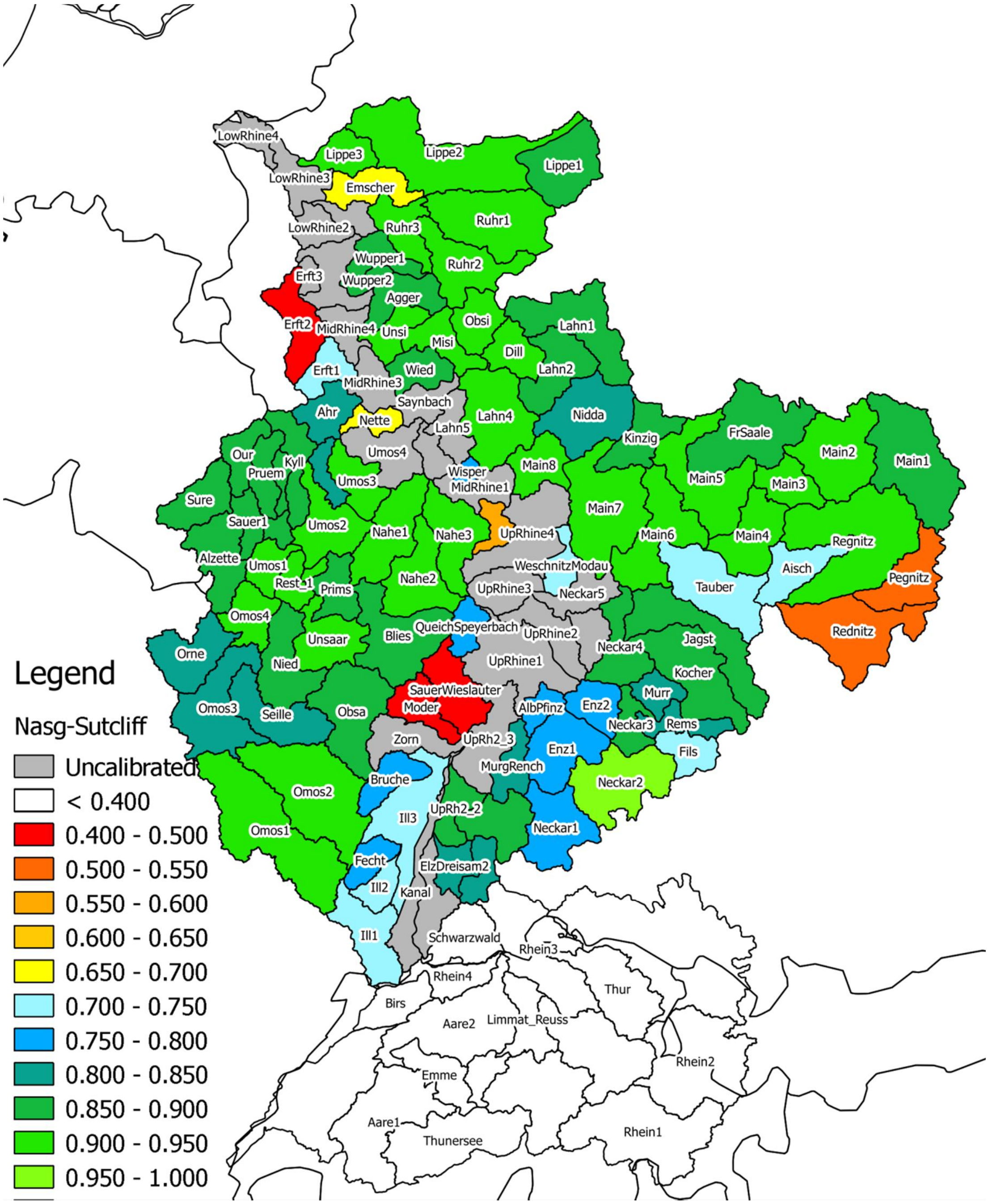


Figure 3.18 Overview of highest obtained Nash-Sutcliffe coefficient during sampling per sub-basin. Basins which were calibrated together receive the same Nash-Sutcliff value

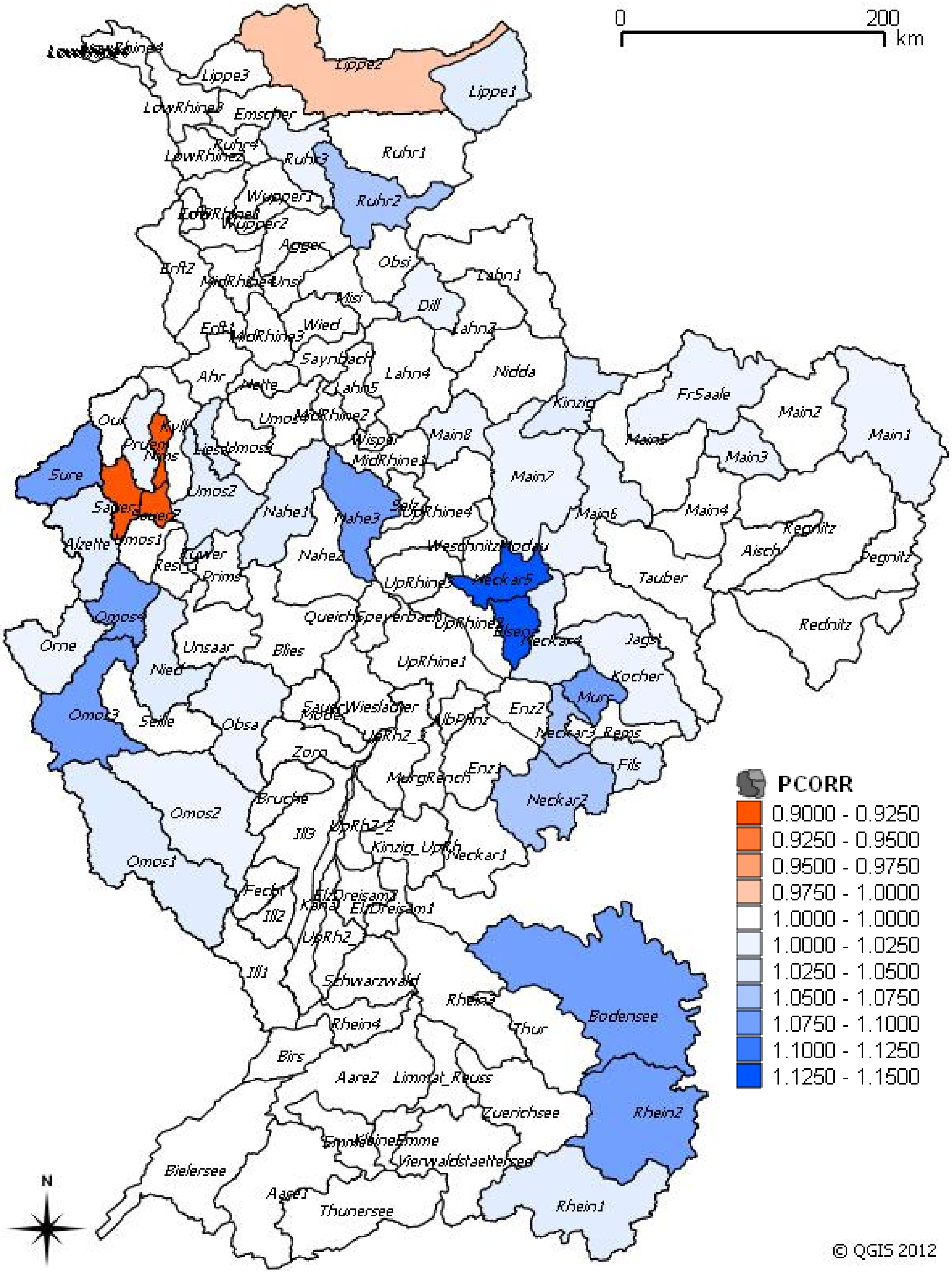


Figure 3.19 Original precipitation correction factors per sub-basin. These precipitation correction factors are now removed from the configuration

4 Parameter selection

The GLUE analysis results in a number of approved parameter sets per subbasin. The number of behavioural parameter sets varies per basin and ranges from a minimum of 10 to several hundreds. When applied in GRADE, it is infeasible to calculate all combinations of all parameter sets from all 13 basins, considered in the GLUE analysis. This would take an enormous computational effort. Therefore, a selection of parameter sets was made that reflects a representative sample for extreme value analysis. In total 5 parameter sets per basin were selected. Each set is able to simulate a certain quantile of extreme discharges. This means that for each parameter set the 1/10 years discharge is determined, these values (one discharge value for each parameter set) are sorted and from that the 5%, 25%, 50%, 75% and 95% quantiles have been selected as representative.

The selection of the 5 parameter sets is done according to the following steps:

1) Run the HBV model for each selected set over the available observation period, in this case from 01-01-1985 until 31-12-2006. From each HBV run, annual maxima are retrieved and sorted into a cumulative distribution function (cdf). This results in a cdf for each parameter set.
2) From each cdf, the value with a relatively high return period is retrieved. From these values, the 5, 25, 50, 75 and 95 percentile values are derived and the associated parameter set saved. A high return period is selected, because GRADE is used for extreme high discharges.

As there may be some sensitivity towards the selected return period, used to retrieve a parameter set, this procedure has been followed for the 2, 5 and 10-year return period. Although the differences were small, the 10-year return period was selected to work with. The 5 parameter sets that were derived are used in an uncertainty analysis. The 5 parameter sets together span the uncertainty band of the HBV model. The 50% parameter set is also used as reference model which is used to do the GRADE calculations.

5 Water balance in the main stem of the Rhine

5.1 Introduction and approach

A water balance analysis has been performed with the newly established parameter sets on the main stem of the Rhine. It should be noted that the intermediate basins (Zwischeneinzugsgebieten, ZWE) in between the sub-basins analyzed and the Rhine itself, have not been considered so far in this analysis. In fact the HBV parameter file has been set such that outflow from these basins is zero at all times. Therefore, if the HBV model is run with the current derived parameter sets, any water that comes from these intermediate basins is assumed to be negligible.

In the water balance analysis, HBV is run over a long period (1985-2006) with the 5 new parameter sets, selected on their extreme value with return period of 10 years (described in Chapter 4) and the volumes of river flow passing along several points in the main stem are compared against the observed flow volumes. This has been done with a double mass curve on the full discharge time series, as well as the > 95% percentile flows, to emphasize the peak flows only. At the upper boundary at Maxau a considerable difference between simulated and observed flow can be observed, which is mainly caused by the fact that the daily HBV model for the Swiss part of the Rhine basin has not been calibrated or constrained properly yet, during the writing of this report. The parameter values of the daily model are based on parameter values of the hourly model instead and have not been altered so far. This large difference propagates downstream through the Rhine's main stem and therefore obscures the water balance differences, due to incompleteness or inaccuracies in the GLUE analyzed HBV models. Therefore, a GLUE analysis will also be performed for the Swiss part of the HBV model. For the time being, the observation-simulation difference at Maxau is imposed on all simulated time series in the downstream locations to remove this bias downstream and reveal the remaining errors in the German part of the Rhine. Finally, a time series of flows at Lobith during the January 1995 event is displayed to show in which domain of the flows most of the differences can be attributed to.

HBV only gives outputs to the main gauging stations on the main stem at Maxau, Worms, Kaub, Andernach, Köln and Lobith (as shown in Figure 5.2). Therefore, this comparison has been performed on these flow stations.

5.2 Results

The resulting Double Mass (DM) curves are given in Figure 5.3 until Figure 5.12. The time series for the 1995 high water at Lobith are shown in Figure 5.13. Finally, the relative differences between simulations and observations over the full time period are summarized in Table 5.1. All results are corrected for the measured discharge at Maxau, hence the titles of the plots ("Maxau corr."). The graphs and table show three main findings:

- Looking at the double mass plots, one could conclude that the average flow is slightly underestimated. This underestimation becomes more as we look further downstream. This can be explained by the fact that the ZWEs have not been accounted for in HBV. Due to the relatively moderate effect of these areas, these could in the short term be included by a simple correction factor, or by a HBV model with a thick unsaturated zone and only a slow linear reservoir outflow (representing groundwater), and no quick flow component. However, to fully understand the behavior of the ZWEs, in particular during extreme events, an in-depth study to their hydrological behavior is required.

- Looking at the double mass plots for the high flows (flows above value), the high flows (> 95% percentile) are slightly overestimated by the model by all parameter sets. In this water balance the routed flows from HBV have been used. HBV uses a simple flow-storage relation (similar to Muskingum) for flow routing and does not account for flow attenuation due to floods or retention areas. Further attenuation in these circumstances may be modeled with the hydrodynamic model SOBEK. In a separate study, the effect of including SOBEK in the model cascade will be investigated.
- Although quite some uncertainty is shown in the peak flows of the main tributaries, this uncertainty largely averages out as we move towards Lobith. This may partly be explained by the fact that no uncertainty is as yet determined upstream of Maxau. This area provides a considerable contribution to the flow at Lobith.

