Flanders
State of
the Art

00_162_4-6
FHR reports

Modelling water availability and water allocation strategies in the Scheldt basin

Sub report 4-6
Evaluation of 4 hydrological rainfall-runoff models
under climate change conditions

DEPARTMENT
MOBILITY & PUBLIC WORKS

www.flandershydraulicsresearch.be

Modelling water availability and water allocation strategies in the Scheldt basin

Sub report 4-6 – Evaluation of 4 hydrological rainfall-runoff models under climate change conditions

Vélez, C.; Maroy, E.; Rocabado, I.; Pereira, F.; Nossent, J.; Mostaert, F.

Flanders Hydraulics Research

Flanders State of the Art

Cover figure © The Government of Flanders, Department of Mobility and Public Works, Flanders Hydraulics Research

Legal notice

Flanders Hydraulics Research is of the opinion that the information and positions in this report are substantiated by the available data and knowledge at the time of writing.
The positions taken in this report are those of Flanders Hydraulics Research and do not reflect necessarily the opinion of the Government of Flanders or any of its institutions.
Flanders Hydraulics Research nor any person or company acting on behalf of Flanders Hydraulics Research is responsible for any loss or damage arising from the use of the information in this report.

Copyright and citation

© The Government of Flanders, Department of Mobility and Public Works, Flanders Hydraulics Research 2021
D/2021/3241/137

This publication should be cited as follows:

Vélez, C.; Maroy, E.; Rocabado, I.; Pereira, F.; Nossent, J.; Mostaert, F. (2021). Modelling water availability and water allocation strategies in the Scheldt basin: Sub report 4-6 – Evaluation of 4 hydrological rainfall-runoff models under climate change conditions. Version 2.0. FHR Reports, 00_162_4-6. Flanders Hydraulics Research: Antwerp, Belgium.

Reproduction of and reference to this publication is authorised provided the source is acknowledged correctly.

Document identification

Customer:	Flanders Hydraulics Research	Ref.:	WL2021R00_162_4-6
Keywords (3-5):	Hydrology, Climate Change, Low Flows		
Knowledge domains	Water management > Hydrology> Conceptual Models> Numerical modelling		
Text (p.):	29	Appendices (p.):	63
Confidentiality:	☒ No	☒ Available online	

Author(s):	Vélez C, Maroy E., Rocabado I.

Control

	Name	Signature
Reviser(s):	Nossent, J.;	Getekend door:Jiri Nossent (Signature) Getekend op:2021-07-08 10:24:26 +02:0 Reden:Ik keur dit document goed Jiri Nossent
Project leader:	Pereira, F.	Getekend door:Fernando Pereira (Signat Getekend op:2021-07-02 15:41:28 +02:0 Reden:Ik keur dit document goed Fernando Pereira

Approval

Head of Division:	Mostaert, F.	Getekend door:Frank Mostaert (Signatur Getekend op:2021-07-07 16:23:50 +02:0 Reden:Ik keur dit document goed Frank Mostaert

CERTIFIED ISO 9001

Abstract

In the framework of the water allocation projects there are four hydrological models (NAM, PDM,VHM and WETSPA) calibrated for the region of Flanders. The aim of the models is to generate the flows needed as input in the basin model in other to analyze water allocation problems and generate strategies. It is also expected to use the hydrological models to analyze possible impacts of climate variability in the context of water allocation. To realize that, the models are forced with projected climate variables that are frequently out of the range of the data used in the calibration process. Thus models are force to extrapolate conditions for which the parameterization is or may not be optimal. This may bring errors in the projection of climate effects. The problem assessing the performance of the models is that there is not data available in the future to compare with.

The aim of this document is to present the proposed approach for testing the capacity of the calibrated hydrological models to project the climate change effects in the context of water allocation for Flanders. The proposed testing framework combine the use of historical data through an adapted form of Differential Split Sample test (Refsgaard et al., 2014) with and approach that use the relative change of statistical properties of flows between historical data and the projected data (van Steenbergen and Willems, 2012). To assess the model we use a robustness criteria based on a Log Nash-Sutcliffe, BIAS on cumulative volumes and relative changes based on Q50/Q90 estimated from the duration curve.

Contents

List of tables

List of figures

1 Introduction

1.1 Background

To analyse the impacts of climate changes in the context of water management, hydrological models are forced with projected variations of climate variables (i.e. precipitation, evapotranspiration, etc.). The hydrological models are frequently calibrated and validated on pass or current conditions; but they are used to project the response of the catchment under future conditions that in principle differ from those for which they were calibrated. Thus, the analysis of impacts of climate variability challenge the usability of the modelling tools and the credibility of its outputs for water management applications. The key problem is that modellers cannot validate the projected effects on the catchment responds to climate change as there is not data reflecting the future condition on the river outlet.

Recently a framework for testing the ability of models to project climate change and its impacts was proposed by a group of researchers (Refsgaard, et al., 2014). In the framework, the authors recommend the use of the differential-split sample test (DSST) in order to build confidence in the model projections. Variations of the DSS test can be found in the scientific literature and also applied in practical situations. For instance, the generalized DSST, which increase the calibration-validation test by using a moving window though the sub-periods of data (Coron, et al., 2012). DSST rely on the existing records of climate variables and hydrological responds, hoping that there is enough of variation in the sub-periods to represent the kind of variability of the climate conditions for which the model will be projecting hydrological responds. The criteria to assess the performance of the models are based on indicators of the error between observed and simulated variables.

Other authors suggest that, since climate models are not able to reproduce single events but rather statistical properties describing the climate, this should be reflected in the test of hydrological models. Thus, performance criteria such as the root mean square error (RMSE) should be replaced by for instance flow duration curves or quantiles in probability distribution functions (Nicolle, et al., 2014). Using this type of performance criteria, a method to test the validity of hydrological models in a climate changing context was proposed by van Steenbergen and Willems (2012). The method is based on the evaluation of peak flow increases due to different levels of rainfall increases. Differently from the DSST, this method use the projected climate variability and it is especially useful to compare different modelling tools.

In the framework of a water allocation project for the region of Flanders (Belgium) we calibrated four hydrological models: NAM, PDM,VHM and WETSPA; for 67 gauged sub-catchments with approx. 40 years of records. The main objective with the hydrological models is to generate the inputs for the Mike Basin model used in the water allocation management. In addition, it is expected that the hydrological models can be used to assess the impacts of possible climate change scenarios in the water allocation context. Therefore, it is needed to assess the capability of the calibrated models to project climate change scenarios and represent the effect on the responds of the catchments.

1.2 Objectives

The objective of this document is to present the methodology proposed to evaluate the capacity of hydrological modelling tools to assess the effects of climate change scenarios in the context of water allocation in the Flanders Region of Belgium.

2 Methodology

2.1 Modelling approach

2.1.1 Calibration of hydrological models

The objective is to compare the capabilities of 4 rainfall-runoff modelling structures : NAM, PDM, VHM and WETSPA. Models of these four structures were previously calibrated for 57 gauged catchments using manual and automatic calibration. Reports about the calibration results can be found in the reports:

- NAM
  WL2020R00_162_5_DO-4-1_NAM_Schelde; Sub report 4-1 – Analyses of hydrological models for climate change modelling – NAM.
  WL2020R00_162_6_DO-4-2_NAM_Meuse; Sub report 4-2 – Developing a rainfall-runoff model of the Meuse.

- PDM
  WL2020R00_162_7_DO-4-3_PDM; Sub report 4-3– Analyses of hydrological models for climate change modelling – PDM modelling

- VHM
  WL2020R00_162_8_DO-4-4_VHM; Sub report 4-4 – Analyses of hydrological models for climate change modelling – VHM.

- WETSPA
  WL2020R00_162_9_DO-4-5_WETSPA; Sub report 4-5 – Analyses of hydrological models for climate change modelling – WETSPA

Input evapotranspiration and precipitation data is available for the period 1967-2013. As general rule, the most-recent 13 years of existing data were used for calibration and the remaining years were used for validation. The years with available discharge records vary from one catchment to the other however. Consequently, the calibration period could not be kept identical for all catchments and some adjustments were necessary. In the case of the distributed WETSPA model, given the very long computation times, shorter calibration periods were used. The table in ANNEX 8 gives the general goodness-of-fit statistics resulting from the calibration and the corresponding calibration periods.

The models were calibrated using historical data from 1990 to 2012. Precipitation data is available since 1967 but there is missing data of discharge therefore the splitting is not always homogeneous and the length of calibration/validation periods varies. The most recent 10 years of existing data were used for calibration and the rest was used for validation. Figure 17 shows as an example of observed data of total flow and the model output for the sub-catchment W08SAMRON000 (Dijle/Zennebekken).

A multi-objective optimization algorithm was used to calibrate the models: the NSGAII was linked to the hydrological modelling tools (NAM, PDM,VHM and WETSPA). Details on the approach can be found in Naranjo et al. (2020). In the calibration process two indictors of goodness of fit (IGoF) were used as objective function in the optimisation process. Overall, the methodology for the calibration of hydrological models in the context of water allocations give priority to calibration of base flows and average flows but at the same time preserve the volume balance. The IGoF used were the Log NSE on the time series of simulated and observed flows and Volume Absolute Error on the cumulative flows. Other IGoFs were calculated in each model to complement in the analysis of the performance of the models in the calibration and validation process. That is on the time series of observed and modelled: the Root Mean Square Error (RMSE), the NSE, the NSE relative, the Kling-Gupta efficiency (KGE); and in the cumulative volume the relative error and the volume bias.

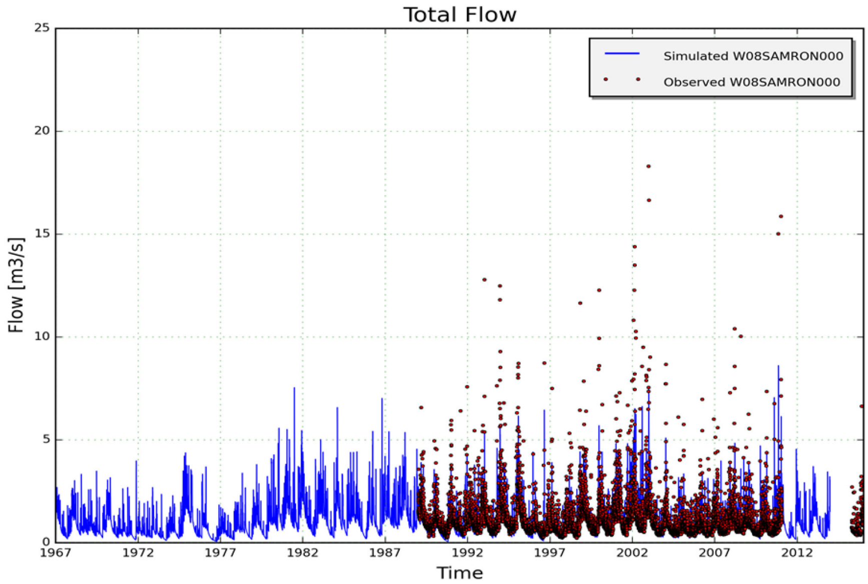


Figure 1 – Illustration of Information of Flows Available for Calibration and Validation in the sub-catchment W08SAMRON000

([a]) source: modelling report from calibration of sub-catchment W08SAMRON000

2.1.2 Perturbation of meteorological timeseries

The climate perturbation tool used for this study was developed at the Katholieke Universiteit Leuven and is a result of the CCI-HYDR project. References for this tool include Ntegeka et al. (2011).

This instrument allows the user to apply a perturbation on time series of different meteorological variables such as rainfall, evapotranspiration or temperature. The logic behind the tool consists in applying a perturbation factor to historical values, depending on the nature of the variable, the location, the season and a chosen projected scenario.

The tool offers the user different scenarios to choose from depending on the objectives of the study. They are based upon four green-house gas emission scenarios (A1B, A2, B1 and B2) developed by the IPCC (2001) and more than 50 runs with global and regional climate models for Belgium.

The results from these runs were statistically analysed and a range of change factors (perturbation factors) from the reference period 1961–1990 till the future period 2071–2100 derived per month and in function of the return period.

The perturbation tool of Ntegeka and Willems (2009) makes it possible to translate observed time series (rainfall, ETo) into future time series for three climate change scenarios:

- High scenario: wet winters and dry summers
- Mean
- Low scenario: dry winters and dry summers

Each scenario, at at different levels of the spectrum of possibilities, will lead to different hydrological conditions however and will be more or less critical for specific planning purposes (e.g. floods, water shortage). All scenarios are possible since each of them is designed as to represent the expected climatological conditions according the state-of-the-art climatic studies in Belgium (Vansteenkiste, 2012). Ideally, all scenarios should be considered to represent variability and uncertainty around future climate scenarios. However, to limit computation and analysis time, a scenario was chosen to reflect the most critical situation in the context of low flows.

The “high winter” scenario has been chosen in this study as the most adequate to study the impact of climate change on low flows in Flanders. This scenario is indeed designed to generate higher precipitation during the winter, but also, more importantly for our purpose, lower precipitation and higher evapotranspiration in the summer, which makes atmospheric conditions the most critical for water shortages.

For a more detailed description of the perturbation tool, refer to 2.4 where the different steps of the perturbation are explained as well as the motivations for choosing the “high winter” scenario for the purposes of this study.

2.2 Integrated approach to assess model projection capabilities

Based on the literature review presented in ANNEX 1, an integrated methodology is proposed to assess the projection capability of the calibrated models in the framework of water allocation. This methodology combines a differential split-sample test, using historical data, with the comparison of the relative change of statistical descriptors between various model projections. The general scheme of the proposed approach is presented in Figure 2. The methodology starts with calibrating the models (which was the object of previous reports), then splits into two branches:

- One based on historical data only, to assess the model robustness when validation forcing is leading to lower flows than calibration forcing, by evaluating change of performance
- the other based on perturbated time series, to compare different model behaviour and variability in projection mode.

Models are evaluated on different criteria and the aim is to generate a hierarchy of modelling structures according to their ability to project the climate change effects and the level of confidence that can be associated to those predictions.

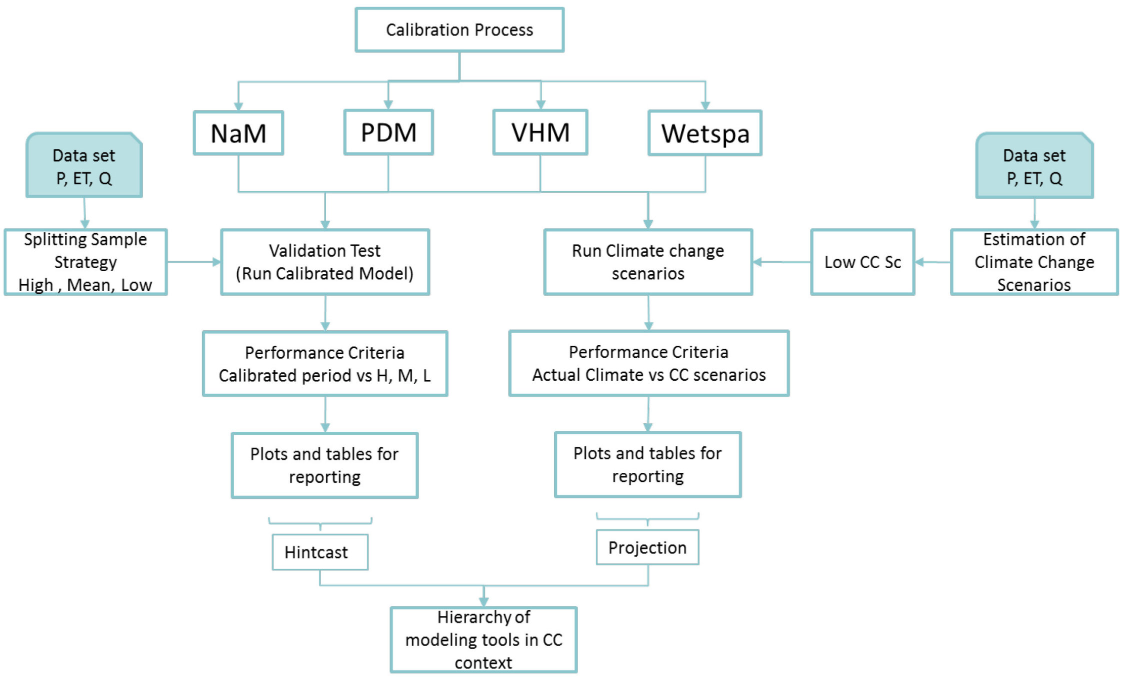


Figure 2 – Illustration of Integrated Approach to Assess Model Projection Capability in Climate Change Context

2.3 Modified differential split-sample test

The methodology applied here is inspired from the differential split-sample test (DSSt) described by Caron et al. (2012). This modified differential splitting method is used to compare the robustness of the different model structures when applied to drier conditions than the calibration period. The methodological note in ANNEX 1 provides more detail on the development and testing of the method.

Generally, the differential split-sample test follows three steps:

1. A small number of sub-periods are selected according to one climate characteristics,
2. Calibration and validation are performed on pairs of periods, then compared both ways,
3. The validation performances are compared to evaluate whether they vary significantly when climatic characteristics differ between calibration and validation.

DSS tests rely on the existing records of climate variables and hydrological response, hoping that there is enough variation in the sub-periods to represent the kind of variability of the climate conditions for which the model will be projecting hydrological response

The modified methodology proposed here used a 3-year window sampling approach, thus minimizing calibration effort and using a maximum of the available data in a way that can be used to isolate significant change in climate variables (here precipitation). The analysis is carried for all four studied model structures (NAM, PDM, VHM and WETSPA) . Ultimately, robustness indicators can be compared for all catchments in order to select the most robust model(s) overall (namely in the various physical and hydrological conditions existing in Flanders).

2.3.1 Sampling approach

Given the large number of catchments to be processed with the four model stuctures described above (2.1.1), it was not feasible to recalibrate each separately for the purpose of the split-sample test. As as simplication, a window of 3 years within the calibration period was chosen as reference (“benchmark”), and then all other windows of both calibration and validation years werecompared to that reference. The figure below (Figure 3) illustrate the choice of the benchmark sample (three years) according to three criteria:

- Part of calibration period,
- Above average annual rainfall,
- Satisfactory conformity between simulated and observed flows and volumes.