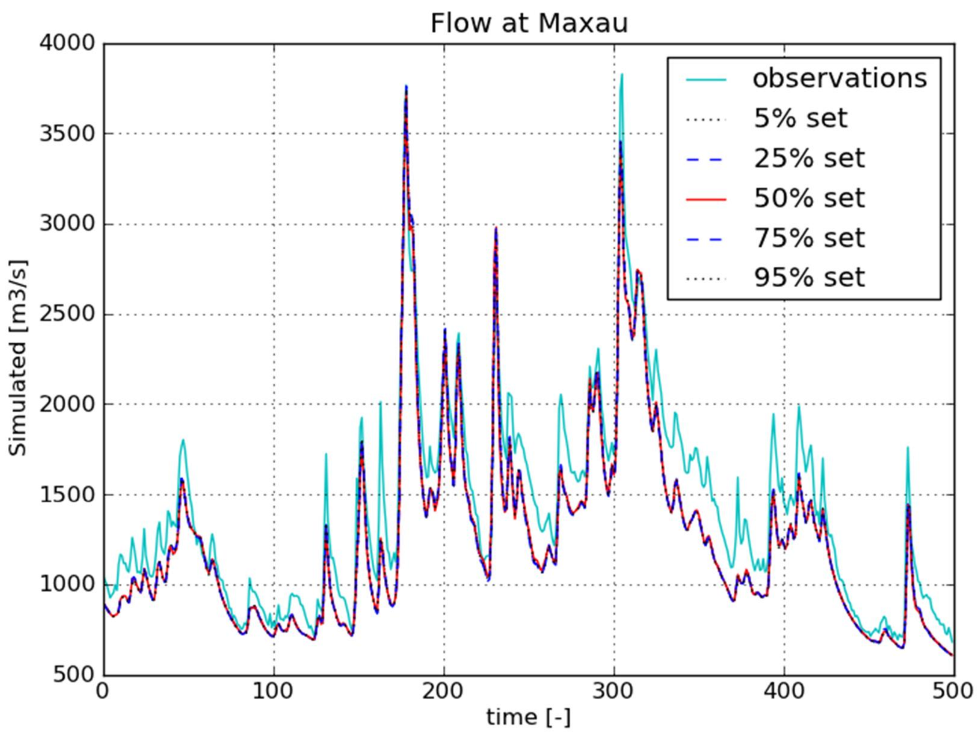


Figure 5.1 Modeled and observed flow at Maxau. The 5 parametersets are identical in the Swiss part of the basin and have not been constrained in a GLUE analysis yet

Table 5.1 Relative difference of simulated and observed flow (sim-obs)/obs of all parameter sets

	GLUE parameterset				
Station	**5%**	**25%**	**50%**	**75%**	**95%**
Maxau	-0.03%	0.08%	-0.05%	-0.04%	0.03%
Worms	-1.95%	-1.79%	-2.06%	-2.00%	-2.03%
Kaub	-3.16%	-3.00%	-3.19%	-3.24%	-3.47%
Andernach	-4.79%	-4.94%	-4.54%	-4.32%	-4.99%
Koeln	-3.95%	-4.05%	-3.68%	-3.42%	-4.19%
Lobith	-3.97%	-3.99%	-3.60%	-3.38%	-4.11%
Maxau (>95%)	-0.23%	0.03%	-0.08%	0.41%	-0.14%
Worms (>95%)	0.70%	0.71%	0.41%	1.03%	0.66%
Kaub (>95%)	-2.75%	-2.67%	-3.08%	-2.10%	-3.03%
Andernach (>95%)	-1.86%	-3.09%	-2.42%	-0.91%	-2.88%
Koeln (>95%)	-0.67%	-1.79%	-1.21%	0.32%	-1.60%
Lobith (>95%)	-1.53%	-1.97%	-1.97%	-0.18%	-1.95%

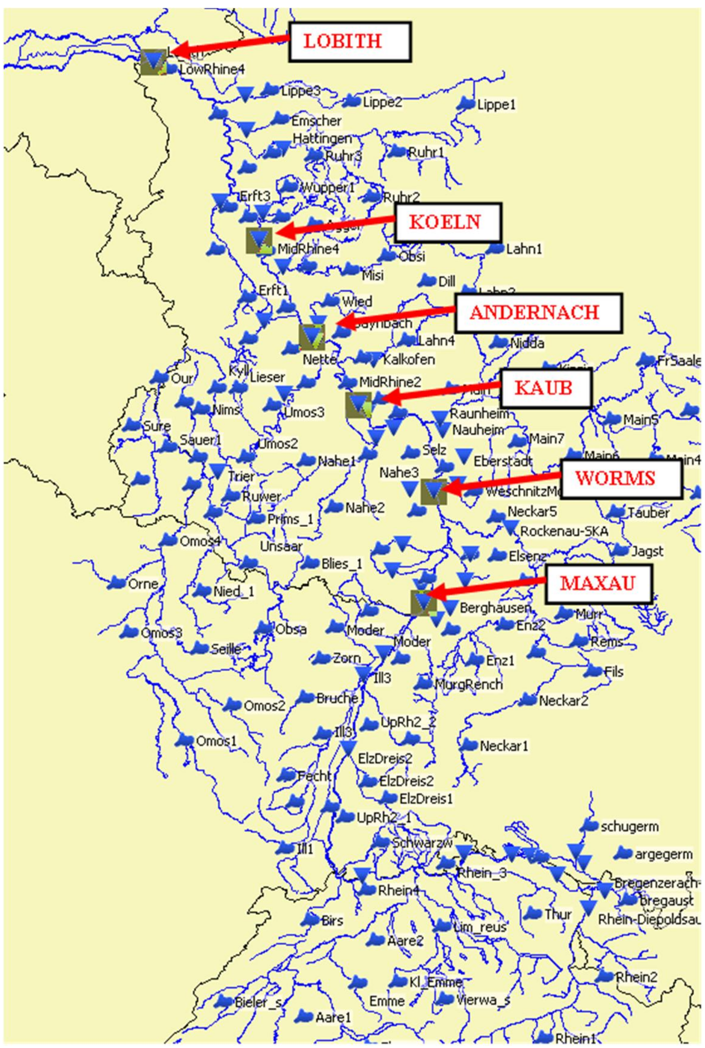


Figure 5.2 Rhine basin with the evaluated discharge locations

Figure 5.3 Double mass (DM) curve of HBV modelled discharge (1985-2006) at Worms

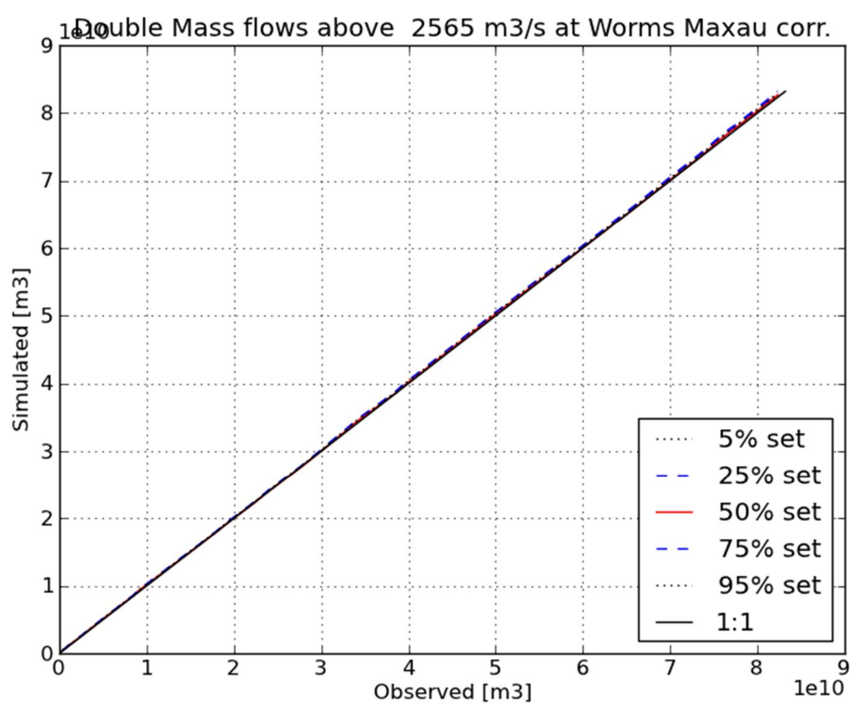


Figure 5.4 DM curve of HBV modelled discharge above 95% percentile(1985-2006) at Worms

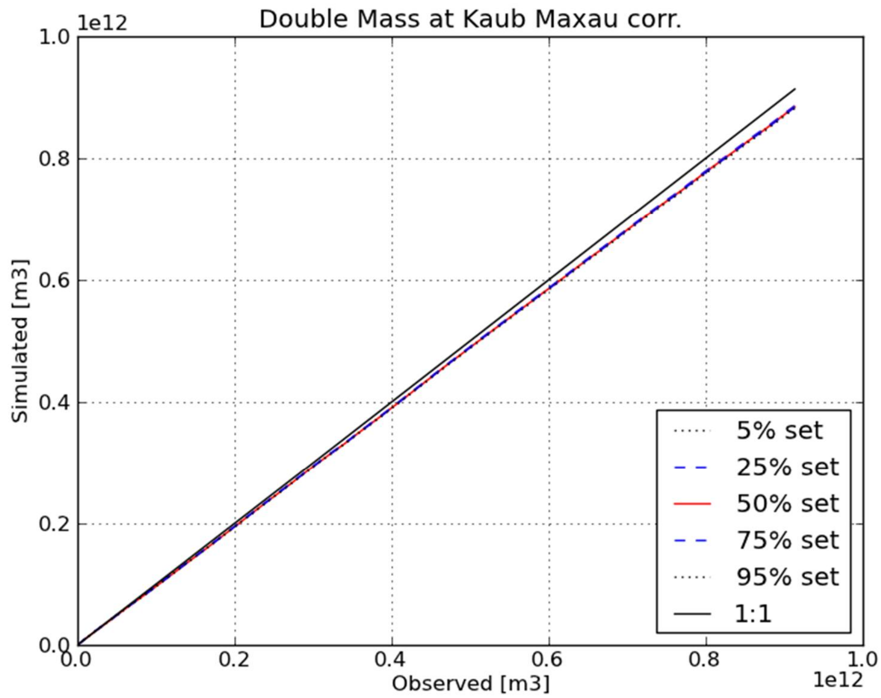


Figure 5.5 DM curve of HBV modelled discharge (1985-2006) at Kaub

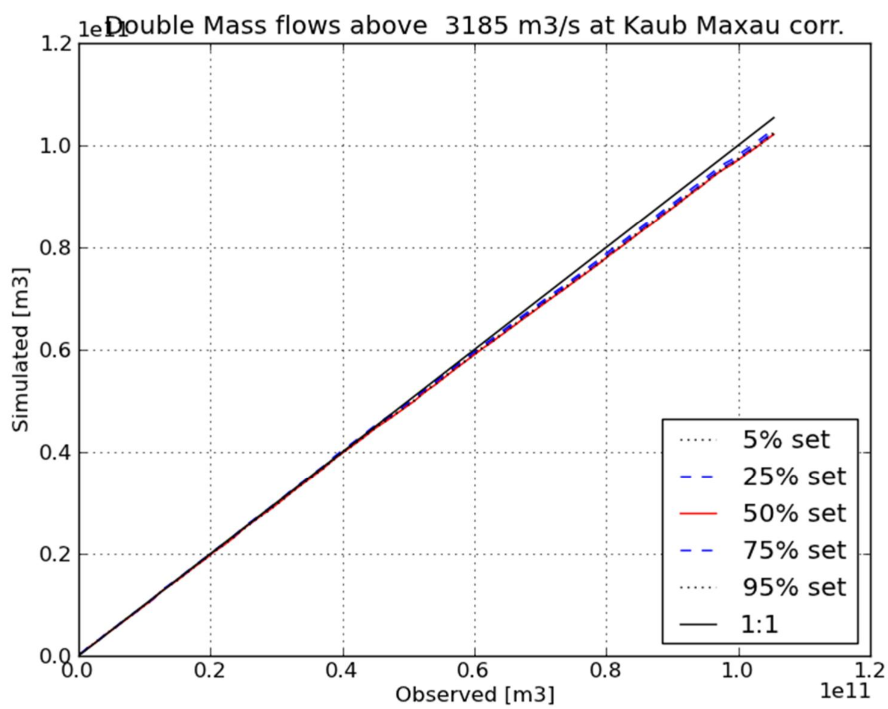


Figure 5.6 DM curve of HBV modelled discharge above 95% percentile(1985-2006) at Kaub

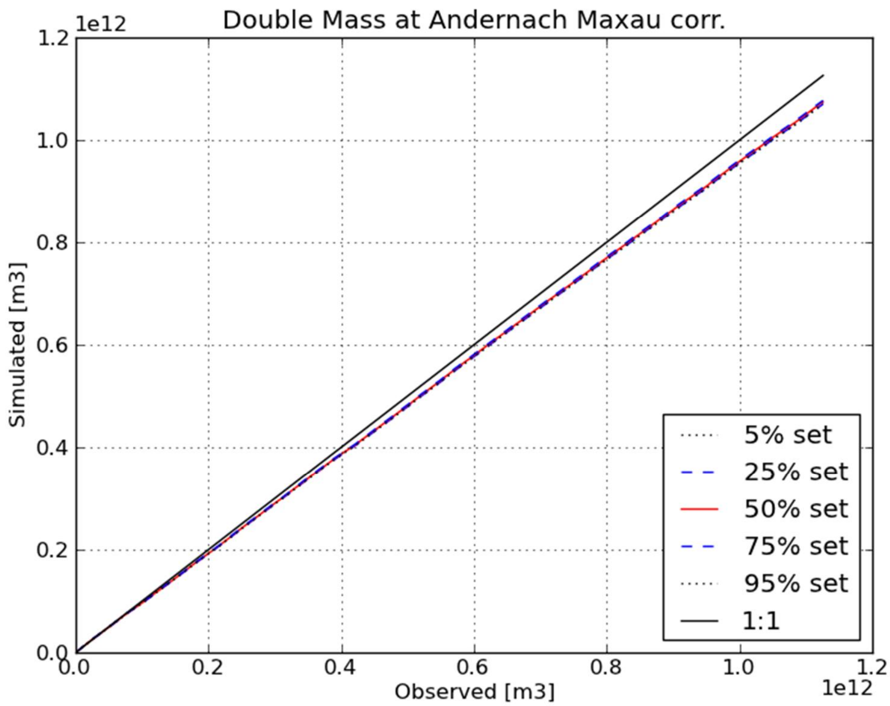


Figure 5.7 DM curve of HBV modelled discharge (1985-2006) at Andernach

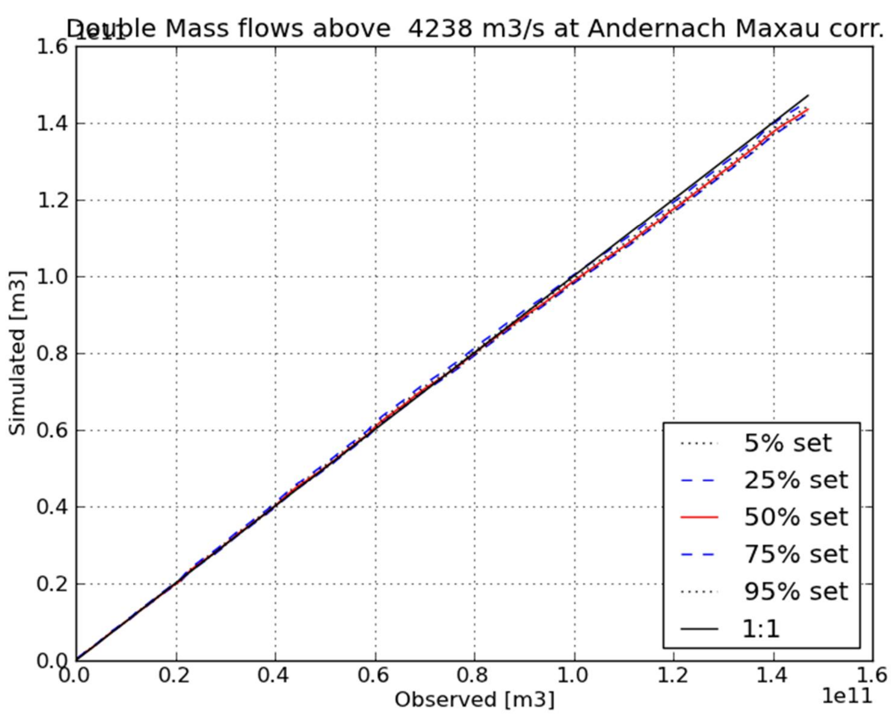


Figure 5.8 DM curve of HBV modelled discharge above 95% percentile(1985-2006) at Andernach

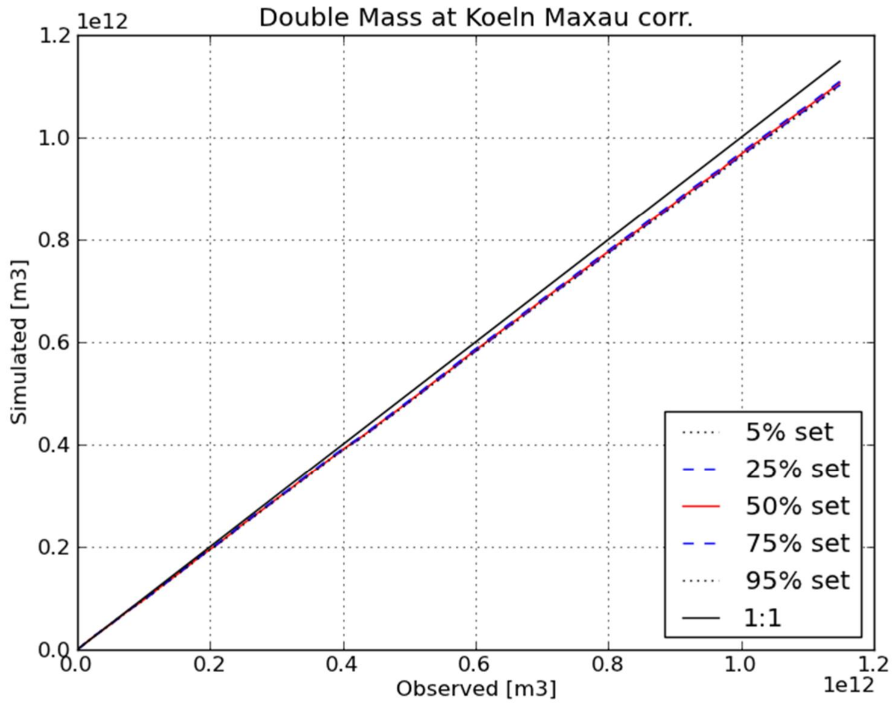


Figure 5.9 DM curve of HBV modelled discharge (1985-2006) at Köln

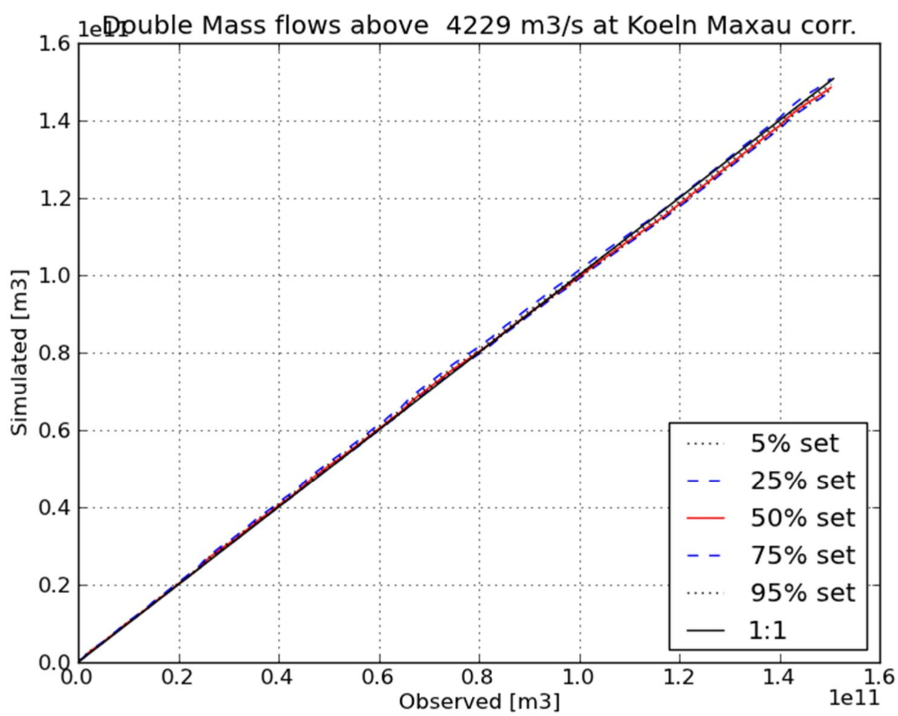


Figure 5.10 DM curve of HBV modelled discharge above 95% percentile(1985-2006) at Köln

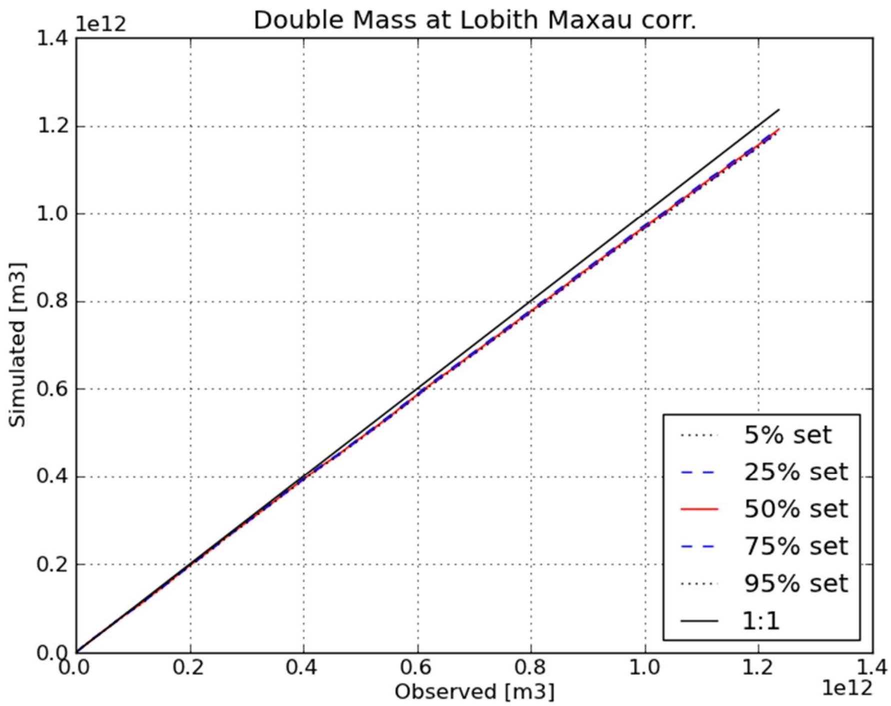


Figure 5.11 DM curve of HBV modelled discharge (1985-2006) at Lobith

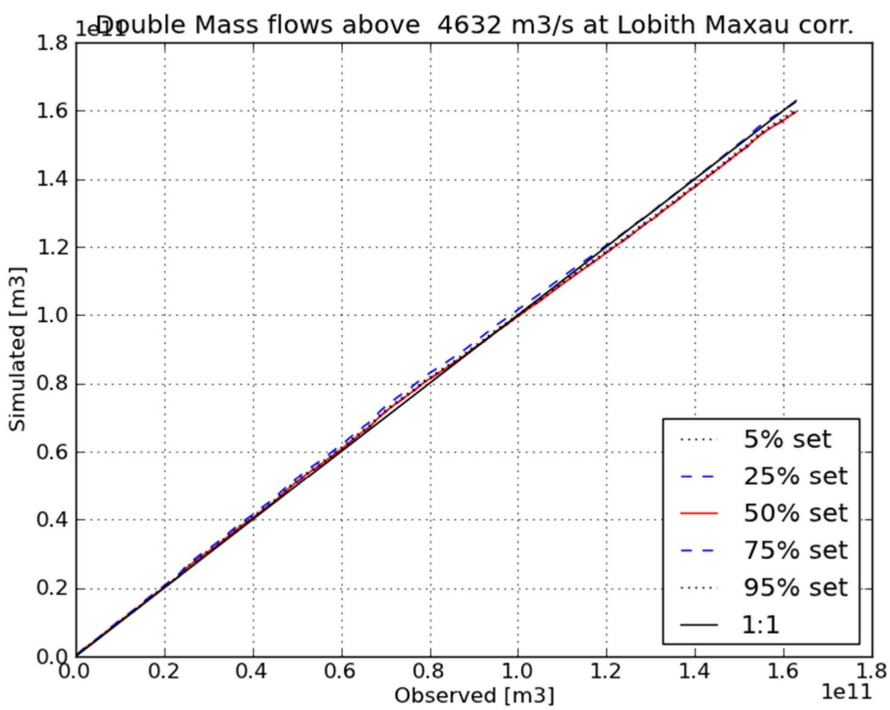


Figure 5.12 DM curve of HBV modelled discharge above 95% percentile(1985-2006) at Lobith

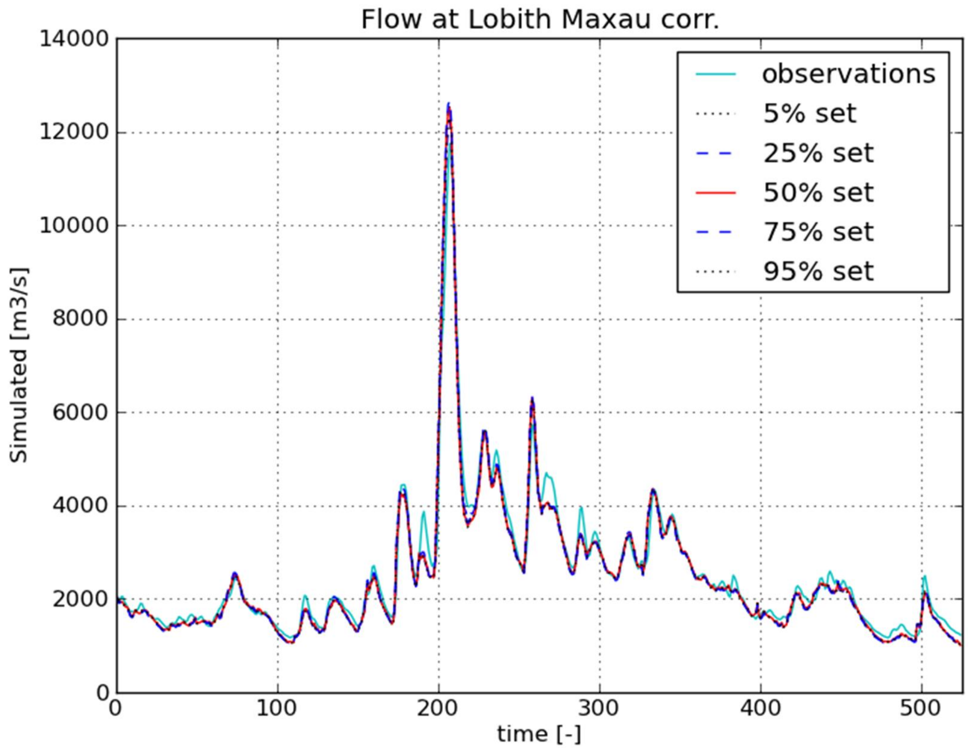
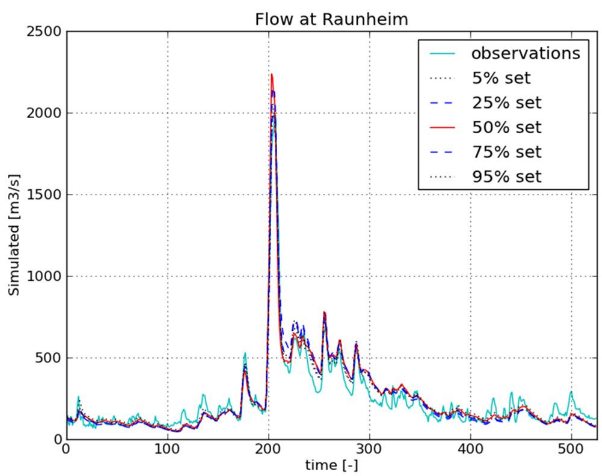


Figure 5.13 Time series of observed a simulated discharge at Lobith during the January 1995 event. The peak is slightly overestimated by HBV, which could be caused by the facto that HBV's routing does not account for flood attenuation in upstream areas

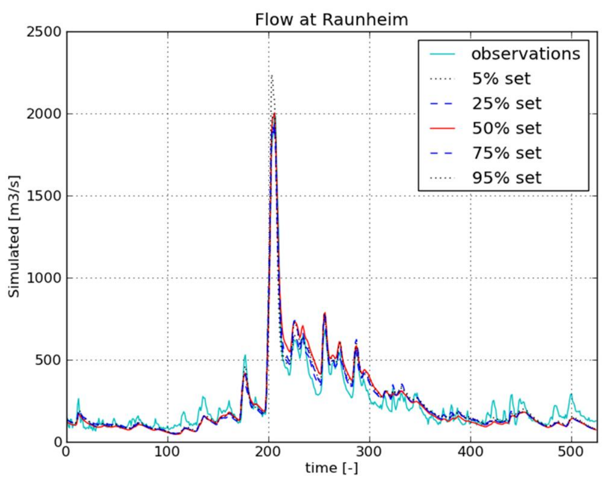
5.3 Sensitivity of parameter selection to chosen return period

The 5 parameter sets were chosen by investigating the extreme discharge value, returned by each parameter set at a return period of 10 years within the extreme value distribution of a run from 01-01-1985 until 31-12-2006. We finally investigated the sensitivity of the choice of the 10-year return period. Instead, we have also made a similar selection over the 5-year and 2-year return period instead. It is expected that other parameter sets are chosen if a different return period is used to condition this choice. However, what matters is whether the resulting discharges and in particular the uncertainty, encapsulated by the 5 parameter sets is significantly different if a different return period is used for the parameter set selection or not. To demonstrate the differences, we show at the large tributaries the simulated discharges at downstream stations as well as at Lobith during the 1995 event. We show this for the selection based on the 10-year, 5-year and 2-year return periods for comparison. The selected stations are Raunheim (Main basin), Rockenau (Neckar basin) and Cochem (Mosel basin).

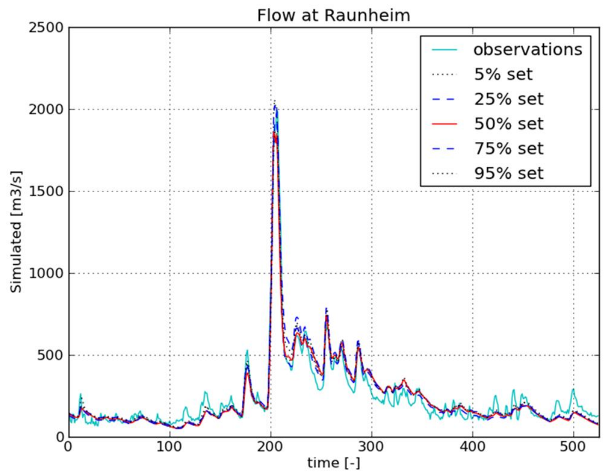
The results are shown in Figure 5.14 until Figure 5.17. Note that the discharge at Cochem is overestimated, but this was already observed and explained during the GLUE analysis (see Section 3.5). The figures show that the discharge, simulated by the selected parameters overall has the same behavior, it can also be seen that when choosing a higher return period, the uncertainty increases slightly, in particular during the highest peaks within the simulation range. This can be observed in particular in the simulations for Raunheim. The 2-year return period shows a lower uncertainty than the 10-year and 5-year return periods. This behavior is difficult to distinguish in the other figures, but can also be observed in the other stations. We therefore recommend keeping the 10-year return period selection as the final selection.