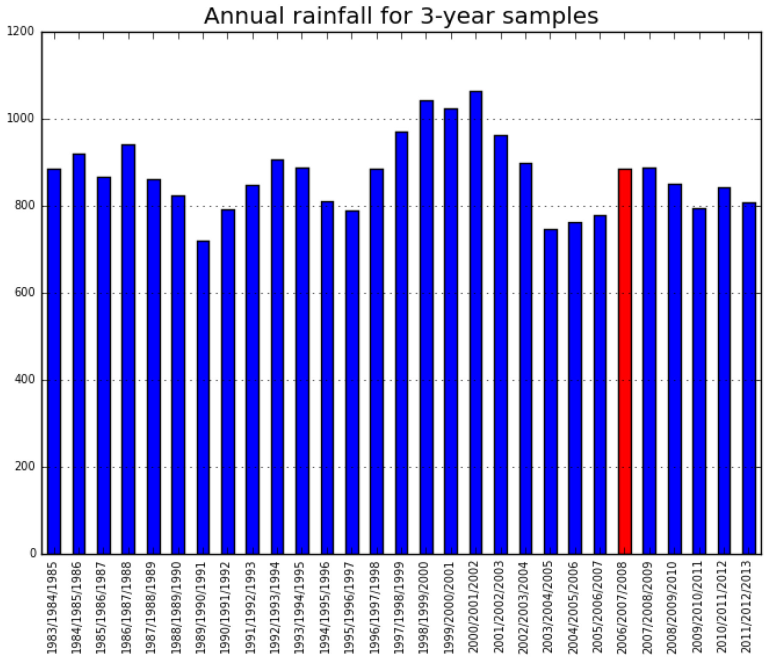


Figure 3 – Selection of benchmark according to annual rainfall (method 2) V09ZWA148120

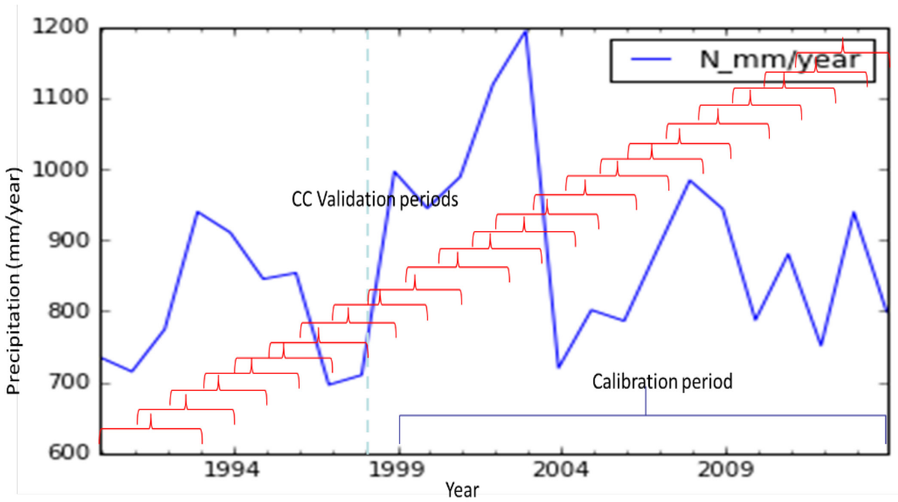


Figure 4 – Illustration of splitting of data for calibration and validation

2.3.2 Performance Criteria

The criteria selected as indicators of the performance of the model are based on the Model Robustness from the GDSSt approach (Coron, et al., 2012) and the Relative Change as in the approach described by van Steenbergen and Willems (2012). The indicators used to calculate the criteria are selected based on the purpose of the model. In principle, a set of 7 criteria were calculated for each sub-period modelled from the historical data. However, priority was given to performance criteria measuring the ability to reproduce mean annual flows or minimum flows and volume balances. The model robustness criteria is expressed below, applicable for either log NSE or bias (ε in the equation of MRC).

1. Model robustness criteria (MRC)

$$\mathrm{MRC}_{D\rightarrow R} = \frac{\varepsilon_{D\rightarrow R}}{\varepsilon_{R\rightarrow R}} - 1$$

εD->R = NSE and BIAS D->R

εR->R = smallest value of ε in calibration period

where ε is the objective function to be minimized during calibration. εD is the error during calibration and εR is during validation

In this case, the model robustness criteria can thus be calculated for logarithmic Nash-Sutcliff efficiency and bias on volumes:

Error criteria
Log Nash-Sutcliffe efficiency (NSLog)

$$NSLog = \frac{\sum_{i=1}^{n}(\log(Qobs,i) - \log(Qobs))^2}{\sum_{i=1}^{n}(\log(Qobs,i) - \log(\overline{Qobs}))^2}$$

Bias on total volumes

$$\mathrm{BIAS}_{D\rightarrow R} = \frac{\sum_{k=1}^{n}\hat{Q}_{R,k}[\theta_D] - \sum_{k=1}^{n}Q_{R,k}}{\sum_{k=1}^{n}Q_{R,k}}$$

2.3.3 Analysis of results

The results can be represented on a dotty plot with, for each sample of each model, the model robustness criteria in function of a climate change indicator. Changes in climate are expressed as ratios (e.g., 10% less rainfall). The plot shows the “Relative Loss of Performance” against the relative evolution of climate conditions. Each dot represent one run of the model. The cloud of dots is for all periods of all catchments and the box plot is a summary of the dots. Figure 1 shows this kind of dotty plot for the NAM models in the Ijzer Basin.

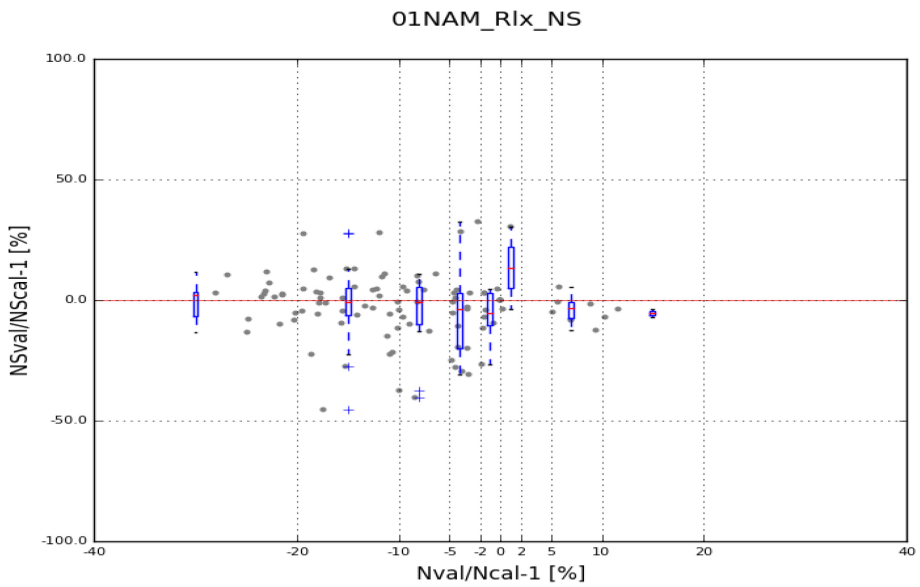


Figure 5 – NSE relative change in function of rainfall relative change, for samples of NAM simulations in catchments of the Ijzer Basin (“01”)

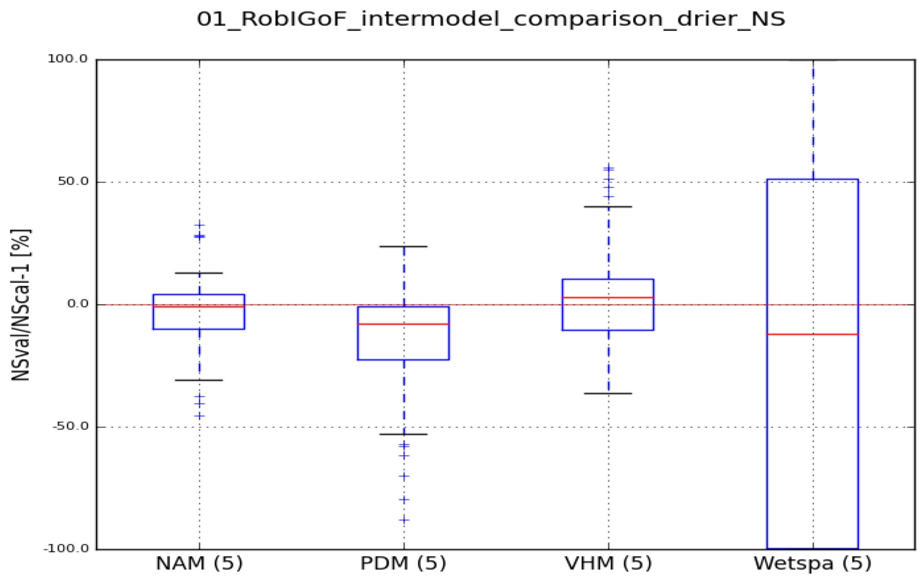


Figure 1: Relative loss of NS performance against relative drier climate conditions (IJzer)

Results are analysed by:

- Plotting the robustness criteria for each sample against relative change in rainfall for that sample.
- Grouping those results in boxplots per category of rainfall relative change
- Comparing boxplots of samples for the 4 types of model: distribution for all samples with rainfall decrease
- Grouping plots per basin.
- Exclusion of models based on robustness criteria outside confidence interval [median – std. dev ; median + std. dev]

2.4 Relative variation on projected climate and hydrological conditions

The methodology is inspired by the approach proposed by Van Steenbergen and Willems (2012). The methodological note is detailed in section 1.3 of ANNEX 1.

To evaluate the projection capability of th four types of models, we want to assess:

- The amplitude of the response to perturbated P and Evap series
- The differences of variability of this response between models

For the purposes of low flow analysis, relative change can be calculated for the median discharge Q50, the 90th dry percentile Q90, and the ratio of Q90/Q50, estimated from the duration curves .

2. Relative Change

$$Rel.Change_{act,cc,T(i)} = \frac{Qr_{cc,\,T(i)} - Qract_{,\,T(i)}}{Qr_{act,\,T(i)}}$$

Qr = ratio Q50/Q90
Qr is estimated from the duration curve and may represent the variability of low-flow discharges

3 Exclusion of the poorest-calibrated catchments

Before performing the model capability assessment regarding climate change for all catchments, we chose to exclude some models because they failed to describe current reality. It was therefore pointless to assess they capability to represent future climate conditions. In particular, catchments were excluded that could not be modelled satisfactorily with more than two or three model structures. Further explanation on the reasons of poor calibration of models can be found in calibration reports. Typically, this can be due to poor quality of the measurements, seasonal disturbances such as vegetation growth, human disturbances to the river discharge etc.

Ultimately, we selected models with sufficient level of confidence, namely for those catchments that have enough quality data to have solid results. Therefore we eliminated the models with poor validation goodness-of-fit statistics (NS, log NS and relative error on cumulative discharge), especially when optimization proved difficult for more than one model structure with, for instance, several possible sets of parameters. Consequently, the catchments listed below were excluded from the analysis (Table 1). Goodness-of-fit coefficients are also indicated in the table.

Table 1 – The excluded catchments and the results of their efficiency coefficients corresponding to the different models

Modelname	NS				NS_log				ReError			
	NAM	PDM	VHM	WETSPA	NAM	PDM	VHM	WETSPA	NAM	PDM	VHM	WETSPA
V01IEP495080	0.391	0.209	0.554	0.484	0.402	0.56	0.45	0.537	0.057	0.097	-0.036	-0.02
V01KEM492060	0.628	0.582	0.575	0.238	0.524	0.519	0.48	0.356	0.011	-0.003	0.02	-0.07
V07MOG288020	0.626	0.331	0.543	0.485	0.02	0.691	0.5	0.515	-0.045	0.03	0.003	-0.06
W07DENLES004	0.587	0.466	0.519	-0.227	0.426	0.666	0.54	0.366	-0.129	0.05	0	0.012
V06MAA347160	0.493	0.407	0.58	-0.453	0.461	0.607		0.509	-0.054	0.007	-0.004	-0.02
W06RHOL54100	0.493	0.445	0.555	0.708	-6.54	0.603	0.38	0.73	-0.168	0.009	-0.009	0.012
V08DIJ093400	-3.412	0.61	0.606	0.329	-0.470	0.618	0.157	0.394	-0.971	0.025	0.017	-0.011
W08SENRON010	0.581	0.567	0.565	0.54	0.45	0.395	0.36	0.454	-0.093	-0.016	-0.066	0.012
V09HUL147150	0.463	0.476	0.404	-0.511	0.474	0.511	0.02	0.47	-0.131	-0.103	-0.015	0.011
V10WIM082050	0.529	0.045	0.624	0.458	0.439	0.586	0.6	0.445	-0.143	-0.003	0.004	0.015
W11HOY5990	0.425	0.291	0.455	0.229	0.372	0.642	0.69	-0.2	-0.031	-0.017	0.004	0.049

4 Validation using the modified differential split-sample test

4.1 Model robustness evaluation

We investigated for each of the catchments, what was the loss of goodness-of-fit from the benchmark sample (selected average rainfall 3-year sample) to drier samples. The statistics taken into account were: NS, the NS_log and the VolBias. In order to select acceptable loss of goodness-of-fit, we tested if the performance criteria is the confidence interval or not, for each model respectively. The confidence interval is defined as:

[median – std. dev; median + std. dev].

For each catchment, we will obtain Table 3. Followed by that, we calculate the number of the criteria (out of the three) that succeeded to be in the confidence interval. The results for all the catchments are explicitly shown in Table 4. We consider that a model is highly robust for a certain catchment if 3 over 3 of the criteria are in the confidence interval. A model is averagely robust if 2 of 3 of the coefficients are in the confidence interval. And a model is poorly robust if less than 2 coefficients are in the confidence interval.

Table 2 – The trust interval for the different variables

Variable	Lower bound	Upper bound
NS	-40.955	29.875
NS_Log	-40.703	21.193
Volume Bias	-182.737	-7.113

Table 3 – Example of the table in which the model indicators are checked if they are in the confidence interval of the robustness evaluation

Catchment	model	NS	NS_log	VolBias
F01IJZ468000	NAM	Accepted	Excluded	Accepted
	PDM	Accepted	Excluded	Accepted
	VHM	Accepted	Excluded	Accepted
	WETSPA	Accepted	Accepted	Excluded

Table 4 – Results of the robustness evaluation for all of the catchments using the differential Split Sample test

Modelname	NAM	PDM	VHM	WETSPA
F01IJZ468000				
V01HAN488180				
V01MAR496120				
V01POP491030				
V01SSV499140				
V02EDE442120				
V02HER426010				
V02KER422030				
V02RIV425020				
V03POE446000				
V04MOL036110				
V04MOM037100				
V07BEL285070				
V07MAR289015				
V07MOE282100				
F05LEI386999				
V05HEU403210				
V05MAN401230				
F06BOS325001				
V06ZWA342190				
V08BAR111370				
V08ZUU233100				
W08SAMRON000				
W08SENTUB030				
V09DEM136000				
V09GET152080				
V09HER163010				
V09MAN161040				
V09MOT144270				
V09VEL145100				
V09WIN141310				
V09ZWA148120				
V10GNE076999				
V10KNE052000				
V10MOP062140				
F11MAA8702				
W11BER551010				
W11MAAPROF				
W11MEH5820				
W11OUR5805				
W11SAM7319				
Total Good	6	11	12	6
Total Average	16	17	13	18
Total Poor	19	13	16	17

For more results, see figures in ANNEX 3 and ANNEX 4, with box-plots per basin and per percentage of rainfall increase respectively.

Table 5 – Results of the robustness evaluation for all of the basins using the differential split sample test

Basin	NAM	PDM	VHM	WETSPA
Ijzer	3	3	3	3
Brugse Polders + Gentse kanalen	2	2	3	3
Benedenschelde + Denderbekken	3	3	2	3
Leie + Bovenschelde	2	2	2	3
Dijle and Zenne	2	3	3	2
Demerbekken	3	3	3	2
Netebekken	2	2	2	3
Maasbekken	3	3	3	1

Table 6 – Results of the robustness evaluation for all of the models using the differential split sample test

Model	Som van NS	Som van NS_log	Som van VolBias	Som	Percentage
NAM	23	22	10	55	11.2%
PDM	22	24	13	59	12.0%
VHM	23	25	13	61	12.4%
WETSPA	17	21	22	60	12.2%
Total	85	92	58	235	47.8%

4.2 General conclusions on model robustness

VHM and NAM appear as the most robust overall. WETSPA seems more volatile.

However, the calibration periods and the benchmark 3-year windows are the same for the lumped models, but not always for WETSPA where a different calibration period was sometimes selected. This makes the results of WETSPA slightly more difficult to compare with the other model structures.

5 Validation using the relative variation on projected climate and hydrological conditions

5.1 The visualization of the results

The first step in order to choose for a model, is to have a look of the outputs of the models and the different indicators related to these latter. We decorticated the data in sets of one year, 6 months of the summer and 6 months of the winter. The interest of the summer period was elaborated to point out the response of the models in low flows.

In order to meet these goals, it is mandatory to display the distribution and the variation of the perturbed and initial rainfall and evapotranspiration series since they represent the input of the models. Followed by that, we will outline the distribution of the cumulative discharge of the different models whether for the perturbed or the current data. Also, we will look out how the distribution of the relative change of the discharge versus the relative change of rainfall is for each of the models. Figure 6 is an example for the catchment V05MAN401230**.**

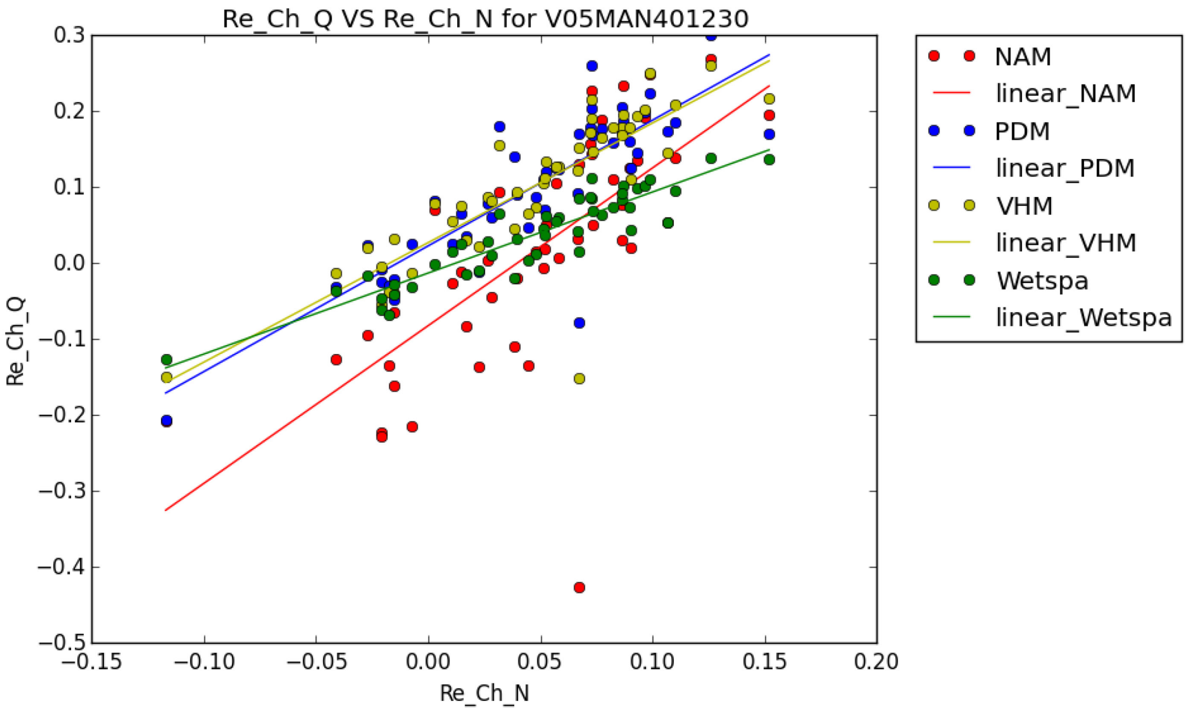


Figure 6 – The Relative change of discharge Versus the Relative Change of rainfall for the catchment V05MAN401230

In addition, the evaluation of the low flow indicators is primordial for the decision making, that’s why different graphs of the distribution of Q_{90} , Q_{50} and Q_{90} /Q_{50} was made to see which of the model simulate the best the low flows.