A: 10-year return period

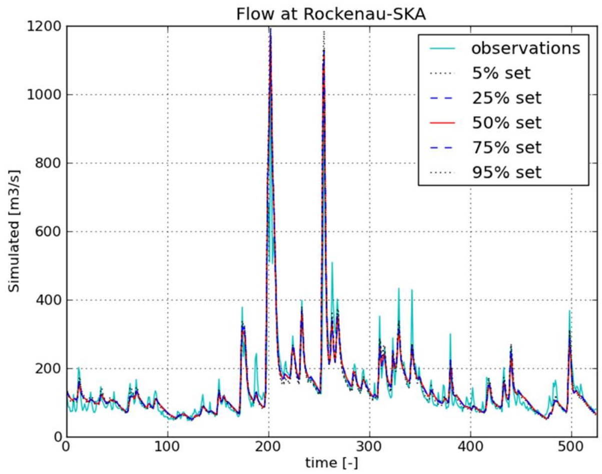


B: 5-year return period

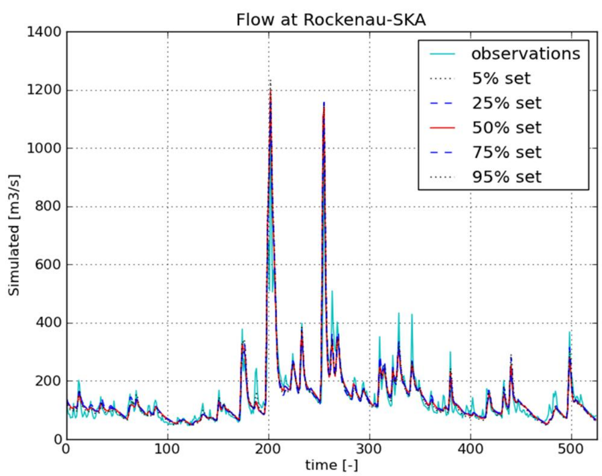


C: 2-year return period

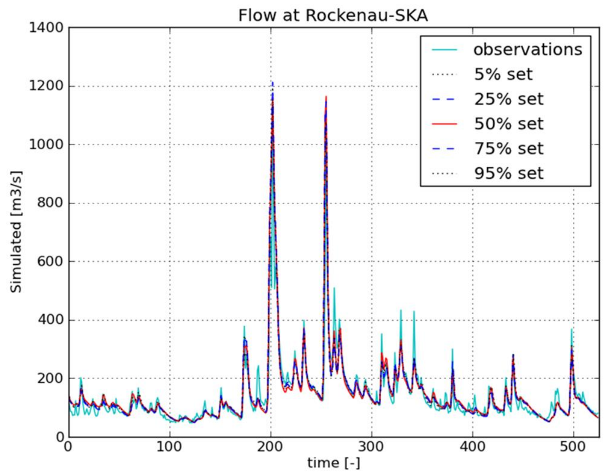
Figure 5.14 Discharge at Raunheim during 1995 event, resulting from 5 selected parameter sets (plotted with the observations) based on A) 10-year; B) 5-year; C) 2-year return period



A: 10-year return period

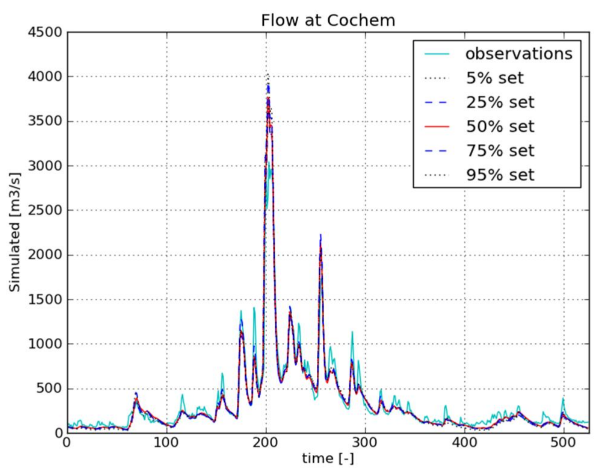


B: 5-year return period

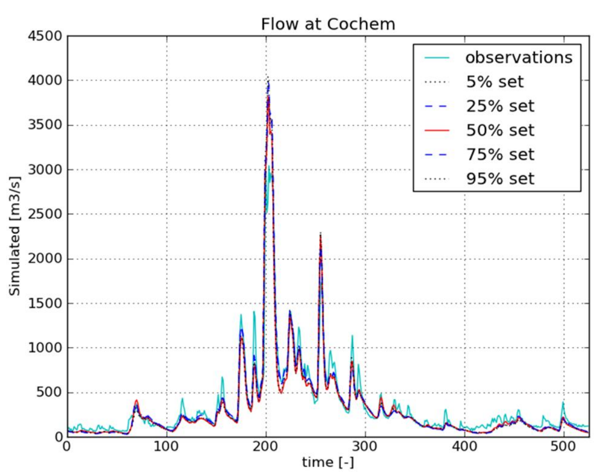


C: 2-year return period

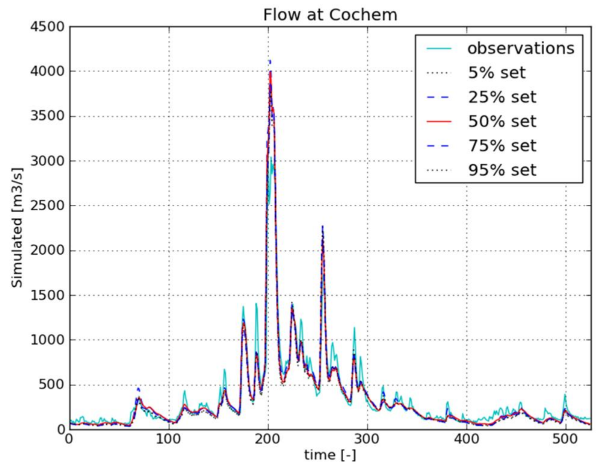
Figure 5.15 Discharge at Rockenau during 1995 event, resulting from 5 selected parameter sets (plotted with the observations) based on A) 10-year; B) 5-year; C) 2-year return period



A: 10-year return period



B: 5-year return period



C: 2-year return period

Figure 5.16 Discharge at Cochem during 1995 event, resulting from 5 selected parameter sets (plotted with the observations) based on A) 10-year; B) 5-year; C) 2-year return period

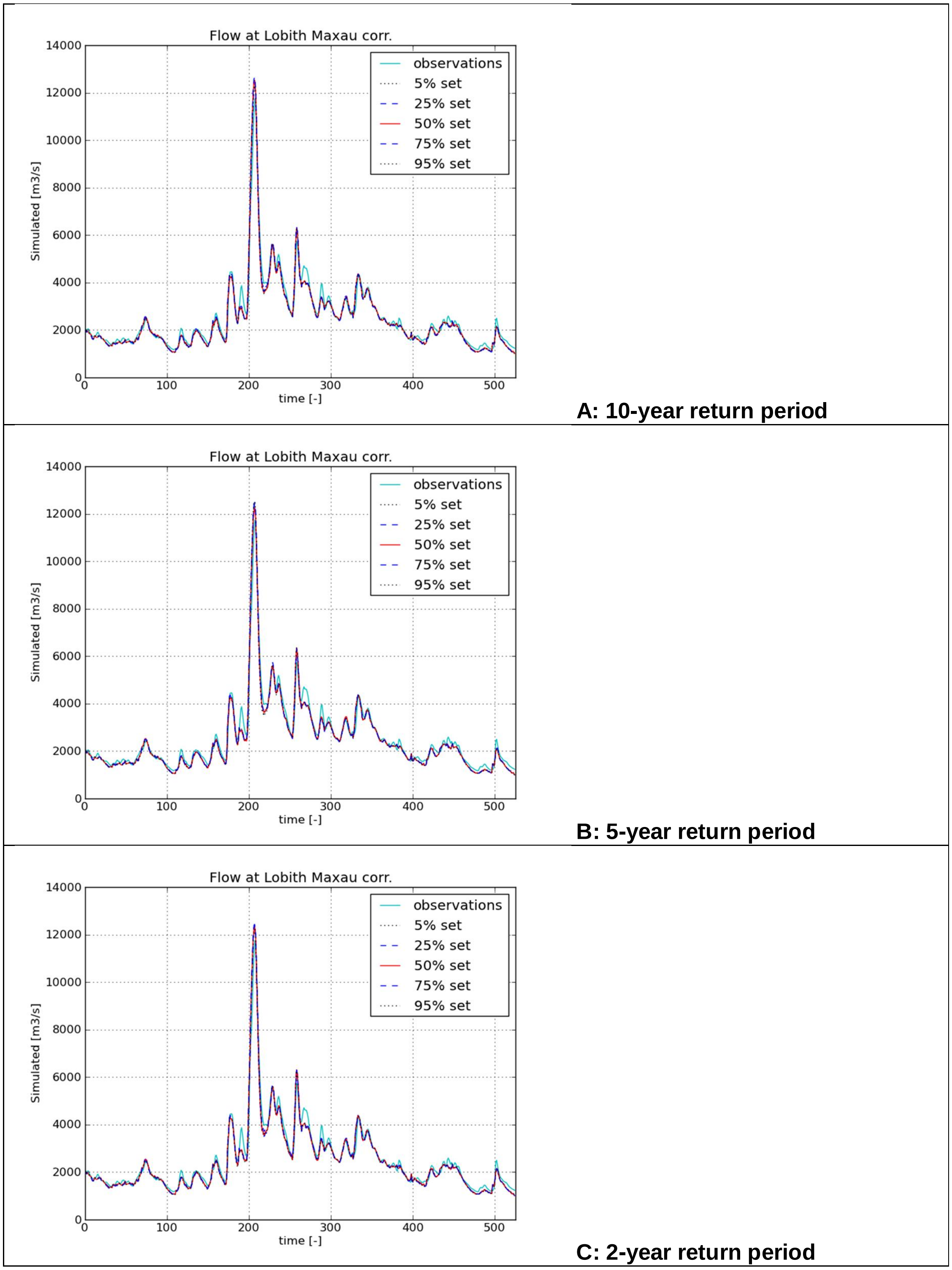


Figure 5.17 Discharge at Lobith during 1995 event, resulting from 5 selected parameter sets (plotted with the observations) based on A) 10-year; B) 5-year; C) 2-year return period

6 Discussion

6.1 Assumptions of the analysis

During the analysis, a number of assumptions were introduced. These are briefly discussed below:

Precipitation correction is not required

We believe that this assumption has been proven to be valid. In fact, we are able to establish good results in most regions of the basin, and any area where a moderate or poor result has been found, an explanation could be found as well. In most cases, the explanation was:

- that man-made interactions or water use are encountered, resulting in reduced performance of HBV. This is because HBV only simulates natural flow; or
- that the hydrological processes in the basin have a faster response time than the chosen (daily time step).

In general limited performance may also be due to the fact that important hydrological processes are not represented or not formulated within HBV according to the way they really occur in the basin.

Independence of parameter sets of upstream basins

During the Monte Carlo sampling in the more downstream end of the 13 basins, a random sampling of the parameter sets that were already conditioned in the upstream basins is performed. This random sampling is done for each of the upstream conditioned basins independently, which may lead to averaging of uncertainty towards the mean of the combined parameter set outputs (central limit theorem). In other words, in one Monte Carlo sample, we may select a parameter set for one upstream basin, which simulates relatively high flows, while we may select a parameter set for another upstream basin which simulates relatively low flows compared to the observations. Together, they may compensate their errors further downstream. This therefore may reduce the uncertainty propagated to the downstream area. Theoretically, we could reduce this effect by exploring all combinations of the upstream parameter sets. This however would drastically increase the required computations for the downstream basin, rendering the experiment infeasible. This issue is still open for debate.

Assumption that each behavioural parameter set is equally likely

This assumption results in a slightly different application of GLUE than the classical GLUE approach. The fundamental difference is that we select a parameter set as behavioural, when it gives a satisfactory value for a number of objective functions, rather than only one (e.g. the Nash-Sutcliffe). This means that we cannot weigh the likelihood of each parameter set such as done in classical GLUE. The only alternative would be to put a weight on each objective function chosen and transform the chosen objectives into one objective function. When this transformation is used, behavioural sets could have been chosen based on a threshold on the combined objective function. The risk of doing this is that a very good value for one objective function may compensate a very poor value for another objective function. Therefore a parameter set may be retained which for instance has a very good volume error closure while misbehaving in terms of the estimation of peak flows. For GRADE it is important that multiple behaviours (volume, overall shape and peak) of the hydrograph are represented satisfactorily.

This is the reason why the multiple objective functions were introduced. The way GLUE was applied is similar to the application shown by Winsemius et al. (2009).

Independence of quantile parameter sets per basin
The 5, 25, 50, 75 and 95% parameter sets are chosen independently from each other, by observing the behaviour of the parameter sets in the individual subbasins of the Rhine. As with the point “Independence of parameter sets of upstream basins”, we should compute all combinations of parameter sets from all basins to make a statistically sounds analysis of accumulated parameter uncertainty throughout the main river. This computationally infeasible. Further study in a smart way to perform sampling from the individual basins is needed to make a more robust choice of the final representative parameter sets.

6.2 Subjective choices

Choice of the acceptance criterion
The choice whether to accept a parameter set based on evidence or not, is always rather subjective in GLUE. There has been many criticism in literature between supporters of formal and informal statistical methods about this subjectivity, but generally, the conclusion was that any uncertainty estimation approach is to some degree subjective (see e.g. Mantovan and Todini, 2006; Beven et al., 2007). Our approach to define whether a parameter set is acceptable was to check whether the uncertainty, left after selection, brackets the observations to a satisfactory degree or not by visual inspection. In case the uncertainty bands seemed overconfident with respect to the observations, we changed the experiment by iteratively stretching the acceptance criteria until the uncertainty brackets the observations to a satisfactory degree. This process results in more parameter sets being accepted and therefore a larger uncertainty band. Although this approach is rather subjective, it means that if for example the combination of the model, input data and output data used gave poor results, the uncertainty due to these poor results is better accounted for by stretching the acceptance criterion.

Zero flow from the Zwischeneinzugsgebieten
We demonstrated in Chapter 5, that this assumption gives satisfactory results within the time series used for the GLUE analysis (1985-2006). The hydrological processes underlying the response of the ZWEs is however not well studied and it could be that during very extreme weather, outside the used measurement series, they will respond differently.The ZWEs are relatively flat areas with deep soils but when saturated, these areas may for instance suddenly generate a lot of saturation overland flow. This requires further study.

7 Conclusions and recommendations

7.1 Conclusions

From the GLUE analysis, we can conclude the following:

- For 13 aggregated catchments (e.g. Main, Neckar) of the Rhine, an uncertainty analysis using GLUE has been performed. The 13 aggregated catchments include the German part of the Rhine basin between Basel and Lobith.
- For each considered sub-catchment, 5 parameter sets were derived from the GLUE analysis. These parameters are selected such, that they reflect as accurate as possible the high flow regime, so that these sets may be used for extreme value estimations in GRADE.
- Although particularly selected for high flow estimation also the overall performance of the parameter sets is good. This is reflected by the high Nash and Sutcliffe coefficients, reached in most parts of the considered sub-catchments.
- Particular catchments such as the upper Main, and parts of the Upper and South Upper Rhine show poor performance. This is likely due to the fact that the hydrological processes occurring in these basins are occurring at a smaller time scale than the model time step (daily).
- The Erft shows poor results as well. This is due to the fact that most of the flow is affected by the lignite mining industry in this area.
- Precipitation correction factors seem to be unnecessary, when good recent rainfall records are used as model inputs (HYRAS 2.0). Therefore, precipitation correction factors have been excluded from all parameter sets.
- A water balance over the Rhine branches, using the new GLUE parameter sets and comparing against observations in the Rhine's main stem shows good results. In this water balance, any flow from the uncalibrated Zwischeneinzugsgebieten (ZWE) in between the 13 GLUE-constrained catchments and the Rhine's main stem was excluded. As expected, HBV therefore underestimates the flow over the complete time period considered (20 years), but the amount of underestimation is rather small (about 4% over the full simulated time series at Lobith). High flows are even slightly overestimated. This could be due to the fact that no attenuation due to upstream flooding and retention is considered in HBV for the main stem of the Rhine.

7.2 Recommendations

- For the long term (beyond the scope of GRADE), we recommend that the hydrological behaviour of the ZWEs is studied in more detail and that the hypothesis that most of the flow from these catchment comes from groundwater is rigorously tested.

8 References

Berglov et al (2009). Improvement HBV model Rhine in FEWS: Final Report. **112**.

Haylock, M.R., Hofstra, N., Klein Tank, A.M.G., Klok, E.J., Jones, P.D. and New, M., 2008. A European daily high-resolution gridded data set of surface temperature and precipitation for 1950-2006. Journal of Geophysical Research, 113, D20119, doi: 10.1029/2008JD010201.

Kramer, N., Beckers, J., Weerts, A. (2008). Generator of Rainfall and Discharge Extremes: Part D & E, Deltares.

Kramer, N. and Schroevers, M., 2008, Generator of Rainfall and Discharge Extremes, Part F, Deltares report, 2008, Delft.

Lindström, G., Johansson, B., Persson, M., Gardelin, M. and Bergström, S., 1997. Development and test of the distributed HBV-96 hydrological model. Journal of Hydrology, 201: 272-288.

Mantovan, P. and Todini, E., 2006. Hydrological forecasting uncertainty assessment: Incoherence of the GLUE methodology. Journal of Hydrology, 330(1–2), p. 368–381

Ogink, H.J.M. (2012) Design discharge and design hydrograph computation for Meuse and Rhine rivers: Comparison of computational methods, Delft, 2012

Rauthe, M., Steiner, H., Riediger, U., Mazurkiewicz, A. and Gratzki, A., 2013. A central European precipitation climatology – Part I: Generation and validation of a high resolution gridded daily data set (HYRAS). Meteorologische Zeitschrift, 22: 235-256.

SMHI (2008). IHMS: Integrated Hydrological Modelling System (Manual: version 4.5), SMHI.

SMHI (2010). HBV-CH Documentation and Manual: For use with the FEWS system at FOWG, SMHI.

Van Verseveld, W., Hegnauer, M., 2013, Generalised Likelihood Uncertainty Estimation for the daily HBV model in the Rhine Basin Part B: Switzerland, Deltares report, Delft 2013.

Weerts, A. and van der Klis, H., 2006. Reliability of the Generator of Rainfall and Discharge Extremes (GRADE): An exploratory study on uncertainty in the hydrological parameters, a GLUE analysis. Report of project Q4268 for RIZA, WL|DelftHydraulics,Delft, The Netherlands.

Winsemius, H.C.W., Kramer, N., De Keizer, O., 2009, GRADE 2009, Deltares report, Delft 2009.

Wit de, K.M. & Buishand, A., 2007. Generator of Rainfall And Discharge Extremes (GRADE) for the Rhine and Meuse basins, Lelystad, The Netherlands: RWS RIZA

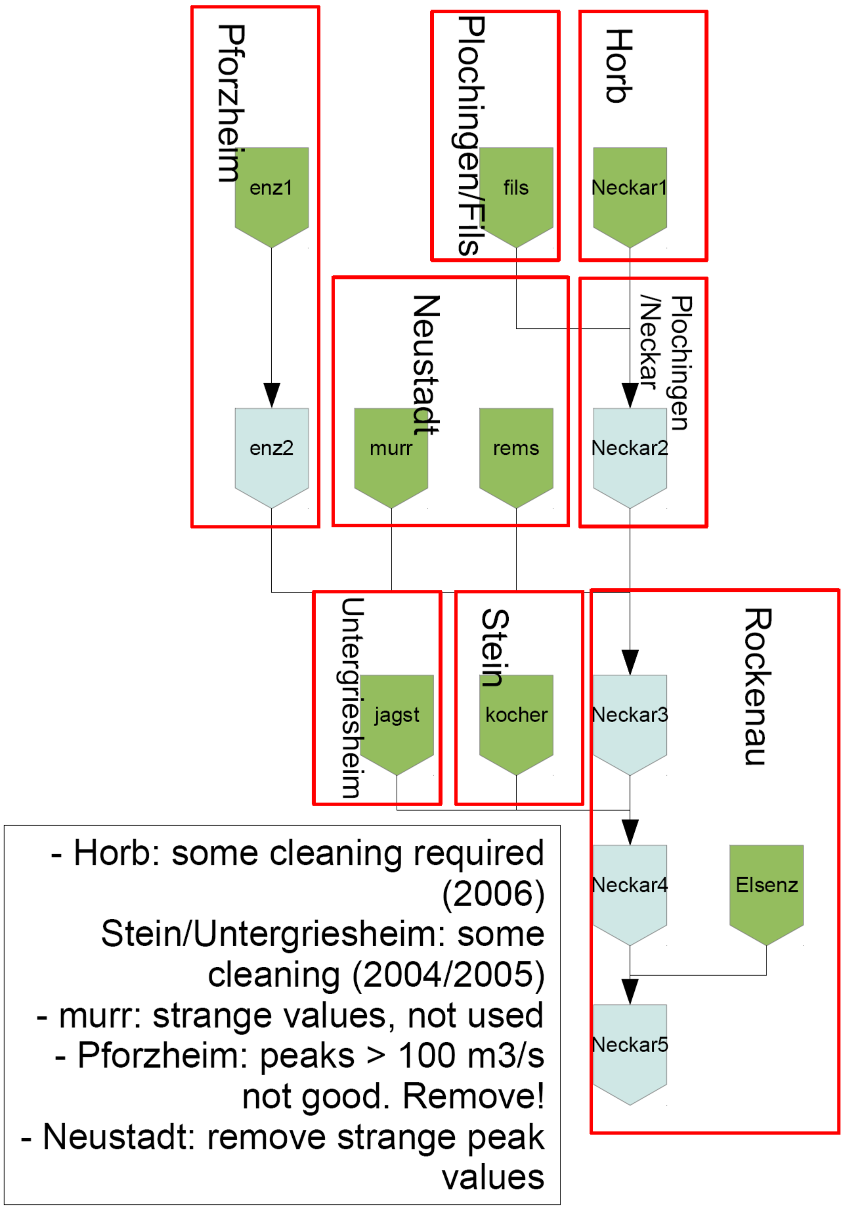
A Flow diagrams per subcatchment

For each subcatchment of interest, a flow diagram has been prepared for the GLUE procedure. The flow diagram indicates the connections between the HBV units. Furthermore, red boxes with the name of a gauging station are drawn. The HBV units, encapsulated in each box are taken together in calibration and are calibrated against the indicated station's discharge series. Green HBV units are located at the boundary of the basin and therefore have no upstream boundary condition. These units typically need to be calibrated first. Then a downstream neighbour may be calibrated using the upstream derived acceptable parameter sets to provide a boundary condition. This procedure is repeated until the most downstream point in the subcatchment is reached.

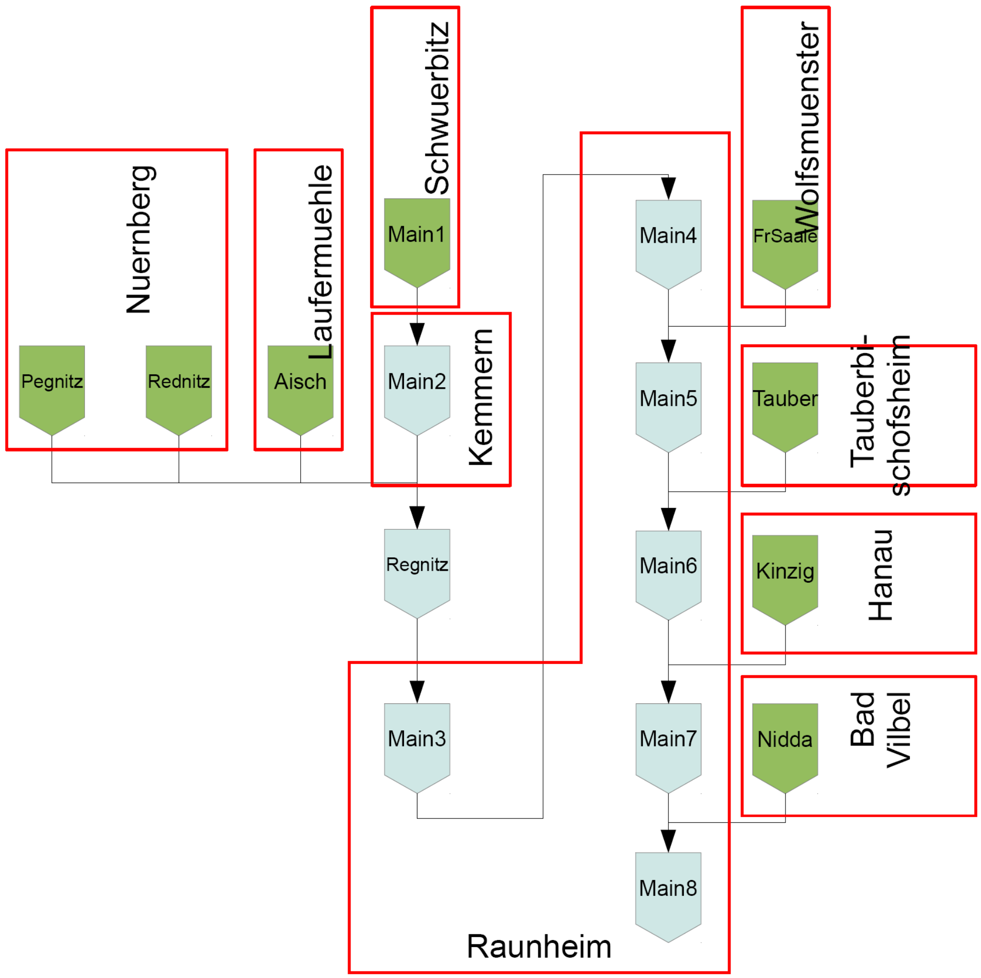
The red HBV units are part of the main stem of the Rhine and are considered ZWE's and therefore not part of the GLUE analysis.

The discharge data used for the GLUE analysis was analysed first to check for missing and incorrect data. This screening (or cleaning) of the data was necessary to get a good performance. In the flow diagrams, for each station which has been "cleaned" from incorrect or missing data, a remark is made about the cleaning process.

A.1 Neckar


Pforzheim
enz1
enz2
Plochingen/Fils
fils
Horb
Neckar1
Neustadt
murr
rems
Plochingen
/Neckar
Neckar2
Untergriesheim
jagst
Stein
kocher
Rockenau
Neckar3
Neckar4
Elsenz
Neckar5
- Horb: some cleaning required (2006)
Stein/Untergriesheim: some cleaning (2004/2005)
- murr: strange values, not used
- Pforzheim: peaks > 100 m3/s not good. Remove!
- Neustadt: remove strange peak values

A.2 Main

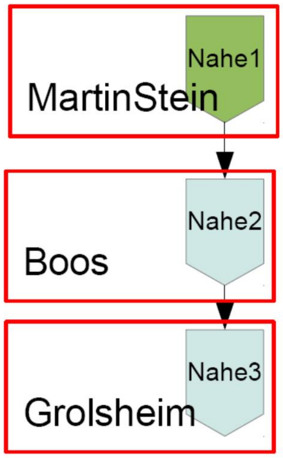


- Quality of Trunstadt/Schweinfurt/Steinbach very doubtful. Suggest to not use them at all. Otherwise:

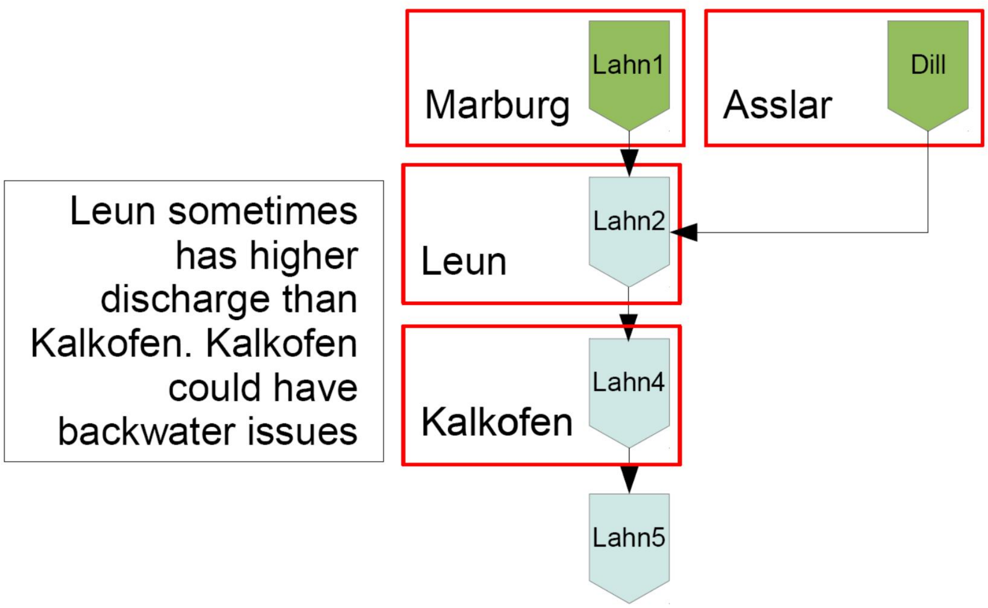
- Trunstadt: remove 01-May-1990 until 01-Nov-1990
- Schweinfurt: remove 01-11-2011 until 01-05-1993
- Steinbach: remove 01-Apr-1991 until 20-Dec-1991
- Trunstadt/Schweinfurt/Steinbach: remove 01-Dec 1993 until 2008

- Kleinheubach: Remove everything after 2002, before 2002, some cleaning required
- Laufermuehle: negative values must be removed
- Bad Vilbel: remove negative values
- Hanau: remove Jan-2002