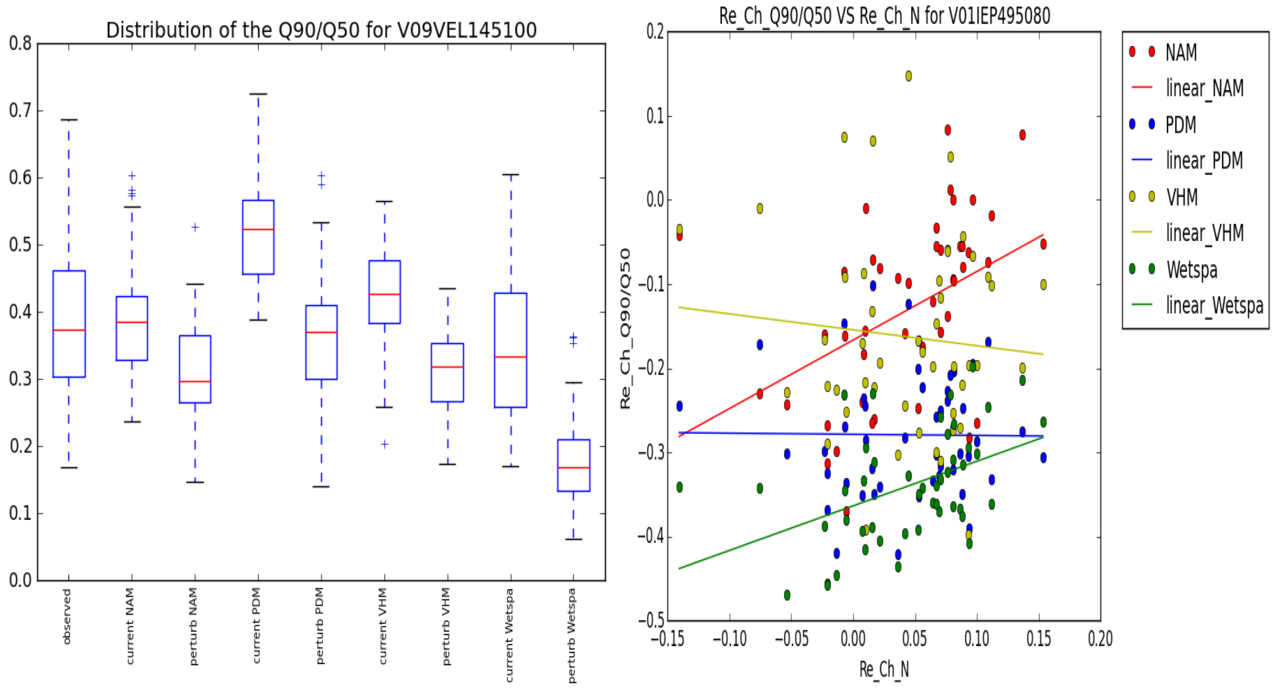


Figure 7 – Left: The distribution of the ratio Q_{90}/Q_{50} for the catchment V09VEL145100.
Right: The Relative change of the ratio Q_{90}/Q_{50} Versus the Relative Change of rainfall for the catchment W08SENTUB030

As mentioned in the previous sections, to choose the adequate model that shows the best robustness, the relative change of the discharge or the different quantiles whether for whole year or only for summer. Their distribution will give us how the models reacts in the same conditions and in which range it occurs. Figure 8 is an example of the distribution of the different variables for the catchment V04MOL036110.

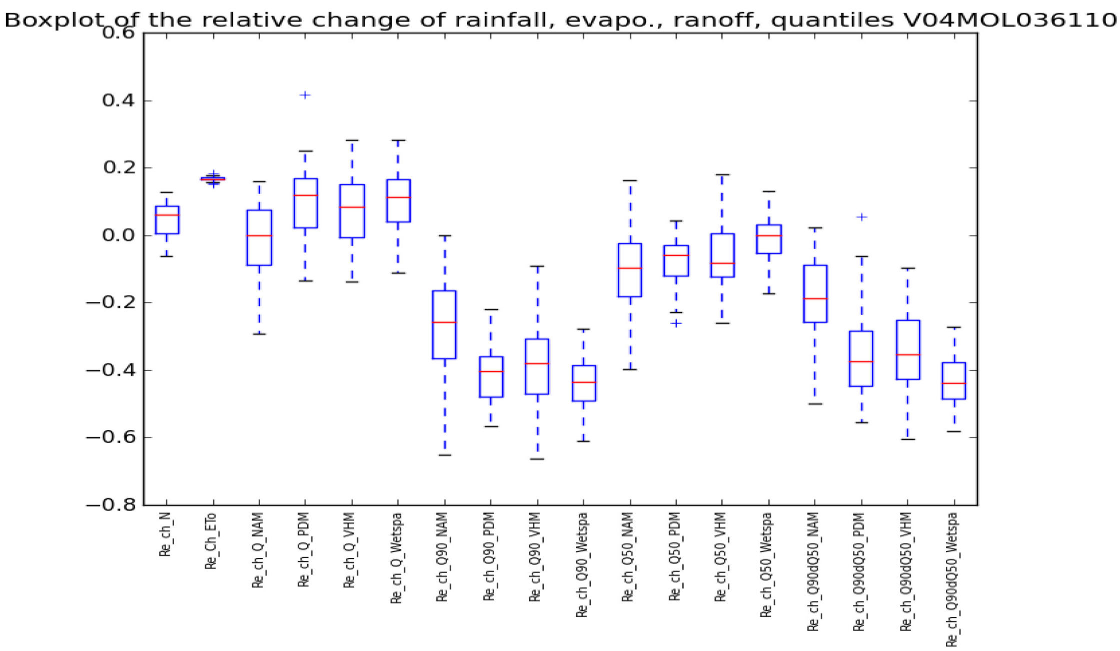


Figure 8 – Boxplot of the relative change of rainfall, evapotranspiration, discharge
and the quantiles of the catchment V04MOL036110

5.2 The evaluation indicators

To evaluate the stability of the models and to eventually choose the robust one to validate all the catchments, we based our assessment on six indicators which are:

- The relative change of the annual cumulative discharge.
- The relative change of the cumulative discharge in summer.
- The relative change of the annual discharge divided by the relative change of the annual rainfall volumes.
- The relative change of the ratio Q90/Q50.
- The relative change of the percentile Q90.
- The relative change of the percentile Q50.

Furthermore, we define the relative change of a variable X by:

$$\frac{X_{perturbed} - X_{initial}}{X_{initial}}$$

The choice of these indicators has been made based on the fact that it does not only consider the discharge volumes variation depending on the rainfall variation, but also it encounters the low flow factors which are the percentile Q_{90}, Q_{50} and the ratio $Q_{90/}$ Q_{50}. Moreover, since we are interested in seeking the model response to the low flows, we took in consideration the relative change of the discharge in summer.

Afterward, for each of the catchments and for each of the models(i.e. NAM, PDM, VHM, WETSPA), we calculate these indicators. Since the validation period was set to 47 years, we ended up with a set of 47 values for each of the variables, for each of the catchments and for each of the models.

Due to considerable amount of data and for better visualization of the results, we calculated the mean, the median and the standard deviation of all of these variables for each of the models and each of the catchments. The results for each of the catchments are in ANNEX 6.

Since the median informs us the most about the range of the variables, we considered it for the further presentations of the results. In fact, the table in ANNEX 7 is the summary of all the mean values of the different variables for each of the catchments.

5.3 The comparison between the different models

To see which of the models is responding inappropriately, we fixed a trust interval that all the models' indicators have to be included in. We calculate the boundaries of the range for the different indicators by calculating their mean value and their standard deviation considering all the models. Therefore, the lower and upper bounds are defined as:

$$Bound_{low} = Mediane - Standard\ Deviation$$

$$Bound_{upper} = Mediane + Standard\ Deviation$$

The obtained boundaries for each of the indicators are as followed:

Table 7 – The boundries table for the different indicators of the models response

Variable	Lower bound	Upper bound
Re.Ch Annual Cumulative Q	-0.00799	0.173919
Re.Ch Annual Summer Cumulative Q	-0.23514	-0.08647
Re.Ch(Vol Q)/ Re.Ch(Vol Rainfall)	0.649159	2.516401
Re.Ch (Q90/Q50)	-0.42965	-0.16773
Re.Ch Q90	-0.46122	-0.20123
Re.Ch Q50	-0.10823	0.060265

Based on whether these indicators are in the appropriate range or not, we evaluated the number of the indicators among the six that fulfill this condition. They were distributed in three types as explained in the following table:

Table 8 – The different types of catchments labeled by their indicators evaluation

Label	Definition	Color Label
Good	**5 to 6 indicators are in the trust interval**	
Medium	**3 to 4 indicators are in the trust interval**	
Poor	**0 to 2 indicators are in the trust interval**	

This procedure was applied to all of the catchments and for each of the models. Hereby, we obtained the following table where the evaluation of the catchments has been made.

Table 9 – The summary table of the evaluation of all the catchments based on their indicators

Modelname	NAM	PDM	VHM	WETSPA
F01IJZ468000				
V01HAN488180				
V01MAR496120				
V01POP491030				
V01SSV499140				
V02EDE442120				
V02HER426010				
V02KER422030				
V02RIV425020				
V03POE446000				
V04MOL036110				
V04MOM037100				
V07BEL285070				
V07MAR289015				
V07MOE282100				
F05LEI386999				
V05HEU403210				
V05MAN401230				
F06BOS325001				
V06ZWA342190				
V08BAR111370				
V08ZUU233100				
W08SAMRON000				
W08SENTUB030				
V09DEM136000				
V09GET152080				
V09HER163010				
V09MAN161040				
V09MOT144270				
V09VEL145100				
V09WIN141310				
V09ZWA148120				
V10GNE076999				
V10KNE052000				
V10MOP062140				
F11MAA8702				
W11BER551010				
W11MAAPROF				
W11MEH5820				
W11OUR5805				
W11SAM7319				
Total Good	14	22	25	23
Total Medium	17	15	10	12
Total Poor	10	4	6	6

Statistically, the VHM model succeeded to enclose the highest number of catchment that showed good indicators with a percentage of 60.97% of the catchments. However, the number of poorly modelled catchments for this model is still considerable with a percentage of 14.63%.

On the opposite side, NAM is the model that revealed the worst statistics with only 34.14% of the catchments that are considered having good indicators.

Furthermore, there are several catchments that have average or bad indicators for all the models (for example W11MAAPROF, F11MAA8702, V09GET15208. This can be due to the bad quality of the input that leads to unstable results or due to the meteorological data that are too high or fluctuant.

To be able to compare the results to the ones of the differential split sample test, we considered another methodology where we group the catchments in basin or a group of basin depending on the number of catchments in each basin and their positions in the map.

After that, we calculate the number of indicators for each basin that verify the condition to be in the trust interval we already specified. Afterwards, we calculate their percentage comparing to the total number of indicators of each basin. The results are explicitly shown in Table 10.

Table 10 – The table of percentage of the number of indicators that are in the trust interval per basin or group of basin

Basin	NAM	PDM	VHM	WETSPA
Ijzer	56.67%	80.00%	73.33%	73.33%
Brugse Polders + Gentse kanalen	66.67%	70.00%	90.00%	66.67%
Benedenschelde + Denderbekken	66.67%	90.00%	80.00%	90.00%
Leie + Bovenschelde	53.33%	73.33%	63.33%	50.00%
Dijle and Zenne	90.00%	76.67%	66.67%	96.67%
Demerbekken	56.25%	66.67%	87.50%	68.75%
Netebekken	55.56%	72.22%	72.22%	83.33%
Maasbekken	66.67%	72.22%	69.44%	69.44%
>=80%				
>=60% and <80%				
<60%				

Unlike the comparison per catchment, PDM and WETSPA have the highest number of acceptable basins (Table 10). When it comes to NAM, this latter wasn't able to model a basin that encounters good statistics. However, if we sum the percentages per model, We will find that VHM is again on the lead; it means that overall this model is scoring higher percentages in the averagely and poorly modelled basins.

In Table 11, for each of models, the calculation of the number of indicators in the trust interval has been done added to their percentage. VHM scored the highest percentage of the accepted indicators with 19.05% of the total indicators. However, NAM has the worst percentage with 15.97% of indicators that are in the trust interval.

Table 11 – Table of the sum of the indicators that are in the trust interval depending of the model

Sum of accepted indicators for :	Re.Ch Annual Cumulative Q	Re.Ch Annual Summer Cumulative Q	Re.Ch(Vol Q)/ Re.Ch(Vol Rainfall)	Re.Ch (Q90/Q50)	Re.Ch Q90	Re.Ch Q50	Sum accepted indicators	% Accepted Indicators
NAM	25	27	27	29	28	25	161	15.97%
PDM	30	28	31	31	34	34	188	18.65%
VHM	38	30	37	26	27	34	192	19.05%
WETSPA	34	34	33	23	25	37	186	18.45%
Total	**127**	**119**	**128**	**109**	**114**	**130**	**727**	**72.12%**

6 Validation using the evalution of the model efficiency coefficients

Besides the projection of the future scenarios, it is interesting to see the behaviour of the models for the current series. For that we choose the same validation period for all of the models and catchment from 01/01/1967 to 31/12/2013 and we evaluate the efficiency coefficients as we did in the reports of the models. However, the calibration period differs from model to another and from catchments to another catchment.

To be coherent with our previous studies, we choose Nash-Sutcliffe, the log of the Nash-Sutcliff and the relative error of the model for our evaluation. In ANNEX 8, is the explicit table of the values of the efficiency coefficients for each of the catchments. To have an overall view at the performance of the models, we calculate the percentage of the coefficients that succeded to be in the trust interval. As mentioned in the previous reports, the trust intervals are defined as followed:

Table 12 – The trust intervals for the model efficiency coefficients

	When value	**Classification**
NSE	>0,6	Good
	0,3-0,6	Average
	<0,3	Poor
LogNSE	>0,6	Good
	0,3-0,6	Average
	<0,3	Poor
RelErr	<15	Good
	15-30	Average
	>30	Poor

In Table 13, there is the summary results for the different basins. It is clearly that PDM has the best performance among the other models, since most of the catchments showed an overall good performance. This is considered as a good indicator for the robustness and the stability of the model. On the opposite side, WETSPA and NAM has proven to have the worst performance overall. For example for WETSPA, for both of NS and NS_log, less than 60% of the catchments for all the basins have coefficients that are in the trust interval.

Table 13 – The summary efficiency coefficient table for the different basins

	NS				NS_Log				Relative Error			
	NAM	PDM	VHM	WETSPA	NAM	PDM	VHM	WETSPA	NAM	PDM	VHM	WETSPA
Ijzer	100%	80%	80%	20%	60%	100%	80%	20%	80%	100%	100%	100%
Brugse Polders + Gentse kanalen	80%	80%	100%	20%	80%	100%	80%	40%	100%	100%	100%	100%
Benedenschelde + Denderbekken	40%	40%	60%	20%	100%	100%	80%	60%	100%	100%	100%	100%
Leie + Bovenschelde	100%	80%	80%	0%	80%	100%	100%	80%	80%	100%	100%	100%
Dijle and Zenne	100%	50%	100%	50%	50%	100%	75%	50%	100%	100%	100%	100%
Demerbekken	75%	62.5%	87.50%	0%	37.5%	50%	37.5%	25%	75%	100%	87.5%	100%
Netebekken	66.67%	66.67%	100%	0%	66.67%	100%	100%	33.33%	100%	100%	100%	100%
Maasbekken	83.33%	100%	100%	16.67%	83.33%	100%	83.33%	50%	100%	100%	100%	100%
>=80%												
>=60% and <80%												
<60%												

Whether for NS_Log, NS or the Relative error, PDM showed great statistics in its efficiency coefficients especially for NS_Log where most of the catchments have high coefficients. In opposition, WETSPA showed really bad results for most of the coefficients that can be explained by the type of the model or its calibration.

This may lead us to the conclusion of not trusting the quality of the comparison results for this particular model. This leave us with two situations; whereas to not consider completely the model as a candidate for the further simulations or to keep it as a candidate but with higher reservations than the other models.