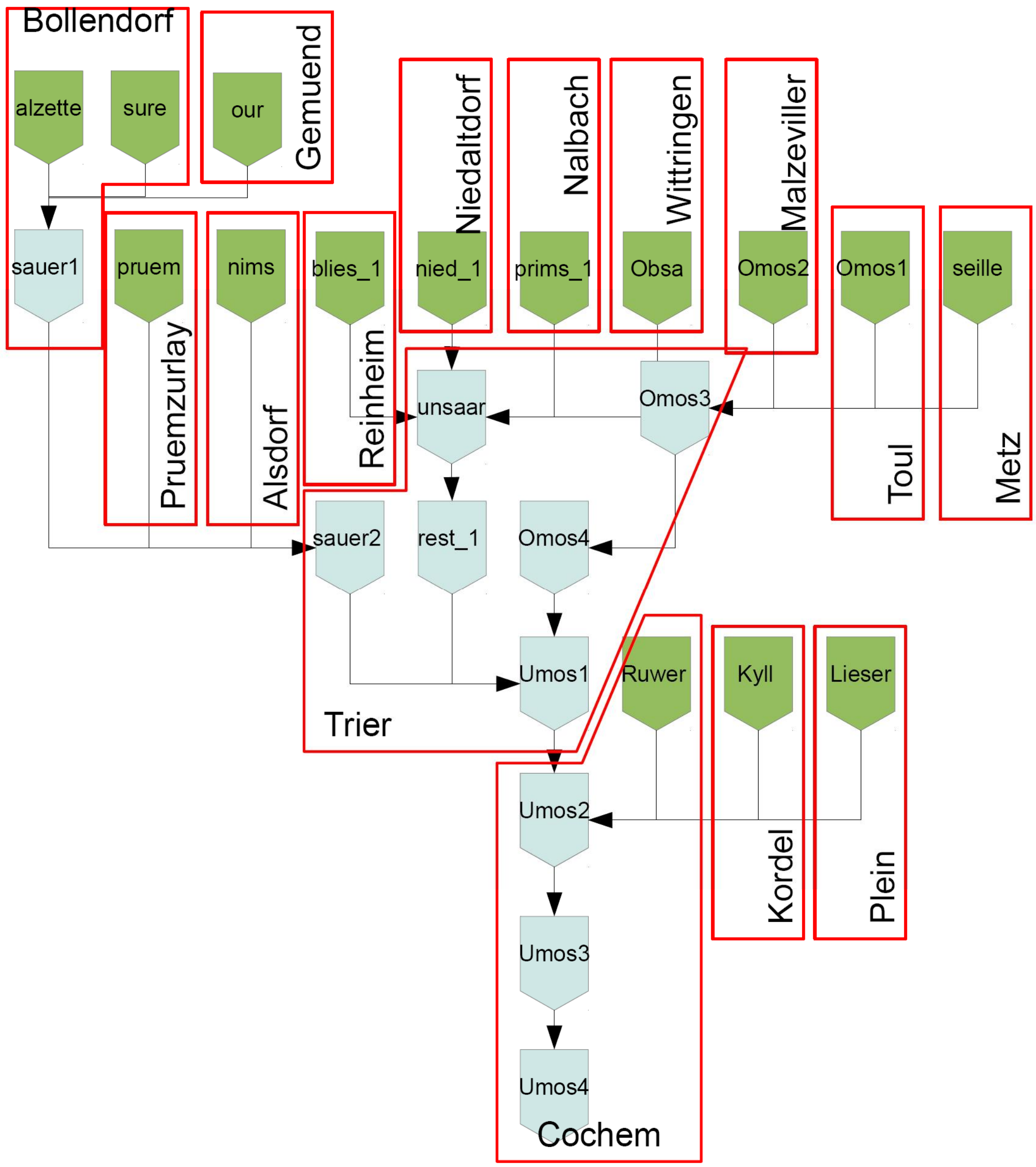
A.3 Nahe



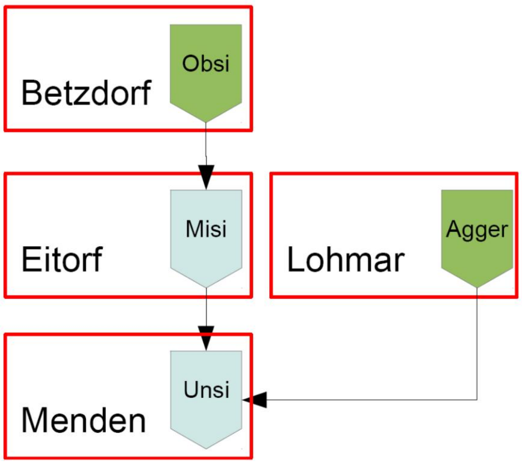
A.4 Lahn



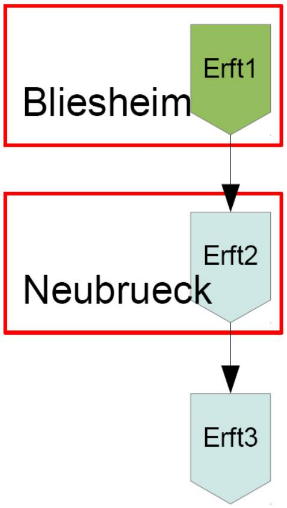
A.5 Moselle


Bollendorf
alzette
sure
our
Gemuend
Niedaltdorf
Nalbach
Wittringen
Malzeviller
sauer1
pruem
nims
blies_1
nied_1
prims_1
Obsa
Omos2
Omos1
seille
Pruemzurlay
Alsdorf
Reinheim
unsaar
Omos3
Toul
Metz
sauer2
rest_1
Omos4
Umos1
Ruwer
Kyll
Lieser
Trier
Umos2
Kordel
Plein
Umos3
Umos4
Cochem