7 Evaluation of the projection capabilities of the models

7.1 Ijzer

46810102 - Ijzer; Roesbrugge Haringe	**F01IJZ468000 (V01HEI468010 and F01YSE468000)**
48810102 - Handzamevaart; Kortemark	**V01HAN488180**
49510102 - Ieperlee; Zuidschote	**V01IEP495080**
49270102 -Kemmelbeek; Boezinge	**V01KEM492060**
49610102 - St. Jansbeek; Merkem	**V01MAR496120**
49110102-Poperingevaart; Oostvleteren	**V01POP491030**
49910102 - Steenbeek; Merkem	**V01SSV499140**

Table 14 – The evaluation tables for the Ijzer Basin

	NAM		**PDM**		**VHM**		**WETSPA**		**Concl<**
	RelChange	**Robustness**	**RelChange**	**Robustness**	**RelChange**	**Robustness**	**RelChange**	**Robustness**	
F01IJZ468000									
V01HAN488180									
V01MAR496120									**PDM**
V01POP491030									
V01SSV499140									

	NAM		**PDM**		**VHM**		**WETSPA**		**Concl<**
	RelChange	**Robustness**	**RelChange**	**Robustness**	**RelChange**	**Robustness**	**RelChange**	**Robustness**	
Ijzer									**PDM**

Apart from NAM, the three other models showed good statistics. One of the reasons is that the catchments seems to be well calibrated and have reliable data. When we check the graphs, PDM seems to be the more stable since most of the boxplots are in a the range of 0 for the NS and the NS_Log. However, for the methodology of the variation of the on projected climate and hydrological conditions,

7.2 Brugse Polders + Gentse kanalen

44210102 - Maldegem	V02EDE442120
42610102 - Hertsbergebeek; Oostkamp	V02HER426010
4220102 - Kerkebeek, Sint-Michiels	V02KER422030
42510102- Rivierbeek; Oostkamp	V02RIV425020
44656122 - Poekebeek; Nevele	V03POE446000

Table 15 – The evaluation tables for the Brugse Polders + Gentse kanalen Basins

	NAM		PDM		VHM		WETSPA		Concl<
	RelChange	Robustness	RelChange	Robustness	RelChange	Robustness	RelChange	Robustness	
V02EDE442120									
V02HER426010									
V02KER422030									PDM VHM
V02RIV425020									
V03POE446000									

	NAM		PDM		VHM		WETSPA		Concl<
	RelChange	Robustness	RelChange	Robustness	RelChange	Robustness	RelChange	Robustness	
Brugse Polders + Gentse kanalen									VHM

7.3 Benedenschelde + Denderbekken

3610102 - Kleine Molenbeek, Liezele	V04MOL036110
3710102 - Grote Molenbeek, Malderen	V04MOM037100
28510102 - Bellebeek, Essene	V07BEL285070
28970102 - Mark, Viane	V07MAR289015
28210102 - Molenbeek, Erpe Mere	V07MOE282100
28810102 - Molenbeek, Geraardsbergen	V07MOG288020
27081002 – Dender, Lessines	W07DENLES004

Table 16 – The evaluation tables for the Benedenschelde + Denderbekken Basins

	NAM		PDM		VHM		WETSPA		Concl<
	RelChange	Robustness	RelChange	Robustness	RelChange	Robustness	RelChange	Robustness	
V04MOL036110									PDM VHM
V04MOM037100									
V07BEL285070									
V07MAR289015									
V07MOE282100									

	NAM		PDM		VHM		WETSPA		Concl<
	RelChange	Robustness	RelChange	Robustness	RelChange	Robustness	RelChange	Robustness	
Beneden Schelde -Denderbekken									WETSPA

7.4 Leie + Bovenschelde

38680122, Leie te Menen	F05LEI386001
40310102 - Heulebeek; Heule	V05HEU403210
40110102 - Mandel; Oostrozebeke (L05_409)	V05MAN401230
32580122 - Bovenschelde; Bossuit	F06BOS325001
34710102 - Maarkebeek; Etikhove	V06MAA347160
34210102 - Zwalm; Nederzwalm	V06ZWA342190
L5412 Amougies – Rhosnes	W06RHOL54100

Table 17 – The evaluation tables for the Leie + Bovenschelde Basins

	NAM		PDM		VHM		WETSPA		Concl<
	RelChange	Robustness	RelChange	Robustness	RelChange	Robustness	RelChange	Robustness	
F05LEI386001									WETSPA
V05HEU403210									
V05MAN401230									
F06BOS325001									
V06ZWA342190									

	NAM		PDM		VHM		WETSPA		Concl<
	RelChange	Robustness	RelChange	Robustness	RelChange	Robustness	RelChange	Robustness	
Leie + Bovenschelde									PDM

7.5 Dijle and Zenne

11110102-Barebeek, Hofstade (Elewijt)	**V08BAR111370**
9310102 - Dijle, Wilsele	**V08DIJ093400**
23310102 - Zuunbeek, St Pietersleeuw	**V08ZUU233100**
2371-10050 Samme, Ronquieres	**W08SAMRON000**
	W08SENRON010
	W08SENTUB030

Table 18 – The evaluation tables for the Dijle and Zenne Basin

	NAM		PDM		VHM		WETSPA		Concl<
	RelChange	Robustness	RelChange	Robustness	RelChange	Robustness	RelChange	Robustness	
V08BAR111370									VHM
V08ZUU233100									
W08SAMRON000									
W08SENTUB030									

	NAM		PDM		VHM		WETSPA		Concl<
	RelChange	Robustness	RelChange	Robustness	RelChange	Robustness	RelChange	Robustness	
Dijle and Zenne									WETSPA

7.6 Demerbekken

13610102 - Demer; Hasselt	**V09DEM136000**
15210102 - Gete; Halen	**V09GET152080**
16310102 - Herk, Kermt (Spalbeek)	**V09HER163010**
14710102 - Zwart Water (affluent of De Hulpe); Molenstede	**V09HUL147150**
14310102 - Grote Losting; Wezemaal	**V09LOS143300**
16110102 - Mangelbeek; Lummen	**V09MAN161040**
14410102 - Motte; Rillaar	**V09MOT144270**
14510102 - Velp; Ransberg	**V09VEL145100**
141 - Rotselaar ; Winge	**V09WIN141310**
14810102 - Zwarte Beek; Lummen	**V09ZWA148120**

Table 19 – The evaluation tables for the Demerbekken Basin

	NAM		PDM		VHM		WETSPA		Concl<
	RelChange	Robustness	RelChange	Robustness	RelChange	Robustness	RelChange	Robustness	
V09DEM136000									VHM
V09GET152080									
V09HER163010									
V09MAN161040									
V09MOT144270									
V09VEL145100									
V09WIN141310									
V09ZWA148120									

	NAM		PDM		VHM		WETSPA		Concl<
	RelChange	Robustness	RelChange	Robustness	RelChange	Robustness	RelChange	Robustness	
Demerbekken									PDM VHM WETSPA

7.7 Netebekken

8610102 - Grote Laak, Vorst	**V10GLA086020**
7610102 Grote Nete/Geel Zammel	**V10GNE076999**
5210102 - Kleine Nete; Grobbendonk	**V10KNE052000**
6210102 - Molenbeek, Pulle	**V10MOP062140**
8210102 - Wiekevorst	**V10WIM082050**

Table 20 – The evaluation tables for the Netebekken Basin

	NAM		PDM		VHM		WETSPA		Concl<
	RelChange	Robustness	RelChange	Robustness	RelChange	Robustness	RelChange	Robustness	
V10GNE076999									WETSPA
V10KNE052000									
V10MOP062140									

	NAM		PDM		VHM		WETSPA		Concl<
	RelChange	Robustness	RelChange	Robustness	RelChange	Robustness	RelChange	Robustness	
Netebekken									WETSPA

7.8 Maasbekken

Chooz, Meuse	**F11MAA8702**
Moelingen, Berwijn	**W11BER551010**
Marchin, Hoyoux	**W11HOY5990**
Chooz, Meuse	**W11MAAPROF**
Wanze, Mehaigne	**W11MEH5820**
Angleur, Ourthe	**W11OUR5805**
Salzinne, Sambre	**W11SAM7319**

Table 21 – The evaluation tables for the Maasbekken Basin

	NAM		**PDM**		**VHM**		**WETSPA**		**Concl<**
	RelChange	**Robustness**	**RelChange**	**Robustness**	**RelChange**	**Robustness**	**RelChange**	**Robustness**	
F11MAA8702									
W11BER551010									
W11MAAPROF									
W11MEH5820									**VHM**
W11OUR5805									
W11SAM7319									

	NAM		**PDM**		**VHM**		**WETSPA**		**Concl<**
	RelChange	**Robustness**	**RelChange**	**Robustness**	**RelChange**	**Robustness**	**RelChange**	**Robustness**	
Maasbekken									**PDM** **VHM**

8 Conclusion

Table 22 – Summary of the results of the chosen model for each basin

Basin Name	Per Catchment	Per Basin
Ijzer	PDM	PDM
Brugse Polders + Gentse kanalen	PDM/VHM	VHM
Benedenschelde + Denderbekken	PDM/VHM	WETSPA
Leie + Bovenschelde	WETSPA	PDM
Dijle and Zenne	VHM	WETSPA
Demerbekken	VHM	PDM/VHM/WETSPA
Netebekken	WETSPA	WETSPA
Maasbekken	VHM	PDM/VHM

Three of the model structures have shown to be performant for the chosen criteria: VHM, PDM and WETSPA. However, WETSPA seems to be the least robust, because not only there were several catchments where the validation was not giving good results but also because of the large variability of the quality of the results between the catchments, especially when it comes to the split sample methodology.

On the other hand, PDM has recorded good statistics in this comparison, but another important factor that can impact the decision is the high number of parameters in PDM models. This model's robustness might therefore be counteract by its high sensitivity to calibration. The comparison of several calibrations methods and strategies was beyond the scope of this analysis however. Literature supports that model robustness in the context of climate change is often decreased when overparametrized and too sensitive to the quality of the measurements.

VHM is the model that counted the highest goodness-of-fit statistics. In fact, in all of the basins, it was one of the top candidates. However, also according to previous research, VHM has the tendency to somewhat overestimates the low flows. This model is really stable however and showed good performance in most of the catchments.

Although 3 of the models appear to be more robust and taking into account the uncertainty of the hydrological modelling, it is recommended to use different hydrological models in the modelling of future climate scenarios.

9 References

Coron, L. et al., 2012. Crash testing hydrological models in contrasted climate conditions: An experiment on 216 Australian catchments. *Water Resources,* Issue 48.

Gorge, K. & al., e., 2010. *Assessment of climate change impacts on discharge in the Rhine River Basin: Results of the RheinBlick project,* sl: CHR.

Jan, S., 2003. Reliability of Model Predictions Outside Calibration Conditions. *Nordic Hydrology,* Issue 34, pp. 477-492.

MIRA, Tabari, H., Taye, M. T. & Willems, P., 2015. *Actualisatie en verfijning klimaatscenario's tot 2100 voor Vlaanderen,* Mechelen, Belgium: Vlaamse Milieumaatschappij Milieurapportering (MIRA).

Naranjo, P; Maroy, E.; Velez, C.; De Boeck K.; Pereira, F.; Nossent, J.; Mostaert, F. (2021). Modelling water availability and water allocation strategies in the Scheldt basin: Sub report 4-1 – Analyses of hydrological models for climate change modelling – NAM. Version 1.1. FHR Reports, 00_162_4-1. Flanders Hydraulics Research: Antwerp, Belgium.

Nicolle, P. et al., 2014. Benchmarking hydrological models for low-flow simulation and forecasting on French catchments. *Earth System Sciences,* Issue 18, pp. 2829-2857.

Ntegeka, V., 2011. *Assessment of the observed and future climate variability and change in hydroclimatic and hydrological extremes,* Heverlee, Belgium: Katholieke Universiteit Leuven – Faculteit Ingenieurswetenschappen.

Ntegeka, V., Willems, P., Baguis, P. & Roulin, E., 2011. *Climate Change Impact on Hydrological Extremes Along Rivers and Urban Drainage Systems in Belgium,* Brussels, Belgium: The Belgian Science Policy.

Refsgaard, J. et al., 2014. A framework for testing the ability of models to project climate change and its impact. *Climate Change,* Issue 122, pp. 271-282.

Taylor, K., Stouffer, R. & Meehl, G., 2012. An Overview of CMIP5 and the Experimental Design. *Bulletin of the American Meteorological Society,* Issue 93.

Van Steenbergen, N. & Willems, P., 2012. Method for testing the accuracy of rainfall – runoff models in predicting peak flow changes due to rainfall changes , in a climate changing context. *Journal of Hydrology,* Issue 415, p. 425–434.

Vansteenkiste, T., 2012. *Climate change and urban expansion impact on high and low flows and the overall water availability,* Heverlee, Belgium: Katholieke Universiteit Leuven – Faculteit Ingenieurswetenschappen.

Appendix 1 Methodological note on the adapted split-sample test

Review of methodologies to assess modelling tools in the context of climate change

Modelling approaches to analyse climate change effects

The common modeling steps needed to quatify the impacts of climate changes on stream flows include (1) selecting emission scenarios, (2) running global circulation models (GCMs), (3) downscaling the GCM's output to a scale that can be used for hydrology, and (4) running hydrological models that simulate the rainfall-runoff (RR) transformation at the catchment scale and for water allocation it will be necessary to add the step (5) runing a basin models to calculate the water balance including users demands. The projection of climate variability (steps 1 to 3) will be based on existing studies done for the Flanders region of Belgium; thus for the water allocation project we focus in the last two steps of the climate change analysis.

Due to the stochastic nature of weather systems, the aim of climate models is to provide information on the statistical properties of the future climate under a given scenario. This is denoted for model projections (Taylor, et al., 2012). As downstream elements of the modelling chain, hydrological and basin models, operate on outputs from climate models, therefore their outputs are projections rather than predictions. Methodologies used for making projections on future climate and its impacts may be classified into three types (see upper part of Figure 9):

- Single model: Projections based on a single model.
- Model ensemble: Projections based on an ensemble of different models, including different model codes, forcings and parameterisations
- Space-time-substitution: Identification of a place or a number of places having a past or current climate similar to the projection of the future climate at the site of interest and use of data from these places as a proxy for the impacts of the projected climate change.

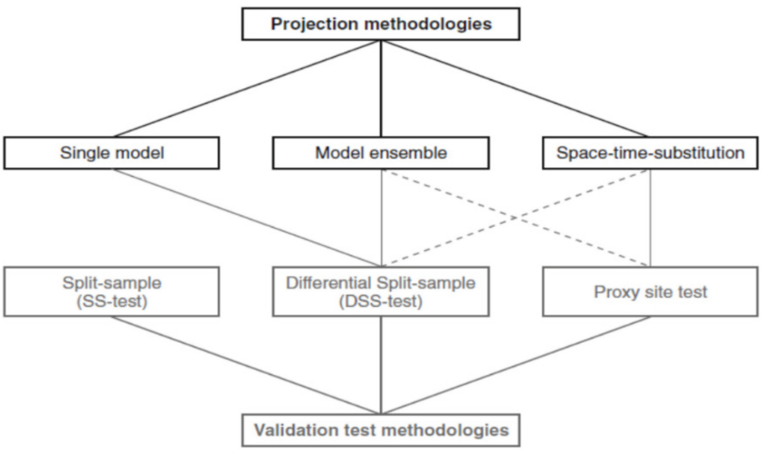


Figure 9 – Methodologies for testing model capabilities to project climate change effects

(a) Source: Refsgaard et al 2014.

In the framework of water allocation project we are calibrating four different hydrological models: NAM, PDM, VHM and WETSPA. The aim is to test different model structures and identify the best alternative to represent the hydrological process per sub-cathemnt/basin. Thus, in the climate change analysis it will be possible to use the best modelling tool per sub/catchment or make an ensemble of models to project the possible impacts.

In climate change analysis, the hydrological models are forced with future conditions that commonly differ from those for which it was calibrated and validated, therefore it is expected to lose performance. The chanllenge is to assess the performance of the models without the hydro-climatological data needed to validate the outputs. Without measurements it is not possible to define which model has the best performance. To face this problem researchers has proposed approaches that are based on the pass records of data and focus in evaluating the loss of accuracy of the model to represent the hydrological responds as it is forced with inputs that differ from the ones used in calibration/validation of the model (Refsgaard, et al., 2014). Contrary other approaches used the projected climate data and assess statistical properties of the hydrological outputs with the variations in the statistical properties of the forcing climate inputs (Van Steenbergen & Willems, 2012). In what follow are describe the two approaches.

Framework for testing the model capability to project climate change effects

Before it is used operationally, a model must demonstrate how well it can perform the kind of task for which it is intended. That is typically done by carring out a a a validation test with data that differ from the one used in the calibration. Validation methodologies may in accordance with Klemes (1986) be classified into three types (see lower part of Figure 9):

- Split-sample test (SSt): The data are, temporally and/or spatially, split into two parts. One part is used for calibration, while the other part is reserved for validation. A more sophisticated version of the SS-test is jack-knifing where the data split and testing are repeated systematically, so that all data are used for both calibration and testing. The underlying assumption behind the SS-test is that climate conditions, as well as physical conditions, can be assumed stationary.
- Differential split-sample test (DSSt): DSS-tests are applicable if climate conditions are non-stationary. The test implies that a model ideally is tested against observation data similar to the future climate conditions. Due to lack of such data, DSS-tests are often made using periods with apparent different climate conditions (e.g. dry/wet or cold/warm) where calibration is performed on one period and validation on another period.
- Proxy site test: A proxy site test is a test of the capability of the model to project conditions without prior calibration, i.e. completely without calibration or by calibration at some locations with site-specific data and projection at the location of interest.

The split-sample test is the most frequently used approach to validate models, probably because it is the easiest and the only test type for which data are readily available. As the underlying assumption behind the SSt is that climate conditions are stationary, it is not an adequate test in climate change impact studies: in the framework for testing the ability of models to project climate change, Refsgaard et al. (2014) recommended the use of the differential-split sample test (DSSt) in order to build confidence in the model projections. Proxy site test are more frequently applied in the context of climate change impacts on ecosystems. Thus in what follow we focus in the DSSt as the one of the strategies that we could use to validate the hydrological models in the climate change context.

Differential Split-Sample test

DSSt have been used for both single models and ensemble models in hydrological applications. For instance, Seibert (2003) used a model to simulate peak flows in four Swedish catchments and Gorge et al. (2010) used the DSSt to assess the reliability of a set of seven lumped hydrological models for the River Rhine. DSS tests rely on the existing records of climate variables and hydrological response, hoping that there is enough variation in the sub-periods to represent the kind of variability of the climate conditions for which the model will be projecting hydrological response. The DSSt method follow three steps:

1. A small number of sub-periods are selected according to one climate characteristics,

2. The calibration – validation test is applied on these periods,

3. The validation performances are compared to evaluate whether they vary significantly when climatic characteristics differ between calibration and validation.

Procedure to select test periods

The most commonly used data source within all disciplines is data from the recent past climate. This provides possibilities to test the capability of models to reproduce climate variability and its impacts as well as spatial differences. In this respect three data sources are interesting: (i) historical time series comprising non-stationarity, (ii) paleo data and (iii) data from controlled experiments. In order to perform DSSt, Refsgaard et al. (2014) recommended to use data sets containing non-stationarity data. Many historical time series exhibit non-stationarity, mostly due to human activity such as land use change or river regulation. An example where climate is the dominant source of nonstationarity is the time series from Skjern River in western Denmark where the precipitation has increased by 26 % and the temperature by 1.3 °C since 1875 (Karlsson et al. 2013). A DSS-test showed that a hydrological model had difficulty predicting the changes in runoff during this transient period. In the context of this water allocation project, we have historical time series for both calibration and validation. The available data extends on approximatelly 40 years, for 57 gauged catchments.

The selection of the sub-periods to be used to test the validity of the models can follow different approaches. One common way of selecting is by difining an indicator of the type of climate variable and used to split the data. An example of the application of this approach for the River Rhine cathemtn is shown in Figure 10. To apply the DSSt, Gorge et al. (2010) selected yearly sub-periods using as index the mean annual rainfall and mean annual temperature. Other indexes could also be used for year selection, for example extreme rainfall or drought events, or indexes based on the variability of conditions on a seasonal basis. Using mean annual precipitation and temperature as spliting index, four possibilities of climate change were analysed for the river Rhine:

- change towards drier years (i.e. calibration on wet years and validation on dry years);
- change towards wetter years (i.e. calibration on dry years and validation on wet years);
- change towards warmer years (i.e. calibration on cold years and validation on warm years);
- change towards colder years (i.e. calibration on warm years and validation on cold years).

One of the limitation of this approach to select the subperiods is that it is often difficult to compare the results because the subperiods used are different (e.g, the driest period may differ from the warmest one). The second limitation pointed out by Coron et al. (2012) is that the number of transfer tests is usually small, as often only two or three contrasted periods can be identified (four test in the case of River Rhine shown above). This limits the possibility of drawing general conclusions and discovering the main drivers of parameter transferability from the results themselves. Indeed, it might be hard to distinguish the effect of the climate difference from other aspects potentially influencing parameter transfer.

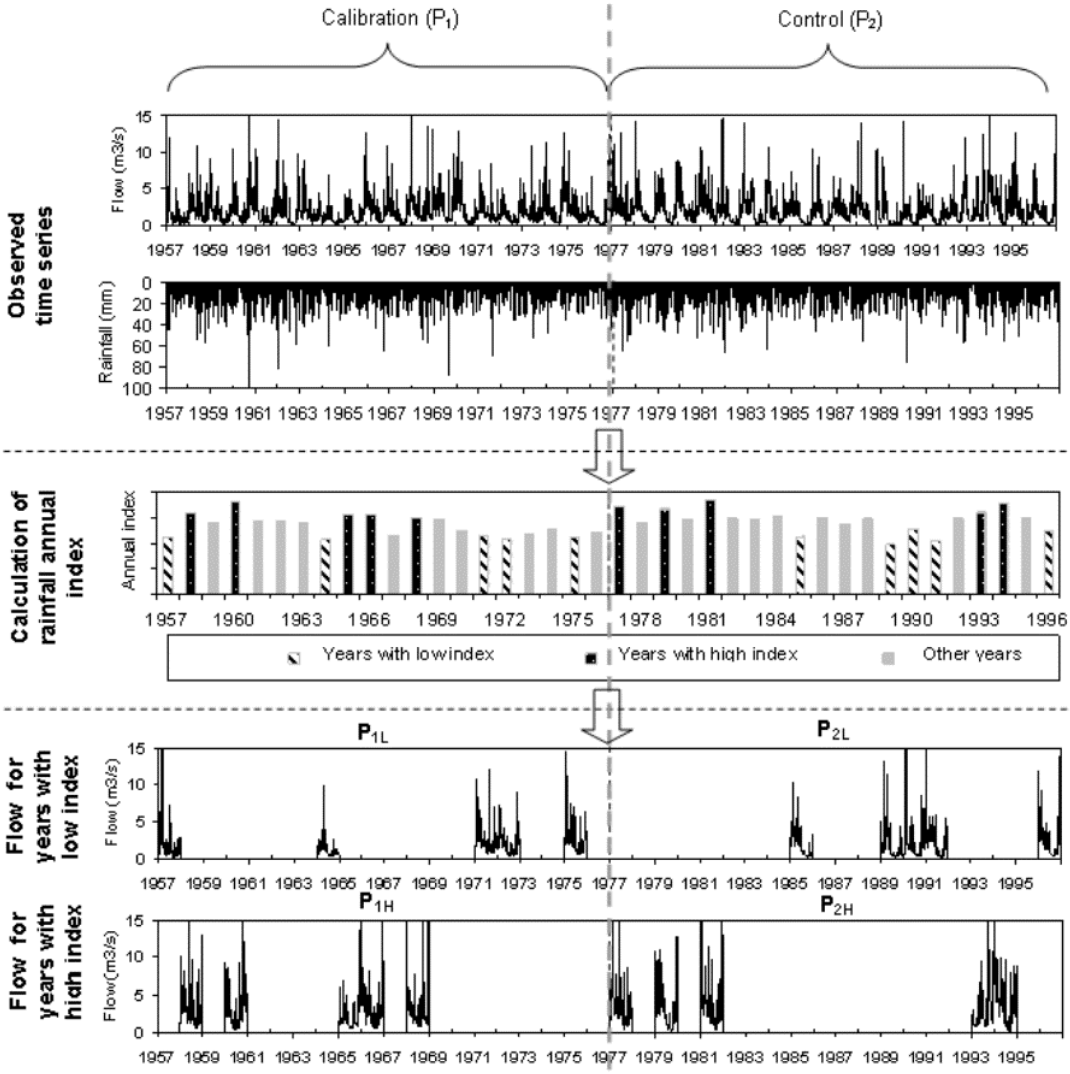


Figure 10 – Illustration of procedure used to selec the test periods for the River Rhine models (Gorge & al., 2010).

Another approach for the selection of the sub-periods is demonstrated in the application of DSSt carried out by Coron et al. (2012) with three hydrological models in 216 catchments in Australia. The authors used a sliding windows of chosen length (e.g. 5 years) to define sub-periods with the historical data. Between two periods the window is moved by 1 year, allowing the sub-periods to overlap. In the left part of Figure 11 it is illustrated the splitting of equal size sub-periods (dark grey bars) used in the Generalized DSSt (GDSSt) by Coron et al. (2012). The aim of the authors with this approach is to test the model in as many and as varied climatic configurations as possible, including similar and contrasted conditions between calibration and validation. In the example, with 18 years of historical data avaible and a 5 year subperiod the authors have 14 sub-periods for calibration/validation of the model. In the rigth part of Figure 11 it is illustrated the approach for the validation. In summary, for each sub-period the model is first calibrated and then, in the validation phase, the model with the best parameter fit for that period is used to test the remaining periods.

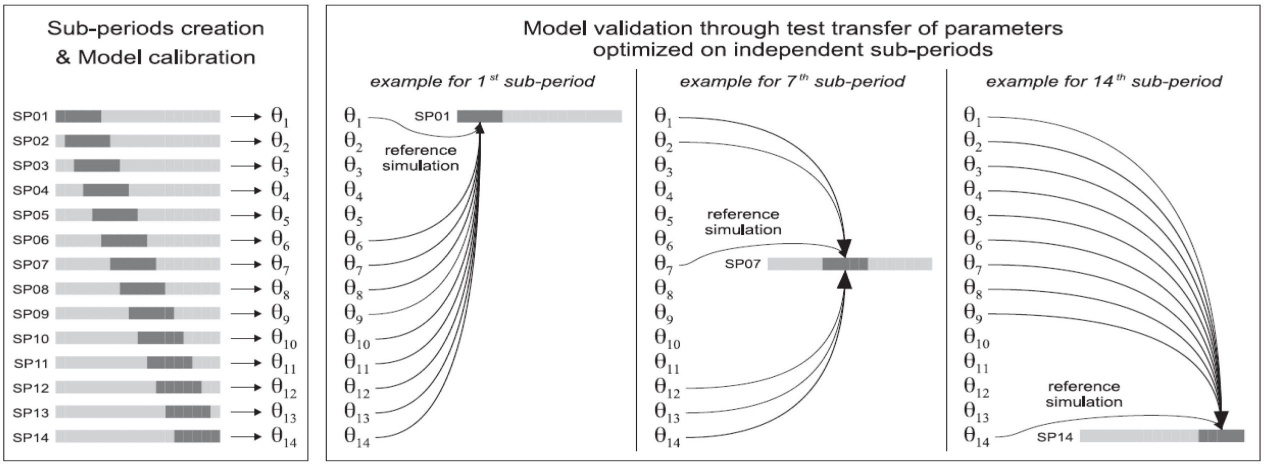


Figure 11 – Illustration of generalized split-sample test procedure

([a]) Source: Coron et al., 2012.

Peformance criteria used in the DSSt and visualization of the results

When a model is used to simulate discharges, errors will arise : (1) for reasons which were already noticeable during calibration (data and model structure errors, identifiability issues, etc.) and (2) by the move from the calibration period to another period leading to the use of less than optimal parameters for this application period (Merz et al., 2011). Separating these two sources of error is essential to achieving an informative evaluation of the extrapolation capacity of hydrological models: a model may work well in calibration but show poor transposability over time. To evaluate the performance of the models different validation criteria are applied. Performance criteria are estimated based on indicators of error between observed and simulated variables. In general, the performance criteria should reflect the conditions in climate change projections and be purpose-specific. Different criteria should be applied for different study purposes (Madsen 2000). As a single criterion cannot assess all the qualities of hydrograph simulation, it is usual to use a set of numerical criteria.

The evaluation criteria used in the analysis of the impacts of climate change for the River Rhine are shown in Table 23. For that application the authors used normalized criteria to give an evaluation of model results comparable between catchments. They put emphasis (1) on mean flow and regime simulation, (2) on low flow simulation and (3) on high flow simulation respectively. Three of the statistics are based on Nash and Sutcliffe [1970] (NS) criterion. Notice that NSLF is calculated on logarithm transformed daily flows to put more emphasis on low flows and NSHF is calculated on daily flows, which puts more emphasis on high flows. These three NS statistics measure the match between simulated and observed series. The three other criteria are ratios between simulated and observed flow statistics. The ratio between simulated and observed mean flows (RMQ) is equivalent to the relative bias. The RFDC_Q90 and RFDC_Q10 criteria are based respectively on the 90% and 10% (exceedance) percentiles of the flow duration curve (i.e. low and high flows respectively) and represent the ratio between simulated and observed values (Gorge & al., 2010).

To get a general assessment of model performance over the Rhine River basin, Gorge et al. (2010) used mean values of efficiency criteria over eight target stations. Note that as values lower and greater than 1 may compensate for the RMQ, RFDC_Q90 and RFDC_Q10 criteria, the authors considered the absolute departure of their values from unity to calculate the mean, which is thus given by:

$$m(x) = \frac{1}{8}\sum_{j=1}^{8}\left|1 - x_j\right|$$

where m(x) is the mean value of criterion x and xj is the value of criterion x on catchment j.

Table 23 – Peformance Criteria used for Evaluation of Models in the DSSt of the River Rhine (Gorge & al., 2010)

	Name	Formulation	Meaning	Range
Mean flow and regime	RMQ	$RMQ = \frac{\sum_{i=1}^{n} Q_{sim,i}}{\sum_{i=1}^{n} Q_{obs,i}}$	Ratio between the simulated and observed mean flows over the evaluation period (also called bias)	[0;+∞[1: perfect fit Values lower than 1: underestimation Values greater than 1: overestimation
	NSMMF	$NSMMF = \frac{\sum_{i=1}^{12} (Qm_{sim,i} - Qm_{obs,i})^2}{\sum_{i=1}^{12} (Qm_{obs,i} - \overline{Qm_{obs}})^2}$	Nash-Sutcliffe efficiency index calculated on mean monthly flows	]-∞;1] 1: perfect fit Values lower than 0: model worse than simulating a constant flow equal to monthly mean observed flow
Low flow	RFDC_Q90	$RFDC\_Q90 = \frac{Q90_{sim}}{Q90_{obs}}$	Ratio between the simulated and observed 90% (exceedance) percentiles of daily flows	[0;+∞[1: perfect fit Values lower than 1: underestimation Values greater than 1: overestimation
	NSLF	$NSLF = \frac{\sum_{i=1}^{n} (\ln(Q_{sim,i}) - \ln(Q_{obs,i}))^2}{\sum_{i=1}^{n} (\ln(Q_{obs,i}) - \overline{\ln(Q_{obs})})^2}$	Nash-Sutcliffe efficiency index calculated on logarithm transformed daily flows	]-∞;1] 1: perfect fit Values lower than 0: model worse than simulating a constant flow equal to mean logarithmic transformed observed flow
High flow	RFDC_Q10	$RFDC\_Q10 = \frac{Q10_{sim}}{Q10_{obs}}$	Ratio between the simulated and observed 10% (exceedance) percentiles of daily flows	[0;+∞[1: perfect fit Values lower than 1: underestimation Values greater than 1: overestimation
	NSHF	$NSHF = \frac{\sum_{i=1}^{n} (Q_{sim,i} - Q_{obs,i})^2}{\sum_{i=1}^{n} (Q_{obs,i} - \overline{Q_{obs}})^2}$	Nash-Sutcliffe efficiency index calculated on daily flows	]-∞;1] 1: perfect fit Values lower than 0: model worse than simulating a constant flow equal to mean observed flow

([b]) $Q_{obs,i}$ and $Q_{sim,i}$ stand for observed and simulated flows at day *i*; *Qm* stands for monthly mean flow; *Q90* and *Q10* stand for the 90% and 10% exceedance percentiles of the flow duration curve; *Q* stands for the mean of *Q*).

An example of the illustration of mean results obtained on the 8 target gauging stations of the River Rhine using the differential split sample test is shown in Figure 12. Results are shown for seven lumped models and six efficiency criteria for the case of model validation for dry years. Base on the plotted performance criteria is possible to observe the effect of the calibration period when the model is validated on dry periods. These plots are usefull to compare the models but it is limited to the three precipitation classes (i.e. low, medium and high).

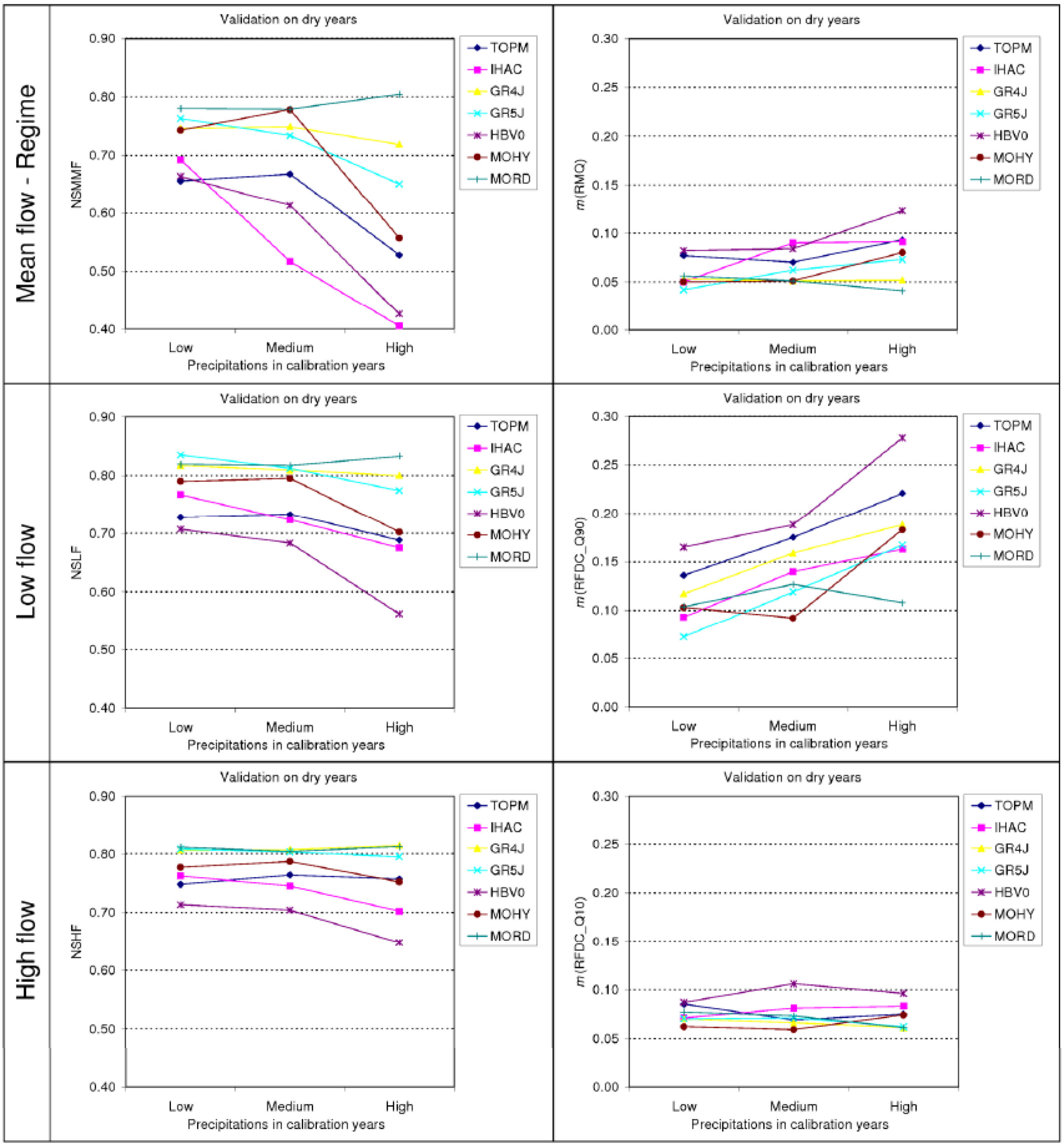


Figure 12 – Sensitivity of model results to the calibration conditions validated for dry years

([a]) Source: Gorge et al., 2010.

Another way to assess the performance losses caused by the parameter transfer from calibration to validation is presented by Coron et al., (2012). For the GDSSt, Coron et al., used the performance criteria shown in Table 24. In their notation the authors enfasis the transfer of parameter set *Θ* from a period D (‘‘donor’’, i.e., calibration) to a period R (‘‘receiver’’, i.e., validation). With these notations, they describe the three performance criteria: the root-mean-square error (RMSE) the Nash-Sutcliffe efficiency (NSE) and the bias on total volumes.

Notice that the criteria are measurements of the error between model and observations, simmillar to the choice make by the authors in the River Rhine. Bias is also calculated using a relative formulation which provide values that are comparable between periods and catchments. The authors also argue that $RMSE_{D\text{->}R}$ or $NSE_{D\text{->}R}$ values for different D periods but a single R period can be directly compared since all errors are calculated on the same time steps. However, the RMSE is dependent on the mean volume and will tend to be greater for periods (or catchments) showing larger discharges. Limitations appear when the periods compared show contrasted climatic properties and hence contrasted flow levels. It becomes even more complicated when results from different catchments are analyzed together. Instead of using differences between criteria, Coron et al., (2012) proposed a model robustness criteria (MRC) to analyse the influence of changes in climate on parameter transferability (right side of Table 24). The main idea is that the quality of a given parameter set is assessed relative to a reference set, obtained through calibration. $\varepsilon_{D\text{->}R}$ is one estimate of the model error on period *R* using the parameters calibrated on period *D*. $\varepsilon_{R\text{->}R}$ should be the smallest value of ε achievable on period *R* with the model. $\varepsilon_{D\text{->}R}$ and $\varepsilon_{R\text{->}R}$ are comparable since they are computed on the same ‘’receiver’’ period.