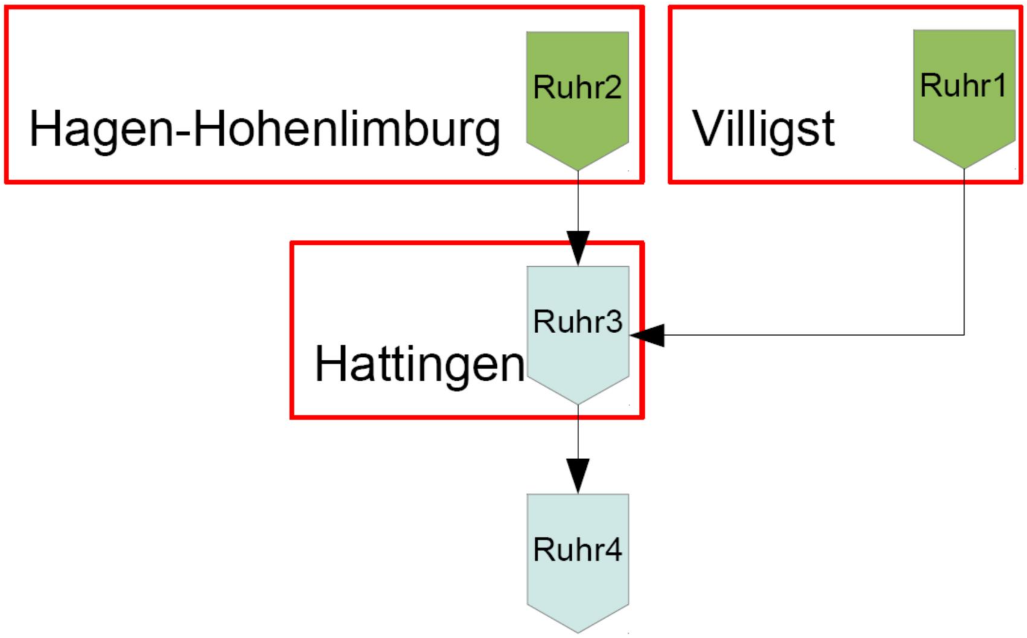
A.6 Sieg



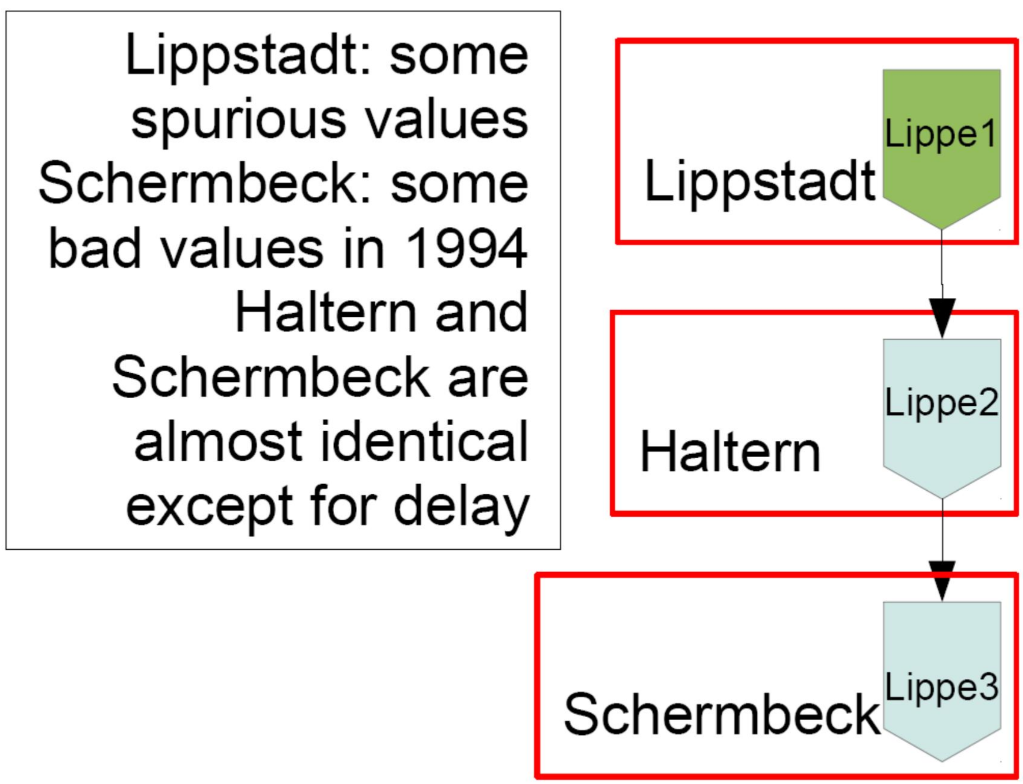
A.7 Erft



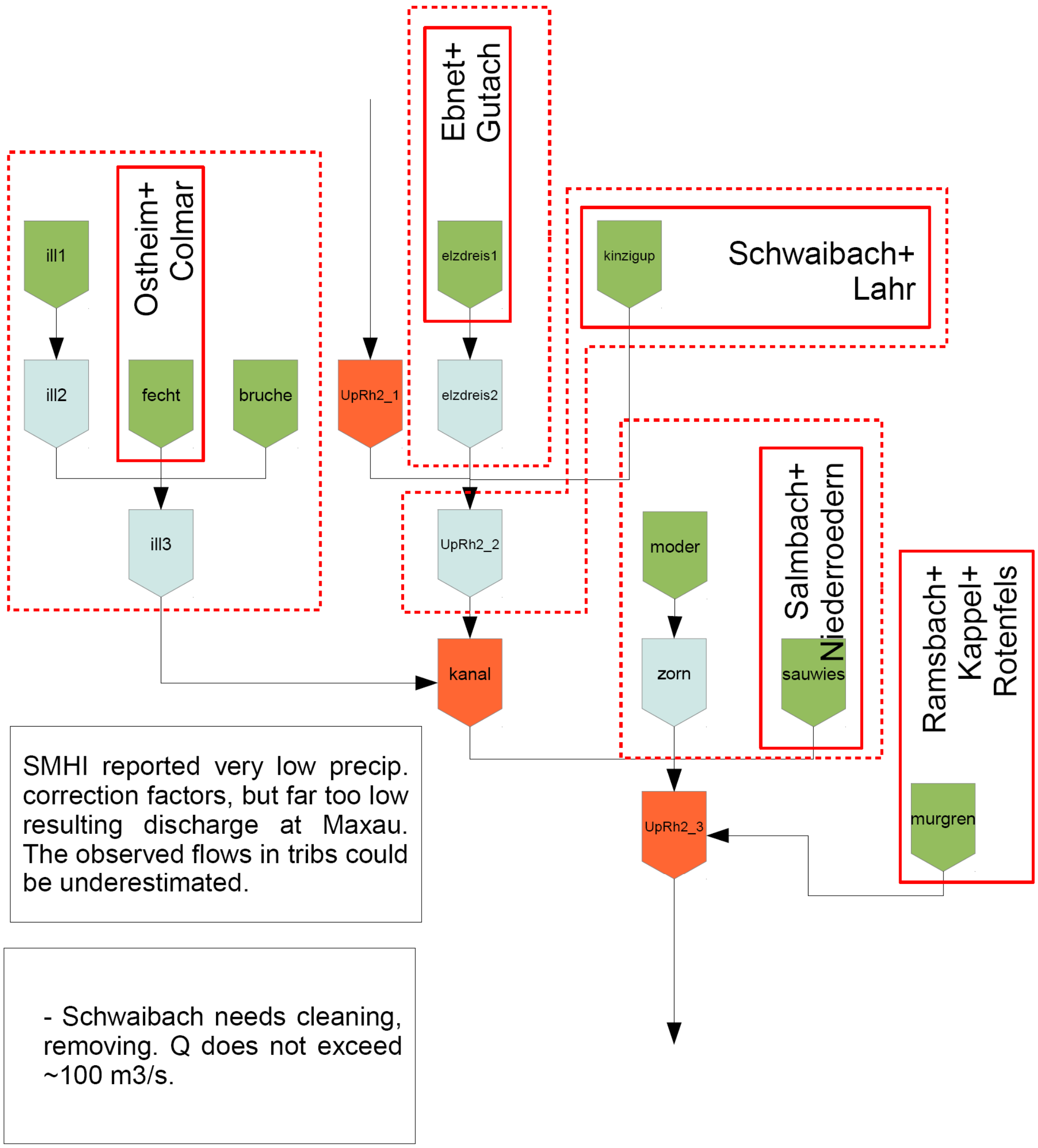
A.8 Ruhr



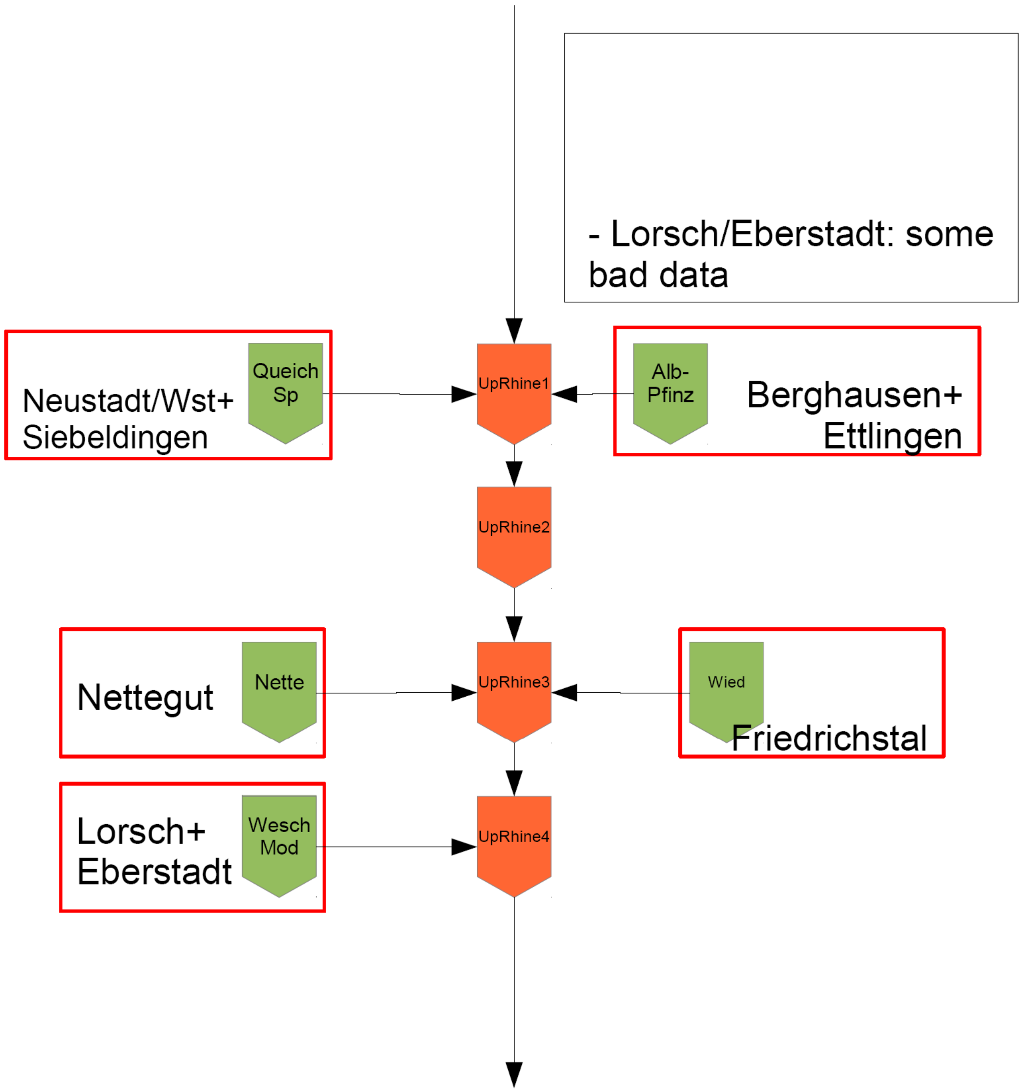
A.9 Lippe



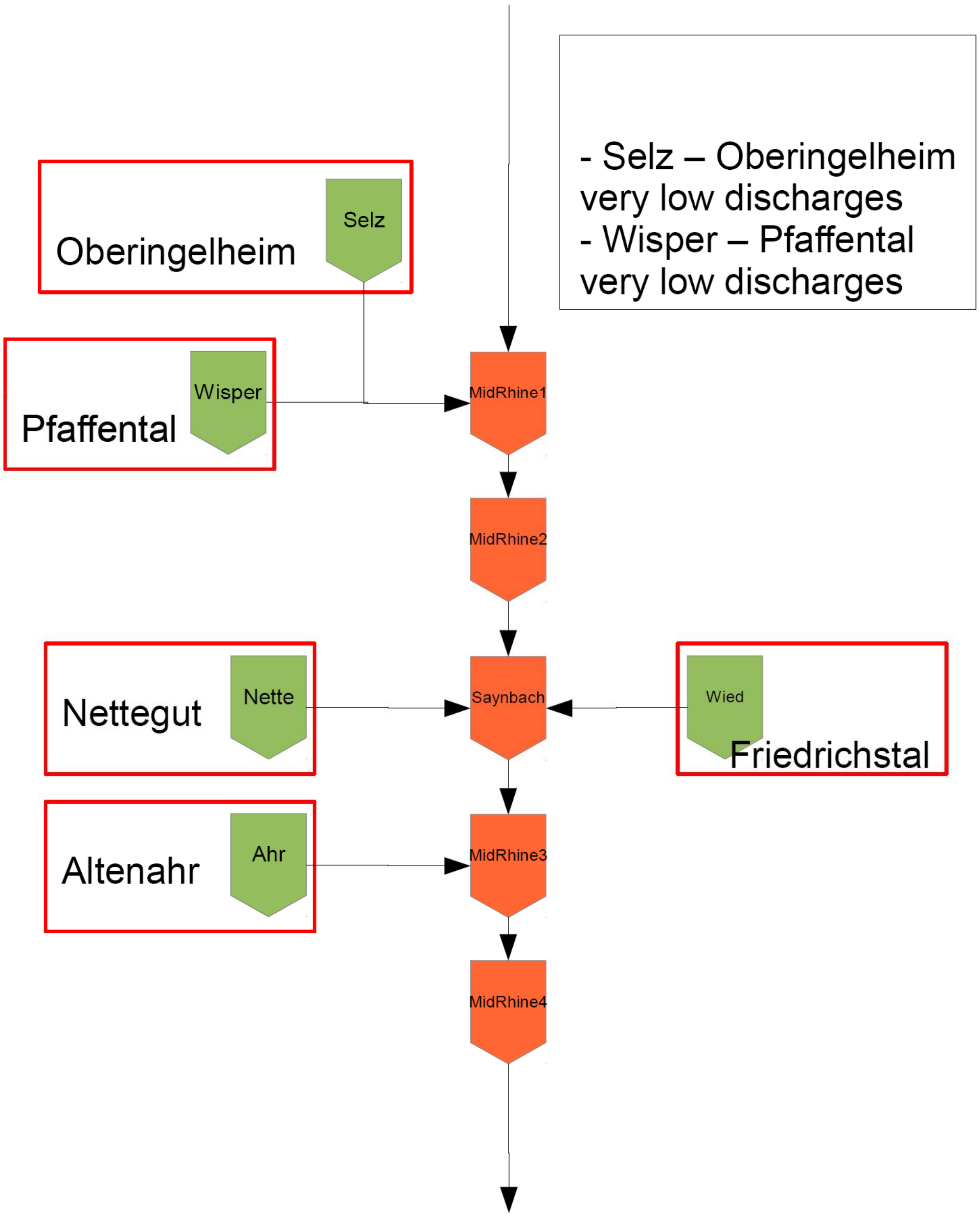
A.10 Southern Upper Rhine



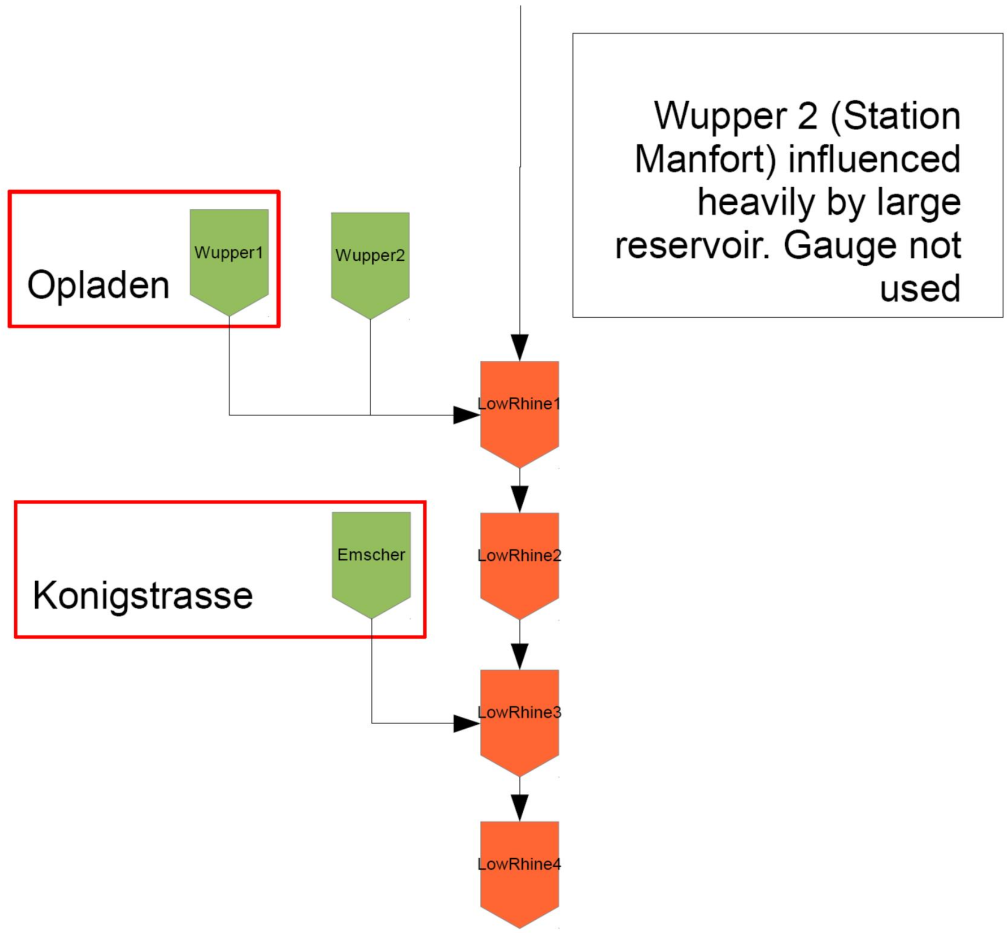
A.11 Upper Rhine



A.12 Middle Rhine



A.13 Lower Rhine



A.14 Cleaning of datasets

Table A.1 Alteration of data in the Neckar subcatchment

Station	Description of alterations
Horb	Strange peaks deleted in 2005, 2006 and 2007 (next to 2006)
Stein	Strange peaks and shifts deleted in 2003, 2004 and 2005.
Untergriesheim	Strange peaks deleted.
Pforzheim	Quite a lot had to be removed, through whole dataset. Not only peaks, also strange shifts. From 2006 onward, peaks higher than 100 are observed.
Neustadt	Deleted some strange peaks, dataset seems quite ok, though

Table A.2 Alteration of data in the Moselle subcatchment

Station	Description of alterations
Nalbach	Strange peaks removed. Also some wobbly peaky measurements removed. In 1997, some 'plateau' values. Not removed.

Table A.3 Alteration of data in the Lippe subcatchment

Station	Description of alterations
Lippstadt	Removed some spurious values (-77.00)
Schermbeck	Removed large parts of 1994 that had plateau values
Haltern	Looked ok

Table A.4 Alteration of data in the Southern Upper Rhine catchment

Station	Description of alterations
Schwaibach	In 2006, suddenly peaks above 100 present. They do not look strange. Before this time, peaks around 100 are strange and topped off. Strange peaks are removed.

Table A.5 Alteration of data in the Upper Rhine catchment

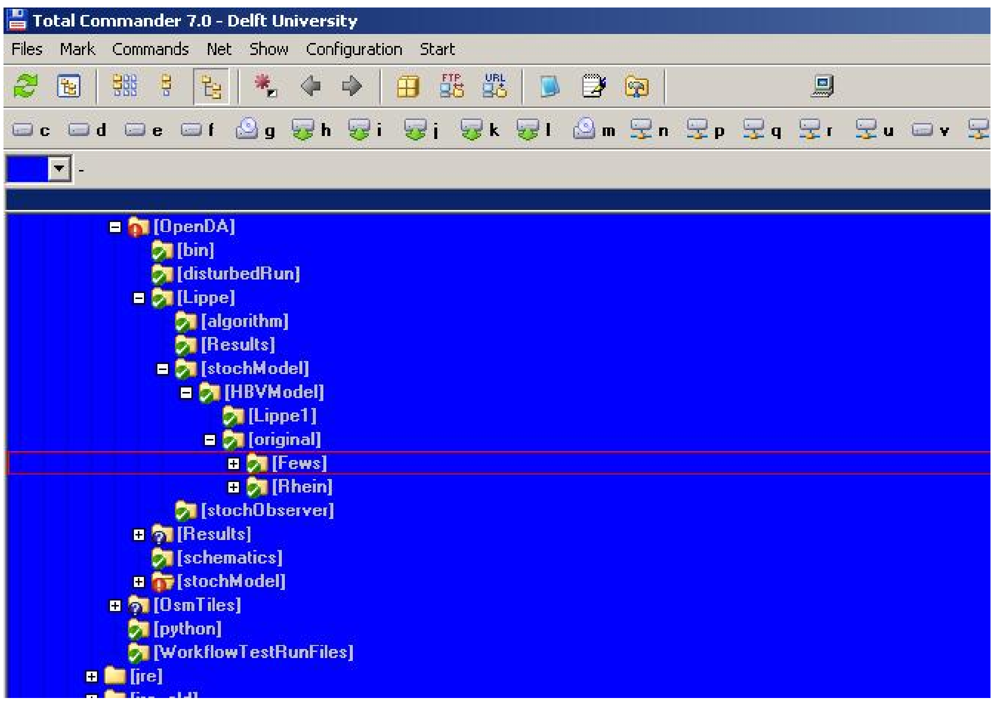
Station	Description of alterations
Lorsch	Deleted steadily sloping or straight values that looked artificial
Eberstadt	Deleted steadily sloping or straight values that looked artificial. Not so much deleted.

B OpenDA setup

GLUE has been implemented in OpenDA and executed from upstream to downstream in each subcatchment of interest. In this appendix, the general way of setting up GLUE for a subcatchment is outlined.

B.1 Folder structure

To ensure the GLUE project can be performed in an organised fashion, a folder structure has been prepared. In this Appendix, we use the Lippe as an example and show the folder structure below. Some more details are given below:



B.1.1 Algorithm

This folder contains details about the algorithm used. We started with an optimisation algorithm called 'Dud' to setup the experimental structure. This was because the GLUE algorithm was not implemented yet at the onset of the sub-project. During the course of this sub-project, GLUE has been implemented and the used algorithm changed. The file structure is very simple. An example is given below. The most important parameter is the ensembleSize (i.e. the number of runs). Note that there is a memory limit to the amount of data, which can be handled by OpenDA. For large experiments (i.e. long time series and many time series), it is recommended to limit the amount of members to e.g. 1000 and run the experiment multiple times, each time saving to a different resultfile.

```
<?xml version="1.0" encoding="UTF-8"?>
<GLUEConfig>
  <likelihoodFunction class="org.openda.algorithms.RMSECostFunction" />
  <ensembleSize>500</ensembleSize>
```

```xml
</GLUEConfig>
```

B.1.2 Results

In this folder, results of each submodel case are saved

B.1.3 stochModel

This folder contains the hydrological model and the configuration used to run it. It also contains the configuration files which determine which parameters are varied and with which distribution function. For GLUE, a uniform sampling space with fixed minimum and maximum sampling boundaries has been added. Uniform sampling is the least informative way of sampling and therefore the most appropriate for GLUE. The configuration files are rather complex. The files are detailed below:

B.1.3.1 hbvModel.xml

This file gives some general information about location of files and variables.

```xml
<?xml version="1.0" encoding="UTF-8"?>
<blackBoxModelConfig xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xmlns="http://www.openda.org" xsi:schemaLocation="http://www.openda.org
http://schemas.openda.org/blackBoxModelConfig.xsd">
	<!-- Specify wrapper configuration for which this model is defined -->
	<wrapperConfig>
		<file>hbvWrapper.xml</file>
	</wrapperConfig>
	<!-- Specify wrapper configuration for which this model is defined -->
	<aliasValues>
		<alias key="templateDir" value="../original"/>
		<alias key="instanceDir" value="../work"/>
		<alias key="modelRootDir" value="./Rhein"/>
		<alias key="piTimeseriesResultFile" value="./Fews/update/FewsRES.xml"/>
	</aliasValues>
	<exchangeItems>
		<vector id="allElementsFromIoObject" ioObjectId="HbvParameters"/>
		<vector id="allElementsFromIoObject" ioObjectId="HbvPiTimeSeriesResults"/>
	</exchangeItems>
	<doCleanUp>true</doCleanUp>
</blackBoxModelConfig>
```

B.1.3.2 stochModel.xml

This file determines how parameters are varied. This can be done in two ways:

- drawing from a uniform distribution, given in a separate xml file uncertaintySpecification.xml
- drawing from a file with parameter sets. This is needed when upstream basins have already been calibrated with GLUE and behavioural sets saved. The behavioural sets are saved in a csv-file which looks as below.

```
runid, fc, lp, beta, alfa, khq, perc
923, -195.04515, -0.19346, -0.88141, -0.19792, 0.03375, 2.33520
939, -148.31140, 0.13983, -0.40364, -0.48997, 0.03695, 1.12274
```

1182, -189.34817, 0.03208, -0.47016, -0.31010, 0.01934, 0.72413
…..
….
…..