Table 24 – Peformance Criteria used for Evaluation of Models in the GDSSt on rivers in Australia
(source : Coron et al., (2012))

$\mathrm{RMSE}_{D\to R} = \sqrt{\frac{1}{n}\sum_{k=1}^{n}\left(\hat{Q}_{R,k}[\theta_D] - Q_{R,k}\right)^2}$ $\mathrm{NSE}_{D\to R} = 1 - \frac{\sum_{k=1}^{n}\left(\hat{Q}_{R,k}[\theta_D] - Q_{R,k}\right)^2}{\sum_{k=1}^{n}\left(\overline{Q}_R - Q_{R,k}\right)^2}$ $\mathrm{BIAS}_{D\to R} = \frac{\sum_{k=1}^{n}\hat{Q}_{R,k}[\theta_D] - \sum_{k=1}^{n}Q_{R,k}}{\sum_{k=1}^{n}Q_{R,k}}$	$\mathrm{MRC}_{D\to R} = \frac{\varepsilon_{D\to R}}{\varepsilon_{R\to R}} - 1$ Where ε = RMSE or NSE or BIAS

([a]) $Q_{R;k}$ is the observed discharge at time step *k* on period *R*, $\hat{}Q_{R;k}\,[\Theta_D]$ the simulated discharge at time step *k* on period *R* using the parameter set Θ optimized on *D*, and *n* is the total number of time steps in period *R*

([b]). MRC is the model robustness criteria

To illustrate the relative loss of performance of a hydrological model, Conors et al., analysed the variations in MRC values relative to the differences in climate between the calibration and validation periods. Changes in climate were expressed as ratios (e.g. 10% less rainfall). The procedure is illustrated in Figure 13 in which each MRC value was plotted against the corresponding change of the selected climate variable. In plot (a) several parameter transfer tests carried out on a single receiving period is shown. Because all values on the x and y axes are relative, the results for all the other receiving periods can be plotted on the same graph. In plot (b) the procedure was then repeated for all the catchments; and at last in plot (c) box plots are draw to facilitate the analysis of the model transferability in the entire range of climate range and for all catchments.

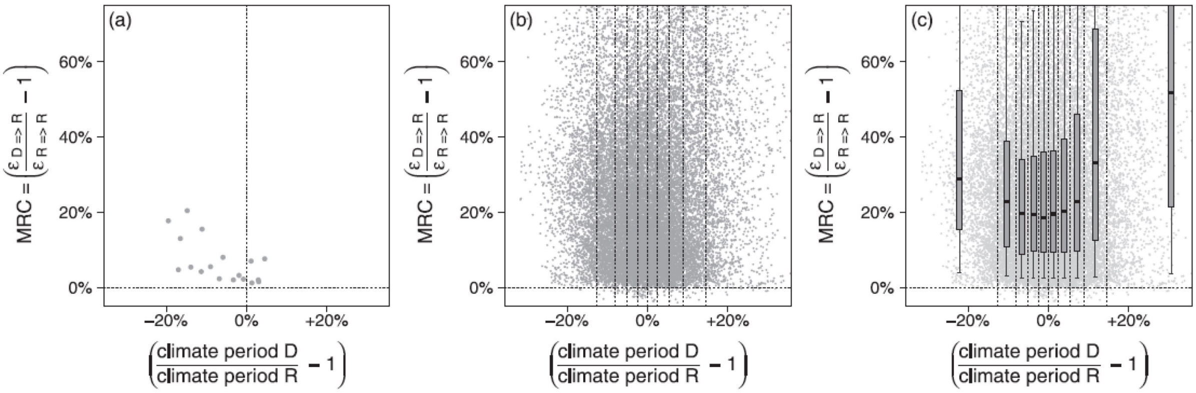


Figure 13 – Procedure followed to illustrate the relative loss of performance of one model against the relative evolution of climate (source : Coron et al. (2012))

([a]) Dotty plot for a single period, ([b]) dotty plot for all periods of all catchments, and ([c]) summary of dotty plots as box plots.

When comparing the two set of performance criteria used for the two independent applications of the DSSt it is possible to see that the authors choose the tranditional indicators of the error between observed and modeled variables. However, Gorge et al. (2010) use variation of the NSE for low flows and for high flows which favor a more accurately evaluation for the type of condition assessed. They both take into account a sort of relative form of the criteria to waranty certain degree of comparativility for the different sub-periods for which the criteria are computed. Limitations in the comparability of the criteria are argue by Conors et al., but their proposed MRC that is in fact a ratio of the error measurement (criteria) does not address the problem. The sppliting strategy of the GDDSt allow a more detail analysis of the change in model accuracy in the range of climate variation on the historical time series available. However, the simple three climate variations of the approach of Gorge et al., has the advantage that all models can be compare in one plot. On the contray the cloud of points for all catchment proposed by Coron et al. (2012), has so much scatter that there is a lost of sensitivity in the trends even when points are sumarized in box plots.

One of the limitations of the DSSt shown is that they are based on historical data, and therefore the variation of climate conditions is limited to the variation exisiting in the system which may not content simmillar variability of that projected with the climate change scenarios. Another limitation of the approaches presented above is thatClimate models are not able to reproduce single events but rather statistical properties describing the climate. This should also be reflected in tests of hydrological models. Therefore, commonly used performance criteria in hydrology such as the root mean square error and the Nash-Sutcliffe coefficient focusing on temporal fits to observed data, are not relevant for this type of test. Criteria based on flow duration curves or quantiles in probability distribution functions are much more appropriate (Van Steenbergen & Willems, 2012). In the following secction is presented an approach that focus in the use of projected climate conditions and the statistical properties of the projected values.

Use of relative change between current climate and climate change scenarios

The methodology proposed by Van Steenbergen and Willems (2012) looks for a relation between the relative change in peak flow for different sub-periods of the same record as a result of rainfall increase. This is done both for the observations and the simulation results. The analysis starts with the split of the observed time series in nearly independent quick flow hydrograph periods. For each period the peak rainfall volume is calculated as the maximum rainfall volume during a period equal to the recession constant of the overland flow. The peak rainfall volume is considered responsible for the overland peak flow. For each nearly independent quick flow hydrograph period the overland peak flow and the peak rainfall volume are extracted. When different periods are being compared the relative change in overland peak flow and the relative change in peak rainfall can be calculated.

Procedure to select quick flow hydrograph **periods**

To select the peak flows, the runoff series are divided in nearly independent quick flow hydrograph periods. These periods are selected based on the time series processing tool of Willems (2009). After this selection of peak flows, the lowest flow values in between two consecutive peaks are defined and the time series is divided in periods based on the time moments of these low flows. A simulated peak flow is paired with an observed peak flow, if the simulated peak appears within a time window of 10 h around the observed peak, allowing small phase errors in the modelling results.

Climate change analysis on peak flows

Since the methodology was developed to assess climate change impacts on floods, the authors use performance criteria that measure how well models predict extreme floods, such as floods with recurrence intervals of 5, 10 or 50 years (Van Steenbergen & Willems, 2012). To construct the empirical extreme value distribution, empirical return periods of the peak flows are calculated based on the rank number of each peak flow after sorting of the peak flows. For the ith highest peak flow in a time series of length n years, the return period of that event is given by:

$$T(i) = \frac{n}{i}$$

An example of the empirical extreme value distribution of peak flows is shown in Figure 14. The comparison shows the peak flows of the observations and the three models assessed by the authors. The example shows that the model performance in terms of peak flow statistics and extremes is more or less similar for the three models.

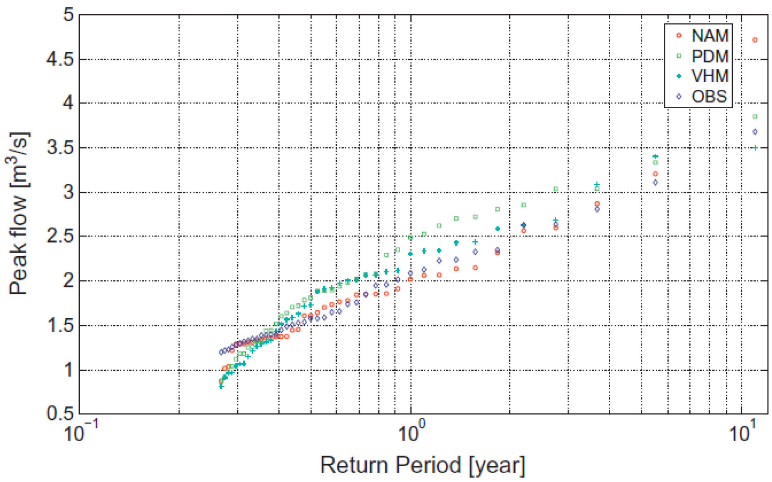


Figure 14 – Example of the empirical extreme value distribution of peak flows

([d]) Source: van Steenbergen and Willens (2012)

Climate change impact analysis **on peak flows**

The climate change scenario used for the authors corresponds to those developed for the Belgian region. The authors used the outputs of the perturbation tool of Ntegeka and Willems (2009) to translate observed time series (rainfall, Evapotranspiration) into future time series for three climate change scenarios: High scenario: wet winters and dry summers; Mean and Low scenario: dry winters and dry summers. According to van

Steenbergen and Willens (2012), the mean and low scenarios give comparable results to the actual climate (1998–2008), because these scenarios produce rainfall and evapotranspiration in the range of the actual climate. The changes in peak flow are much larger for the high scenario. The largest rainfall events simulated in the high scenario are higher than the ones covered by the calibration period; thus forcing the models to make an extrapolation.

To asses the performance of the models for this type of extrapolation, an analysis of the change in peak flow as a result of rainfall increase is made using the denominated Relative Change in peak flow between the actual climate and the climate change scenario. The equiation to calculate the relative change is shown below:

$$Rel.Change_{act,cc,T(i)} = \frac{q_{p,cc,T(i)} - q_{p,act,T(i)}}{q_{p,act,T(i)}}$$

where Rel.Change$_{act,cc,T(i)}$ is the relative change in peak flow (q_p) between the actual climate (*act*) and the climate change scenario (*cc*) for a return period *T(i)*.

To represent the results and compare between different models, boxplots of the relative changes in the peak flows can be compared for each climate change scenario. An example of the box plots used to compare three hydrological models and three climate scenarios is presented in Figure 15.

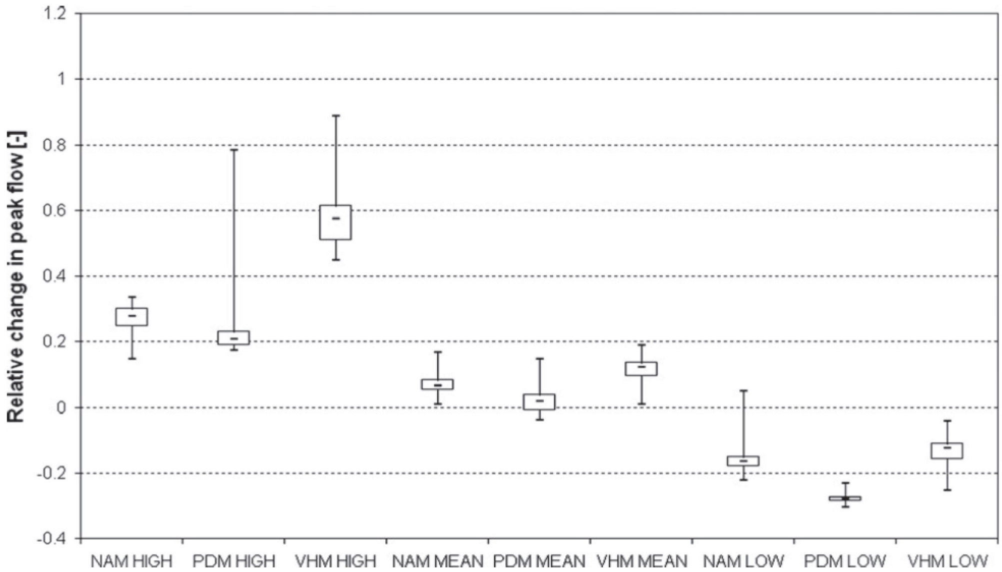


Figure 15 – Illustration of relative change in peak flow in comparison with the actual climate for different climate change scenarios and hydrological models.

([a]) HIGH: high climate change scenario, ([b]) MEAN: mean climate change scenario, and ([c]) LOW: low climate chage scenario.
([d]) Source: van Steenbergen and Willens (2012)

The advantege of this approach is that use statistical characteristics of precipitations and flows from both the historical data and the projected data. It also seems to be useful to compare different modelling approaches which is one of the aim in the water allocation project. However, it seems that for one model the approach is limited by the fact that it is not clear how to define which one is better. The criteria seems to be weak in the sense that modeller can not define what is performance. Within the limitations is also the fact that the appoach was mainly developed for flood management, therefore in the context of the project the approach need to be adapted.

Methodology to assess model projection capability for water allocation

Integrated approach to assess model projection capabilities

Based on the literature review presented above, an integrated methodology is proposed to assess the projection capability of the calibrated models in the framework of water allocation. This methodology combines a differential split-sample test, using historical data, with the comparison of the relative change of statistical descriptors between various model projections. The general scheme of the proposed approach is presented in Figure 2. The methodology starts with calibrating the models (which was the object of previous reports), then splits into two branches:

- One based on historical data only, to assess the model robustness when validation forcing is leading to lower flows than calibration forcing, by evaluating change of performance
- the other based on perturbated time series, to compare different model behavior and variability in projection mode.

Models are evaluated on different criteria and the aim is to generate a hierarchy of modelling structures according to their ability to project the climate change effects and the level of confidence that can be associated to those predictions.

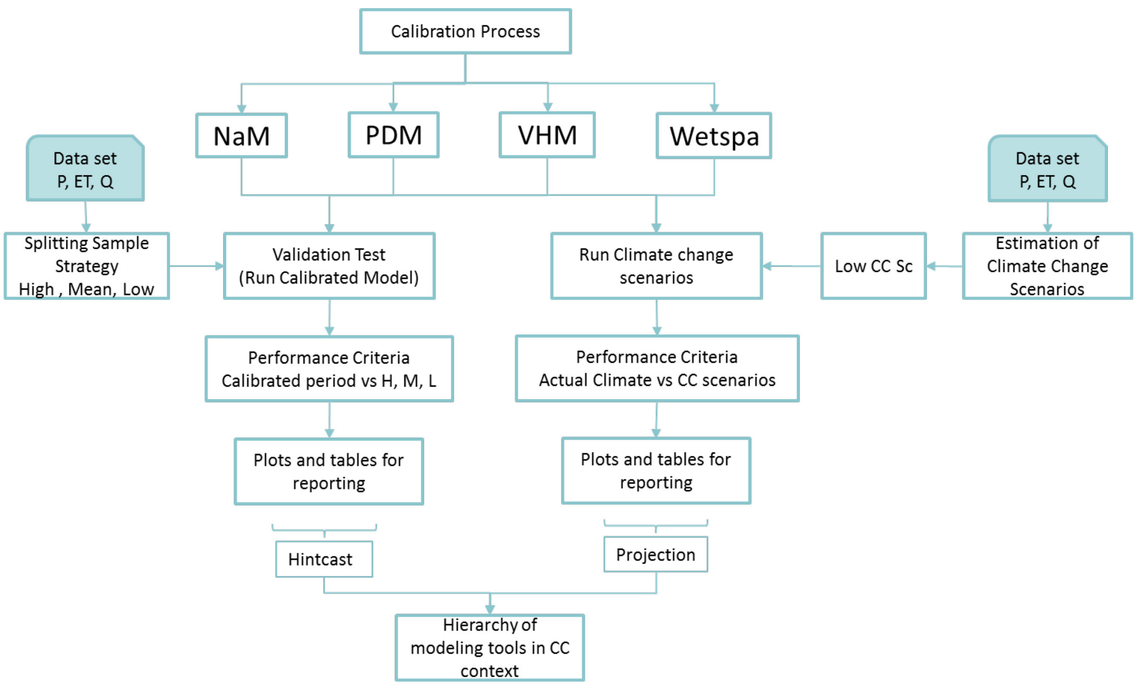


Figure 16 – Illustration of Integrated Approach to Assess Model Projection Capability in Climate Change Context

Calibration and validation of hydrological models

The models were calibrated following the split sample approach using historical data from 1990 to 2012. Precipiation data is available since 1967 but there is missing data of discharge therefore the splitting is not always homogeneous and the length of calibration/validation periods variesThe most recent 10 years of existing data were used for calibration and the rest was used for validation. Figure 17 shows as an example of observed data of total flow and the model output for the sub-catchment W08SAMRON000 (Dijle/Zennebekken).

To calibrate the models a multi-objective optimization algorithm was used. The NSGAII was linked to the hydrological modelling tools (NAM, PDM and VHM). Details on the approach can be found in Naranjo et al. (2020). In the calibration process two indictors of goodness of fit (IGoF) were used as objective function in the optimisation process. Overall, the methodology for the calibration of hydrological models in the context of water allocations give priority to calibration of base flows and average flows but at the same time preserve the volume balance. The IGoF used were the Log NSE on the time series of simulated and observed flows and Volume Absolute Error on the cumulative flows. Other IGoFs were calculated in each model to complement in the analysis of the performance of the models in the calibration and validation process. That is on the time series of observed and modelled: the Root Mean Square Error (RMSE), the NSE, the NSE relative, the Kling-Gupta efficiency (KGE); and in the cumulative volume the relative error and the volume bias.

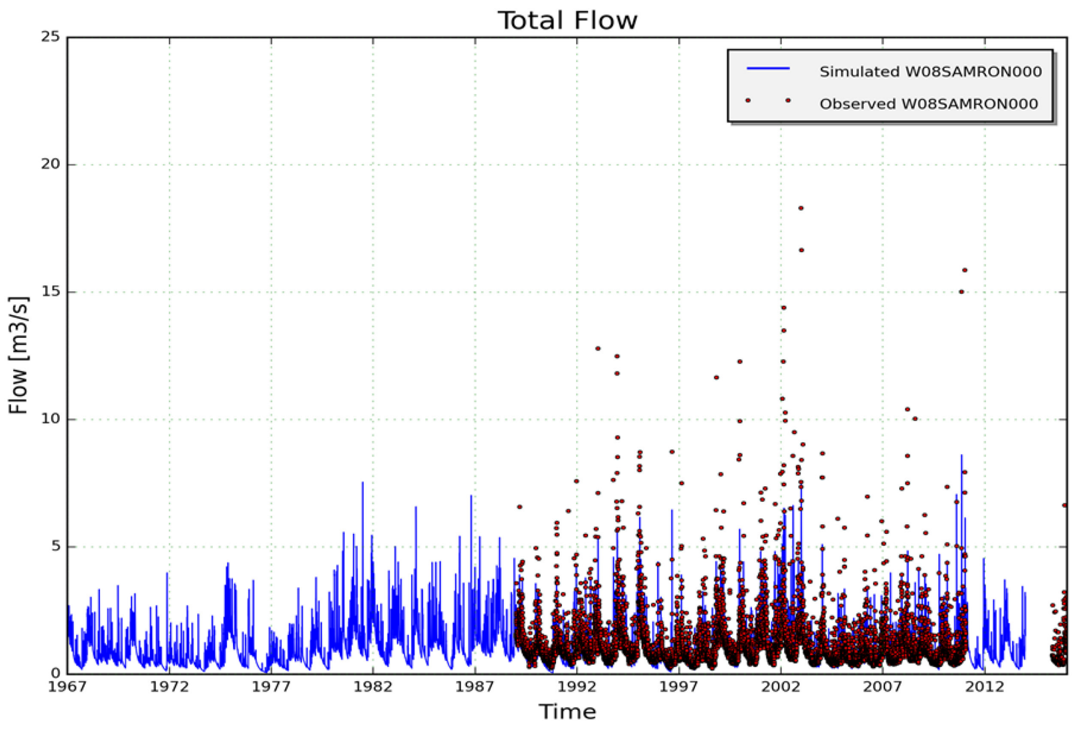


Figure 17 – Illustration of Information of Flows
Available for Calibration and Validation in the sub-catchment W08SAMRON000

(a) source: modelling report from calibration of sub-catchment W08SAMRON000

Modified split-sample test to assess model projection capabilities

Based on the literature review presented above, a modified split-sample approach is proposed to assess the projection capability of the calibrated models in the framework of water allocation. Splitting Approach

A modified version of the differential splitting method described by Coron et al. (2012) is used to compare the robustness of the different model structures when applied to drier conditions than the calibration period.

To choose the moving window strategy, care was taken that significant changes was found in the climate variables. Initially, a window of 10 years was used to assimilate the period used for calibration but there were no significant differences of the climate variables between the samples (e.g. maximum difference found in periods of 10 year was 3% in precipitation volume). By reducing the windows to 3 years there are periods that reach up to 25% differences in volume of precipitation (Figure 18).

Some authors may argue that validation should be done with data that differ from the calibration, but in this approach we have decided to use all the available data to verify the model capability in the most contrasting climate conditions possible. Note that in the approach of Coron et al. (2012), all available data sets were used as calibration and validation successively.

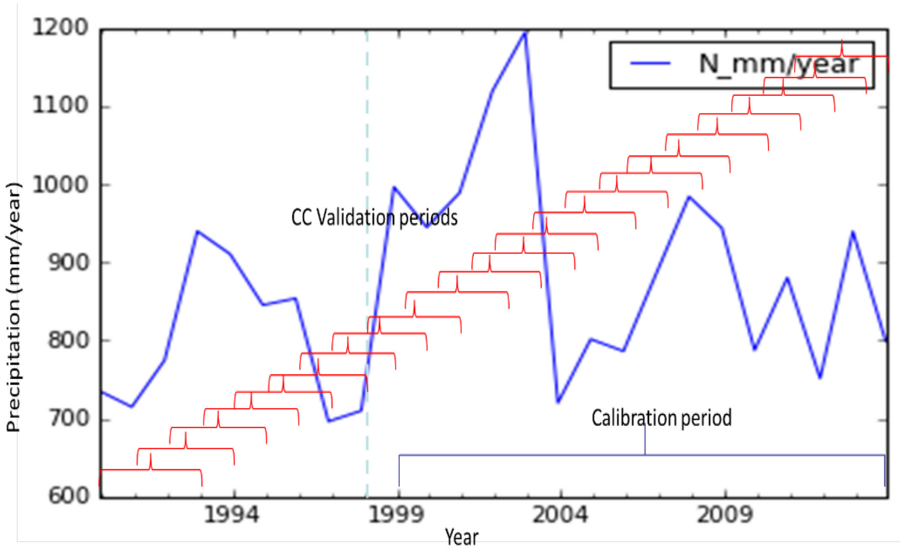


Figure 18 – Illustration of splitting of data for calibration and validation

Performance criteria

The criteria selected as indicators of model performance are based on the “Model Robustness” indicator from the GDSSt approach (Coron, et al., 2012) and the “Relative Change” ratio as in the approach described by van Steenbergen and Willems (2012). The indicators used to calculate the criteria are selected based on the purpose of the model. In principle, a set of 7 criteria were calculated for each sub-period modelled from the historical data. However, priority was given to performance criteria measuring the ability to reproduce mean annual flows or minimum flows and volume balances. Thus for the MRC error the Log Nash-Stcliffe and the Bias Volume were used and for the Relative Change the Ration of Q50/Q90 estimated from the duration curves Mandal and Cunnane, (2009). The criteria selected is presented in Table 25.

Table 25 – Peformance criteria used for evaluation of models

1. Model robustness criteria (MRC) $$\mathrm{MRC}_{D\to R} = \frac{\varepsilon_{D\to R}}{\varepsilon_{R\to R}} - 1$$ εD->R = NSE and BIAS D->R εR->R = smallest value of ε in calibration period	Error criteria Log Nash-Sutcliffe efficiency (NSLog) $$NSLog = \frac{\sum_{i=1}^{n}(\log(Qobs,i) - \log(Qobs))^2}{\sum_{i=1}^{n}(\log(Qobs,i) - \log(\overline{Qobs}))^2}$$ Bias on total volumes $$\mathrm{BIAS}_{D\to R} = \frac{\sum_{k=1}^{n}\hat{Q}_{R,k}[\theta_D] - \sum_{k=1}^{n}Q_{R,k}}{\sum_{k=1}^{n}Q_{R,k}}$$
2. Relative Change $$\mathrm{Rel.Change}_{act,cc,T(i)} = \frac{Qr_{cc,T(i)} - Qract_{T(i)}}{Qr_{act,T(i)}}$$	Qr = ratio Q50/Q90 Qr is estimated from the duration curve and may represent the variability of low-flow discharges

Validation test using the differential split sample

After the splitting of the historical data and calculating the criterias proposed above the performance of the model can be assessed in comparison with the variation in the climate conditions. For every sub-period (3 years) the indicators are calculated by comparing the model results with the observations of flows. An example of the evaluation of the performance of the NAM model on 20 sub-periods is shown in Table 26. The second column shows the sub-periods and the third column the average annual precipitation in mm. Columns 4 to 10 shows the results of the error criteria calculated for each period.

Table 26 – Example of the evaluation of performance of the NAM model for the sub-catchment W08SAMRON000

	TestPeriod	Nannual	NS	NSlog	kge	RMSE	VolBias	AbsErr	RelErr
0	1989-1991	742	0.64	-0.48	0.67	0.45	0.74	154	-0.239
1	1990-1992	810	0.62	-0.36	0.65	0.47	0.75	166	-0.293
2	1991-1993	875	0.65	0.47	0.59	0.56	0.78	154	-0.237
3	1992-1994	898	0.70	0.66	0.66	0.50	0.86	122	-0.167
4	1993-1995	870	0.75	0.66	0.73	0.49	0.90	89	-0.115
5	1994-1996	798	0.81	0.47	0.85	0.40	0.92	40	-0.053
6	1995-1997	753	0.80	0.45	0.82	0.45	0.89	60	-0.091
7	1996-1998	801	0.70	0.49	0.70	0.51	0.86	80	-0.175
8	1997-1999	884	0.71	0.75	0.69	0.53	0.90	63	-0.114
9	1998-2000	977	0.71	0.76	0.71	0.47	0.96	56	-0.078
10	1999-2001	1017	0.73	0.75	0.74	0.44	1.04	22	-0.005
11	2000-2002	1101	0.65	0.73	0.61	0.56	1.06	56	0.068
12	2001-2003	1011	0.66	0.73	0.62	0.61	1.05	68	0.070
13	2002-2004	905	0.66	0.71	0.63	0.70	0.97	21	-0.001
14	2003-2005	769	0.68	0.59	0.71	0.59	0.92	38	-0.050
15	2004-2006	824	0.68	0.60	0.67	0.53	0.89	69	-0.136
16	2005-2007	885	0.66	0.63	0.69	0.49	0.95	41	-0.085
17	2006-2008	938	0.65	0.71	0.64	0.53	0.99	24	-0.043
18	2007-2009	905	0.61	0.61	0.67	0.55	1.06	17	0.009
19	2008-2010	870	0.57	0.46	0.60	0.74	1.17	55	0.082

([a]) Nannual = Annual average precipitation volume in mm for the test period, NS = Nash-Suffclife, NSlog = Log NS, KGE = Kling-Gupta efficiency, RMSE = Root Mean Square Error, VolBias = Volume Bias on cumulative flows, AbsErr = Absolute error and RelErr = Relative Error on cumulative flows.

In the approach proposed by Coron et al., (2012), the authors used the smallest value of the error in the calibration period to define the reference period to compare the other sub-periods performances. In the preliminary test shown in Table 26 it is possible to see that depending on the indicator there are differences in the performance. Therefore not always the best performance coincide for the same period (see numbers in red and purple for best performance in calibration periods). When comparing the clould of points form by the performance criteria using diiferent reference period it can be noticed that the reference is not significantly important but the spread of the points as the climate vary (see in Figure 19 the cloud shifted by the effect of choosing different reference sub-period). Therefore, in the approach it is proposed to use the climate variability an center the cloud with respect to mean.

To compare evaluate the performance of each model the cloud of variables will be analyzed using the box plots as in Coron et al (2012) per each catchment. Comparison between models can be achieved by comparing how their performance is affected by the relative change in the climate variable. An example of the NAM and PDM models for the sub-catchment W08SAMRON000 (Dijle/Zennebekken) is shown in Figure 20. Notice that for this preliminary results the cloud is not yet center based on climate variable, and there is only one sub-catchment included therefore the box plots are not yet calculated. In Figure 20 negative percentages of precipitation index (Nval/NCal-1) indicates a reduction in the precipitation with respect to the reference period . In the vertical axis, negative percentages of model robustness means that the errors are increasing as the precipitation change and positive shows and improvement. Althoug the example in Figure 20 is a basic evaluation, it shows that In general the trend is similar to that expected; that is, that the model performance is reduced as the climate input is deviating further from the period used for calibration.

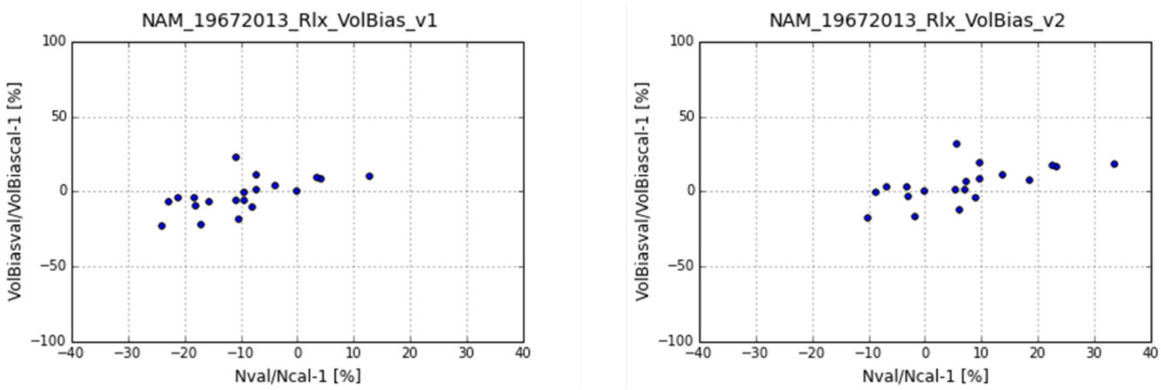


Figure 19 – Comparison of the Peformance of Models with Difference Reference Period for one sub-cathment

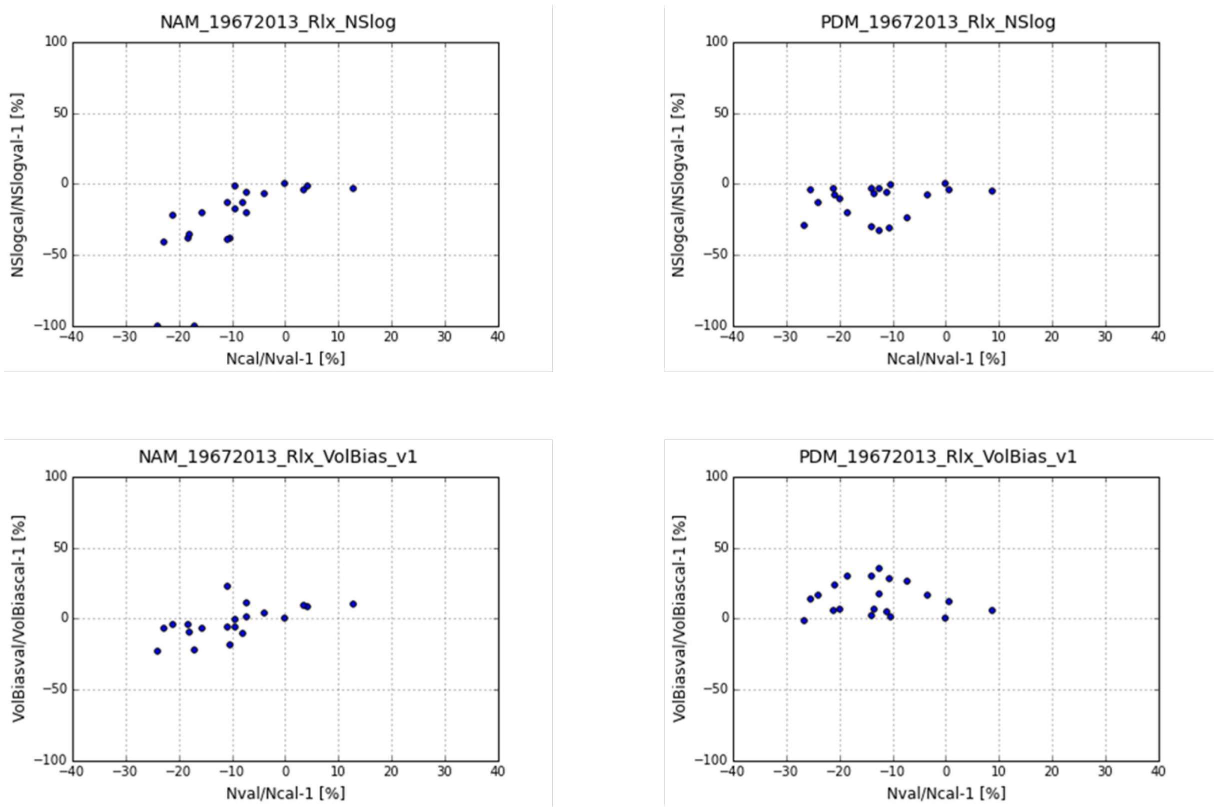


Figure 20 – Comparison of the Peformance of NAM and PDM Models for the sub-catchment W08SAMRON000 (Dijle/Zennebekken)

Validation test using the relative variation on projected climate and hydrological conditions

For this approach it is needed the projected climate conditions. For water allocation project, we will used the existing climate change scenarios for Belgium. They are based upon four green-house gas emission scenarios (A1B, A2, B1 and B2) developed by the IPCC (2001) and more than 50 runs with global and regional climate models for Belgium. The results from these runs were statistically analysed and a range of change factors (perturbation factors) from the reference period 1961–1990 till the future period 2071–2100 derived per month and in function of the return period. The perturbation tool of Ntegeka and Willems (2009) makes it possible to translate observed time series (rainfall, Evapotranspiration) into future time series for three climate change scenarios: High scenario: wet winters and dry summers; Mean and Low scenario: dry winters and dry summers. Considering that for water allocation the critical conditions occur during dry periods the scenario selected to assess the model is the Low one.

In principle the approach will follow the estimation of the relative change of the ratio Q_{90}/Q_{50}. whereQ_{90} and Q_{50} are the 90th and 50th percentiles of the flow duration curve, respectively. The Q_{90}/Q_{50} ratio represents the difference between low flows and medium flows, thus indicating the severity of low flows. Low flow frequency distributions can be also used to calculate the relative change. The main idea is to adapt the comparison of the models proposed by van Steenbergen and Willems (2012). Therefore, low flow statistics could be used. The Water Laboratorium must provide the time series of climate variables (i.e. precipiation and evapotranspiration) for each sub-catchment. More detail on the approach will be described when the data is available and pre-liminary test can be carried out.

Appendix 2 Methodological note on the climate perturbation tool

Methodological note on the climate perturbation tool

Presentation of the tool

The "climate perturbation tool" is the result of the CCI-HYDR project. It was developed and designed by a multi-disciplinary team from the Katholieke Universiteit Leuven and the Royal Meteorological Institute (RMI) of Belgium (Ntegeka, et al., 2011).

The project's purpose was to explore the climate change impact on the risk of hydrological extremes along rivers and urban drainage systems in Belgium. The pertubation tool allows the user to predict future meteorological data based on current series of precipitation, evapotranspiration and temperature. In fact, it applies perturbation on the rainfall and the evapotranspiration observed series (ETo) depending on different scenarios and the period.

The perturbation steps

The regime of precipitation is modified by changing the frequency of rain storms or the rainfall intensity (depending on the month in the year, and on the return period or storm frequency). The perturbation of the ETo series concerns intensities, depending on the month and the return period.

For the precipitation series, the number of wet days to be added or removed is determined for the "high", "mean" and "low" wet day frequency perturbations. Wet days are selected from the set of empiricial (observed) wet days. Then, this selection is removed from the original series and replaced by dry days or wetter days (Vansteenkiste, 2012). This procedure is followed by a perturbation of rainfall intensities. Perturbation is applied by calculating for each wet day in the historical series the exceedance probability, and multiplying the intensity of this wet day with the corresponding quantile-perturbation factor as projected by the regional climate model (Ntegeka, 2011).

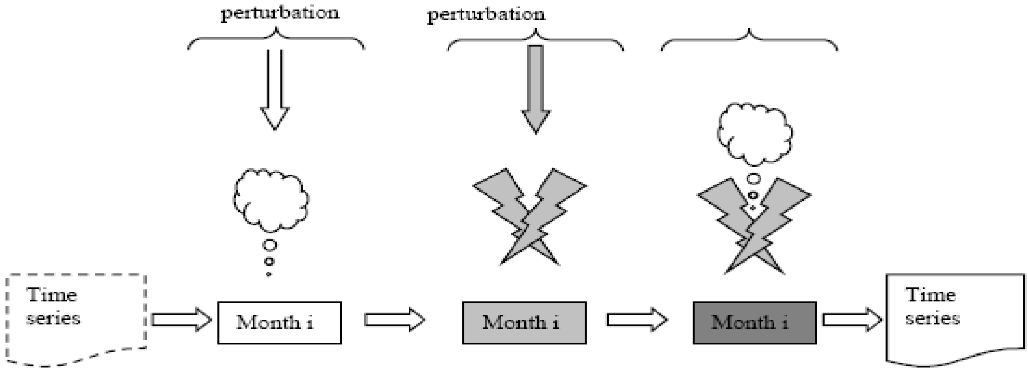


Figure 21 – Rainfall series perturbations (Ntegeka, et al., 2011)

When it comes to the ETo, it is expected to increase for all seasons, reaching about 50% in the winter and autumn. However, its variability is considered low, and a quantile-based perturbation is used. To perturb the observed series, a point is located, above which the perturbation remains fairly constant. Then, all the perturbation above this point are averaged. Afterwards, each ETo value is multiplied by an average factor. This procedure is valid for the whole year and is repeated for each regional climate model (Ntegeka, 2011).