The parameter names in each csv files are mapped onto parameters in specific sub-catchment folders. In this example, parameter 'fc' from the csv-file is mapped onto 'fc.lippe1' in the xml configuration. The extention 'lippe1' means that a sub-catchment folder should be located in the model-instance directory with the name 'lippe1'. The parameter fc will be imposed in this folder in the bmod.par file. An example of the stochModel.xml is given below. In the example, parameters for subbasin 'lippe1' are drawn from a csv-file called paramsFromLippe1_H-RN-BFG036.Q.m.csv. Parameters for lippe2 are drawn from uniform distributions, defined in 'uncertaintySpecification.xml'.

```
<?xml version="1.0" encoding="UTF-8"?>
<!-- edited with XMLSpy v2009 sp1 (http://www.altova.com) by ICT (Stichting Deltares) -->
<blackBoxStochModel xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xmlns="http://www.openda.org" xsi:schemaLocation="http://www.openda.org
http://schemas.openda.org/blackBoxStochModelConfig.xsd">
	<modelConfig>
		<file>./hbvModel.xml</file>
	</modelConfig>
	<vectorSpecification>
		<parameters>
			<!-- Below the parameters belonging to the already calibrated areas
upstream of Umos1 are given (see also Moselle1 config) -->
			<uncertaintyModule
className="org.openda.uncertaintyModels.ParameterDrawerFactory" workingDirectory=".">
				<configFile>paramsFromLippe1_H-RN-
BFG036.Q.m.csv</configFile>
				<exchangeItems>
					<exchangeItem id="fc">
						<modelExchangeItem id="lippe1.fc"/>
					</exchangeItem>
					<exchangeItem id="beta">
						<modelExchangeItem id="lippe1.beta"/>
					</exchangeItem>
					<exchangeItem id="alfa">
						<modelExchangeItem id="lippe1.alfa"/>
					</exchangeItem>
					<exchangeItem id="lp">
						<modelExchangeItem id="lippe1.lp"/>
					</exchangeItem>
					<exchangeItem id="khq">
						<modelExchangeItem id="lippe1.khq"/>
					</exchangeItem>
					<exchangeItem id="perc">
						<modelExchangeItem id="lippe1.perc"/>
					</exchangeItem>
				</exchangeItems>
			</uncertaintyModule>
			<uncertaintyModule
className="org.openda.uncertainties.UncertaintyEngine" workingDirectory=".">
				<configFile>uncertaintySpecification.xml</configFile>
				<exchangeItems>
```

```
                        <exchangeItem id="lippe2.fc">
                            <modelExchangeItem id="lippe2.fc"/>
                        </exchangeItem>
                        <exchangeItem id="lippe2.beta">
                            <modelExchangeItem id="lippe2.beta"/>
                        </exchangeItem>
                        <exchangeItem id="lippe2.alfa">
                            <modelExchangeItem id="lippe2.alfa"/>
                        </exchangeItem>
                        <exchangeItem id="lippe2.lp">
                            <modelExchangeItem id="lippe2.lp"/>
                        </exchangeItem>
                        <exchangeItem id="lippe2.khq">
                            <modelExchangeItem id="lippe2.khq"/>
                        </exchangeItem>
                        <exchangeItem id="lippe2.perc">
                            <modelExchangeItem id="lippe2.perc"/>
                        </exchangeItem>
                    </exchangeItems>
                </uncertaintyModule>
            </parameters>
            <predictor>
                <vector id="H-RN-BFG022.Q.m" sourceVectorId="I-RN-0076.Q.uh"/>
            </predictor>
    </vectorSpecification>
</blackBoxStochModel>
```

hbvWrapper.xml

This file determines how each model instance should be run. It has similarities to the generalAdapter configurations of Delft-FEWS. It can run a number of commands (‘actions’) with arguments from each work-folder. An example is given below.

```
<?xml version="1.0" encoding="UTF-8"?>

<blackBoxWrapperConfig xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
    xmlns="http://www.openda.org"
        xsi:schemaLocation="http://www.openda.org
http://schemas.openda.org/blackBoxWrapperConfig.xsd">

    <aliasDefinitions defaultKeyPrefix="%" defaultKeySuffix="%" >
        <alias key="templateDir"/>
        <alias key="instanceDir"/>
        <alias key="modelRootDir"/>
        <alias key="piTimeseriesResultFile"/>
    </aliasDefinitions>
    <run>
        <!-- for each model instance, the template directory will be cloned to create the
instance directory -->
        <initializeActionsUsingDirClone instanceDir="%instanceDir%%instanceNumber%"
templateDir="%templateDir%"/>

        <!-- exe's / classes to be run for each computations -->
        <computeActions>
            <action exe="../../../../bin/move_files.bat"
workingDirectory="%instanceDir%%instanceNumber%">
                <arg>..</arg>
```

```
					<arg>.\Fews\update</arg>
					<arg>FewsPTQ.xml</arg>
				</action>
				<action exe="../../../../bin/runHbvModel.bat"
workingDirectory="%instanceDir%%instanceNumber%">
					<arg>Rhein/</arg>
				</action>
				<action exe="../../../../bin/runFewsRes.bat"
workingDirectory="%instanceDir%%instanceNumber%">
					<arg>Rhein/</arg>
				</action>
				<action exe="../../../../bin/runHbvPostModelAdapter.bat"
workingDirectory="%instanceDir%%instanceNumber%">
					<arg>Rhein</arg>
				</action>
				<action exe="../../../../bin/move_files.bat"
workingDirectory="%instanceDir%%instanceNumber%">
					<arg>.\Fews\update</arg>
					<arg>..</arg>
					<arg>FewsPTQ.xml</arg>
				</action>
			</computeActions>
			<finalizeActions/>
	</run>

	<inputOutput>
			<ioObject
className="nl.deltares.openda.models.hbv.HbvParamsReaderWriter">
				<file>%modelRootDir%</file>
				<id>HbvParameters</id>
			</ioObject>

			<ioObject className="nl.deltares.openda.fews.io.PiTimeSeriesIoObject">
				<file>%piTimeseriesResultFile%</file>
				<id>HbvPiTimeSeriesResults</id>
			</ioObject>

	</inputOutput>

</blackBoxWrapperConfig>
```

B.1.3.3 *uncertaintySpecification.xml*

This file establishes uniform distributions for each parameter of interest. An example is given below.

```
<?xml version="1.0" encoding="UTF-8"?>
<uncertainties xsi:schemaLocation="http://www.wldelft.nl
http://datools.wldelft.nl/schemas/v1.3/uncertainties.xsd" version="1.0"
xmlns="http://www.wldelft.nl" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
	<uncertaintyType>ProbabilityDistributionFunction</uncertaintyType>
	<probabilityDistributionFunction id="lippe2.fc" isActive="true">
		<uniform min="-200" max="200" stdvIsFactor="false"/>
	</probabilityDistributionFunction>
	<probabilityDistributionFunction id="lippe2.lp" isActive="true">
		<uniform min="-0.4" max="0.25" stdvIsFactor="false"/>
```

```
        </probabilityDistributionFunction>
        <probabilityDistributionFunction id="lippe2.beta" isActive="true">
            <uniform min="-1.0" max="1.0" stdvIsFactor="false"/>
        </probabilityDistributionFunction>
        <probabilityDistributionFunction id="lippe2.alfa" isActive="true">
            <uniform min="-0.5" max="0.5" stdvIsFactor="false"/>
        </probabilityDistributionFunction>
        <probabilityDistributionFunction id="lippe2.khq" isActive="true">
            <uniform min="0.015" max="0.165" stdvIsFactor="false"/>
        </probabilityDistributionFunction>
        <probabilityDistributionFunction id="lippe2.perc" isActive="true">
            <uniform min="0.5" max="3.5" stdvIsFactor="false"/>
        </probabilityDistributionFunction>
</uncertainties>
```

B.1.4 StochObserver

This folder contains configuration files, defining which time series are used for calibration. The folder also contains the PI-timeseries file, containing the actual time series themselves.

To retain missing value entries for further analysis (e.g. this is required for extreme value analysis, because one needs a complete time series), one should include a flag: removeMissingValues and set this to ‘false’. An example stochObserver config file is given below.

```xml
<?xml version="1.0" encoding="UTF-8"?>
<ioObjectStochObserver xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xmlns="http://www.openda.org" xsi:schemaLocation="http://www.openda.org
http://schemas.openda.org/openDaStochObserver.xsd">
	<!-- Specification of the uncertainty for the observations -->
	<uncertaintyModule workingDirectory="."
className="org.openda.uncertainties.UncertaintyEngine">
		<arg>stochObsUncertainties_Moselle1.xml</arg>
	</uncertaintyModule>
	<!-- IoObject that reads the observations from a PI time series file -->
	<ioObject workingDirectory="."
className="nl.deltares.openda.fews.io.PiTimeSeriesIoObject">
		<!-- HBV result file containing the "truth", i.e. the results of the original run -->
		<fileName>FewsRES.xml</fileName>
	</ioObject>
	<removeMissingValues>false</removeMissingValues>
</ioObjectStochObserver>
```

B.2 Preparing a HBV subcatchment model.

The HBV model for the complete Rhine basin consists of 134 HBV units. This is a quite heavy model and it would be a waste of computational efforts to run the complete HBV model while calibrating only a very small portion of it. Therefore, the first step is to make a submodel from the complete model. Table .. lists the actions which should be undertaken to make a model structure for a subcatchment. Everywhere where <submodel> is given, the folder name for the submodel should be replaced.

No.	Action
1	Copy the folder structure of one of the existing submodels (e.g. Lippe)
2	Replace the folders in <submodel>/stochModel/HBVModel/original by the complete model (stored in ./stochModel/HBVModel/original)
3.	In the folder with the HBV configuration (called “Rhein”) edit the following files, so that only the HBV units of interest are referred to: - FewsPTQ.key (make sure both the ‘p’ and ‘t’ entries are given. The number of lines should be equal to 2 * nr. Of HBV units) - FewsRES.key (contains results in discharge, so number of lines is equal to nr. Of HBV units) - Ptqw.key (only show the ‘p’ and ‘t’ entries of the HBV units of interest - Basin.par. For each HBV unit, 5 lines are given. Include these lines -
4.	Ensure that the parameter values are all set to default (e.g. copy bmod.par from existing config). Ensure that MAXBAS is given the same parameter value as the

	original model.
5.	Adapt HQ and K4 based on the analyses given in Section 2.2 of this report.
6.	Prepare input files from FEWS as outlined in the section below and overwrite FewsPTQ.dat, dos_ptqw.dat, FewsPTQ.par and FewsRES.par.

B.3 Preparing submodel input series

To ensure the submodel receives the correct inputs, new input files for HBV (so-called PTQ files) should be generated. FEWS can do this with its HBV pre-adapter. To make these files, make a new locationSet, which includes all HBV units of interest (no more, no less!!). Then open the module that runs HBV Rhine (HBV_Rhine.xml) and change the locationSetId in the generalAdapter export section to the newly prepared locationSetId. Run the workflow over the period 01-01-1985..31-12-2006 (calibration period) to yield the correct period.

When this is done, overwrite the files FewsPTQ.dat, dos_ptqw.dat, FewsPTQ.par and FewsRES.par from the FEWS modules folder, which holds the HBV configuration. FewsPTQ.dat and dos_ptqw.dat are identical and can be written from the same file. Also replace FewsPTQ.xml (in <submodel>HBVModel) by the newly made from FEWS.

B.4 Preparing submodel output calibration series

Each OpenDA run should use one or more observed time series to compare simulated values against. The observed series should be exported very much in the same way as the input series.

First prepare a locationSet, containing all the flow observation stations, used in the stochObserver of the OpenDA setups of the subcatchment of interest. Then adapt the HBV_Rhine.xml in such a way that the flow observations of the locationSet are exported instead of the precipitation and temperature series. The workflow should be run over the period 02-01-1985 / 31-12-2006. (not 01-01-1985 due to the fact that HBV only produces data one day after the first inputs). The resulting FewsPTQ.xml should be copied to the stochObserver folder in the OpenDA setup, and renamed to FewsRES.xml.

B.5 Preparing the OpenDA configuration

The OpenDA configuration consists of a number of different files, tied together by an XML-structure. In each submodel, a number of GLUE runs should be performed in sequence. Each GLUE run has its own XML structure, GLUE run specific files are indicated with <submodel_case>, this should be replaced by a case-specific name, which clarifies in which order the GLUE runs should be performed (e.g. in the case of the Lippe: Lippe1, Lippe 2, Lippe3). The files which should be adapted are described below.

File (all in <submodel>)	Change
./hbvDud_ <submodel_case>.oda	- Check the workingDirectory entry of the stochModelFactory. The other workingDirectory entries should be ok - **(!) Change the filename for results into a different filename. If you do not change this, the results of other GLUE runs (in the example e.g. Lippe1) will be overwritten.** - Change the stochObserver config file to a case-specific config file (e.g. stochObsConfig_Lippe3.xml)
./stochObserver/ stochObsUncertainties_<submodel_case>.xml	Prepare this file by copying from a different config file. Change the id in the probabilityDistributionFunction into the id of the observation time series in FewsRES.xml, which is used for calibration in this submodel-case (ensure FewsRES.xml contains all observation required)
./stochObserver/ stochObsConfig_<submodel_case>.xml	Prepare this file by copying from a different config file. Replace the reference to the stochObsUncertainties file to the correct one for this case.
./stochModel/HBVModel/FewsPTQ.xml	If not already done in the previous section, replace this file for the FewsPTQ.xml coming from the FEWS update (see section above). This file is quite large and is moved into the right place in each modelInstance. This saves some copying time.
./stochModel/HBVModel/<submodel _case>/hbvStochModel.xml	Change the vector id and sourceVectorId to match the simulated and observed variable respectively.