Before developping the perturbation tool, different climate change scenarios were tested from multiple regional and global climate models. It was chosen at the end to adapt the PRUDENCE regional climate model adjusted by scaling factors from the global climate model considered from the 4th Assessment Report of the IPCC (United Nations Intergovernmental Panel on Climate Change) (Ntegeka, et al., 2011).

The climate scenarios

In Belgium and Flanders, the climate scenarios signpost an increase of the annual average temperature by 0.7 to 7.2°C over 100 years. Three main scenarios were considered as the most likely to occur: “high” (summer or winter), “mean” and “low” scenarios, based on their hydrological impact. They can also be referred to as “wet” (summer or winter), “mild” and “dry”. For the “high” (or “wet”) impact scenario, a winter and a summer scenario are distinguished to address future wet conditions in winter and summer respectively. However, evapotranspiration is at its highest level for both these scenarios, unlike the other scenarios. The “mean” scenario represents the expected average scenario (mean flow impact). The low scenario reflects the most pessimistic change in the low flow situation, with dry winters and dry summers (strongest low flow impact) (Vansteenkiste, 2012).

Table 1 below summarizes the observations of the possible climate scenarios. The rate of the possible change for the different hydrological and climatological factors is presented for an anticipation up to 100 years. These values are taken into consideration for perturbing time series with the perturbation tool.

In Table 2, the different combination of the level of perturbation to obtain each scenario is explained. Clearly, the evapotranspiration is at its highest level in the High winter scenario while the precipitation is at its lowest value during the low scenario. What is also remarkable, the choice of adapting mean precipitation for the spring and autumn during all the seasons and also for most of the seasons for the evapotranspiration. Overall, there is a correlation between the adapted methodology presented in table 2 and the researches results presented in table one.

Table 27 – Overview of possible climate change for Flanders and Belgium according to the low, medium and high climate scenario, over 30, 50 and 100 years (MIRA, et al., 2015)

Change	Time horizon	Climate scenarios			Additional info
		low	medium	high	
annual average temperature	30	+0.2 °C	+1.1 °C	+2.2 °C	The coast has a mitigating effect on warming, but the effect is small with respect of the expected climate change
	50	+0.3 °C	+1.8 °C	+3.6 °C	
	100	+0.7 °C	+3.7 °C	+7.2 °C	
average number of extremely hot days per year	30	0	5	+19	The number of extremely cold days decreases the most in the Ardennes
	50	0	8	+32	
	100	0	16	+64	
average number of extremely cold days per year	30	0	-2	-10	Winter precipitation increases more along the coast
	50	-1	-4	-17	
	100	-1	-7	-33	
total winter precipitation	30	-0.40%	3%	+11%	Extreme summer precipitation intensities may increase significantly. Spatially, a north south pattern is emerging with greater desiccation in the south
	50	-0.60%	6%	+19%	
	100	-1%	12%	+38%	
total summer precipitation	30	-16%	-4 %	+5%	
	50	-26%	-7 %	+9%	
	100	-52%	-15 %	+18%	
numbers of wet days in winter	30	-1%	+0.5 %	+2%	
	50	-2%	+0.8 %	+4%	
	100	-5%	+1.5 %	+8%	
numbers of wet days in summer	30	-12%	-5 %	+1%	
	50	-21%	-8 %	+2%	
	100	-41%	-15 %	+4%	
total potential evapotranspiration in winter	30	+0.50%	+3 %	+11%	
	50	+1%	+6 %	+18%	
	100	+2%	+12 %	+35%	
total potential evapotranspiration in summer	30	0.50%	+5 %	+14%	
	50	1%	+8 %	+23%	
	100	2%	+17 %	+47%	
daily average speed in winter	30	-8%	0 %	+3%	
	50	-14%	-0.5 %	+6%	
	100	-28%	-1 %	+11%	

To predict the effect of the driest conditions in the future in Belgium, a scenario of climate change has to be chosen among the three main ones. Based on Table 1 and Table 2, major candidates are the high winter and the low scenarios. The first one has the specificity to involve the driest summer of all of the scenarios and the higher evapotranspiration for most of them. However, the low scenario has the lowest rate of precipitation during all the seasons.

Table 28 – Seasonal correlations and scenario definitions (Ntegeka, et al., 2011)

Season	Eto	Precipitation	Scenario
Winter	High	High	High Winter
Spring	Mean	Mean	
Summer	High	Low	
Autumn	Mean	Mean	
Winter	Mean	Low	High Summer
Spring	Mean	Mean	
Summer	Low	High	
Autumn	Low	Mean	
Winter	Mean	Mean	Mean
Spring	Mean	Mean	
Summer	Mean	Mean	
Autumn	Mean	Mean	
Winter	Low	Low	Low
Spring	Low	Mean	
Summer	Low	Low	
Autumn	Low	Mean	

Comparison between the scenarios

The next step is to investigate which scenario is the best to model the Scheldt basin regarding water allocation.

There are mainly two major candidates that can father the driest hydrological conditions: the “high winter” scenario and the “low” scenario. The “low” scenario is quite logically considered as a candidate, due to the fact that it is considered as a dry one in the bibliographical reports and, also, it has the lowest precipitations rate predictions. However, the evapotranspiration rate along the year is at its lowest level for this dry scenario, whereas it is maximum in the high winter scenario. This causes in fact a dryer summer in the “high winter” scenario.

Since the tool instructions were to investigate all the possible scenarios before taking a final decision on which one to use, we followed these instructions and made a comparison between those two expected-dry scenarios.

For that aim, we chose four random catchments, situated in different part of Belgium and of different scales. These catchments are: F05LEI386999, V01HAN488180, V08BAR111370, W11OUR5805.

Since the evapotranspiration in the low scenario is designed as low, we added an extra personalized scenario to the comparative analysis, where the precipitation during all the year is low (even autumn and spring) and the evapotranspiration is high throughout all the seasons (considering also autumn and spring). This “extremely low” scenario was created given the fact that the “low” scenario might not generate such dry conditions after all, due to limited evapotranspiration. This extra scenario is analyzed for the sake of comparison and might be beyond the scope of state-of-the-art downscaled climate models for Belgium however. This additional scenario is called “dry” in the following.

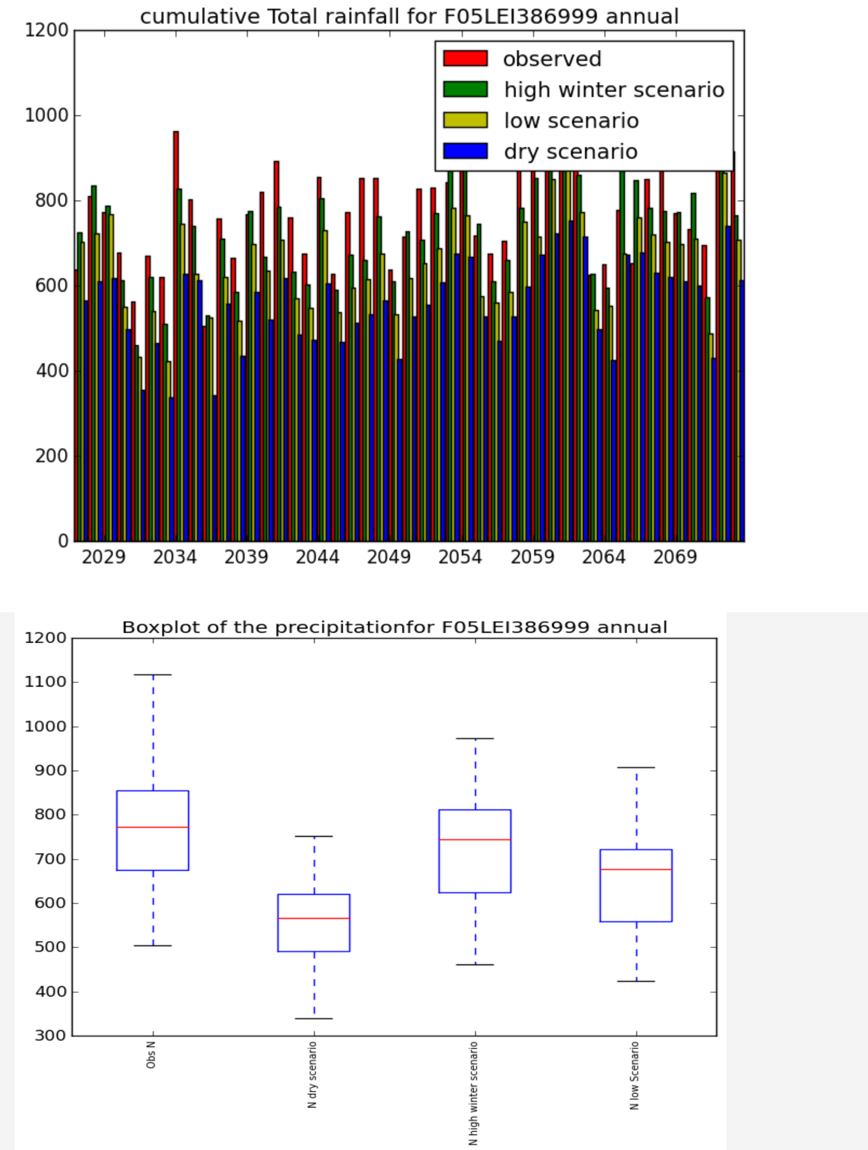
According to the table above, this artificial “dry” scenario can be described as follow:

Season	Eto	Precipitation	Scenario
Winter	High	Low	Dry
Spring	High	Low	
Summer	High	Low	
Autumn	High	Low	

Comparative graphs are shown below for the different climate and hydrological variables, with examples out of the four test catchments.

Annual cumulative rainfall

The figures below show the case of the Leie catchment F05LEI386999:

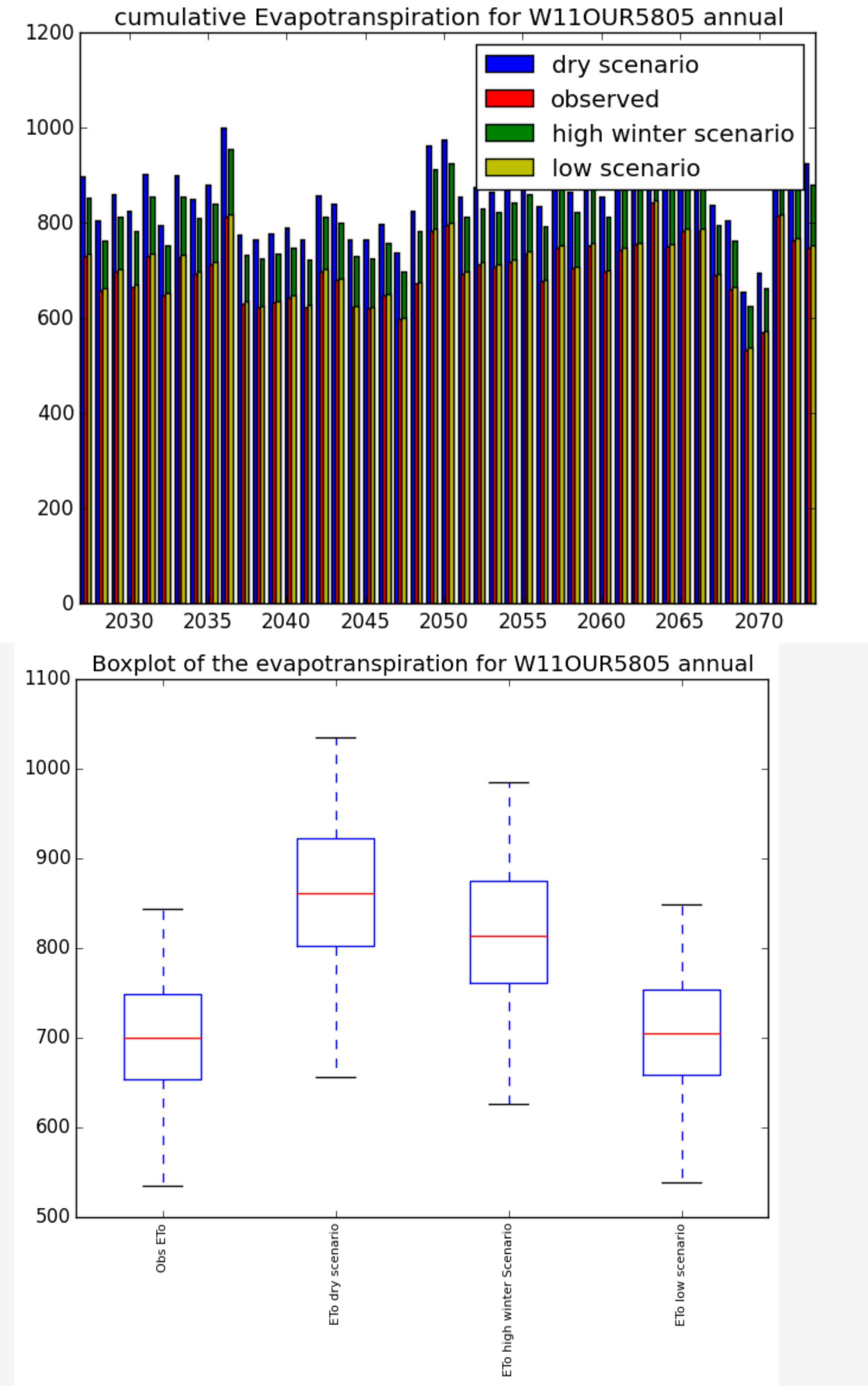


Overall, the three scenarios have a range of precipitation lower than the historic one. The observed median is around 780 mm. Precipitation is slightly lower in the “high winter” scenario, with a median of 750 mm.
As expected by design, the “dry” scenario has the lowest annual cumulative precipitation (with a median around 570 mm), while the “low” scenario stands in an intermediate position (with a median precipitation of 680 mm). Variability is the lowest for the “dry” scenario.

Figure 22 – Comparison of observed and perturbed rainfall in catchment F05LEI386999.
Top: Histogram of the annual cumulative rainfall. Bottom: Box-plots of the annual cumulative rainfall

Annual cumulative evapotranspiration

The figures below show the case of the Ourthe catchment W11OUR5805.

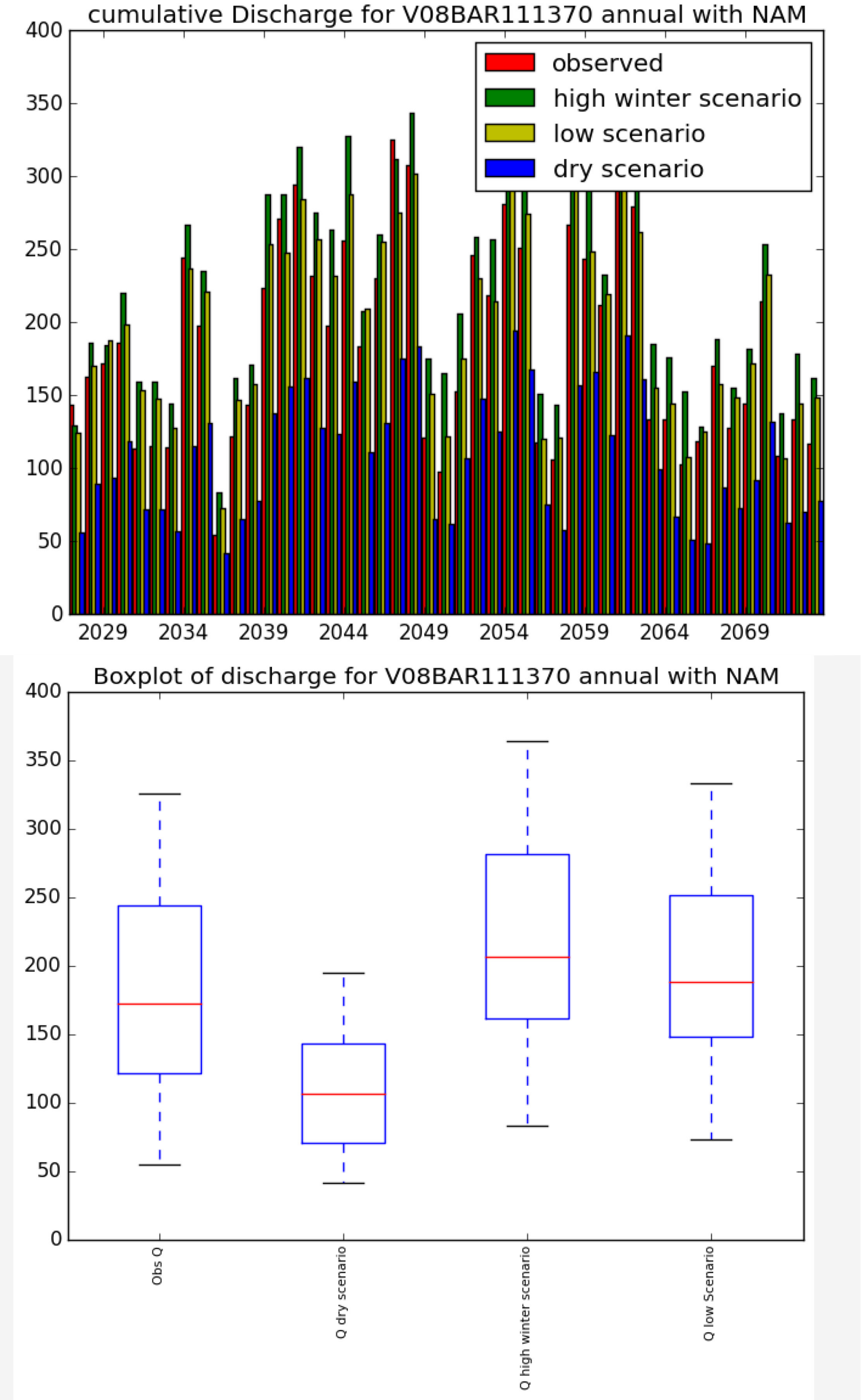


For this catchment, as well for the others, the perturbed evapotranspiration is higher than the observed one for all scenarios, though not significantly for the "low" scenario where the median ETo remains practically unchanged.
As expected from the tool design, evapotranspiration is higher throughout the year for the driest scenario. When we compare the "low" and the "high winter" scenarios, ETo ranges much higher for the latter, which indeed alternates high and medium ETo instead of low ETo (Table 28).

Figure 23 – Comparison of observed and perturbed evapotranspiration in catchment W11OUR5805.
Top: Histogram of the annual cumulative evapotranspiration. Bottom: Box-plots of the annual cumulative evapotranspiration

Annual cumulative flow comparison (with NAM)

The figures below show the case of the Barebeek V08BAR111370



Naturally, the “dry” scenario generates the lowest annual cumulative discharge of all.

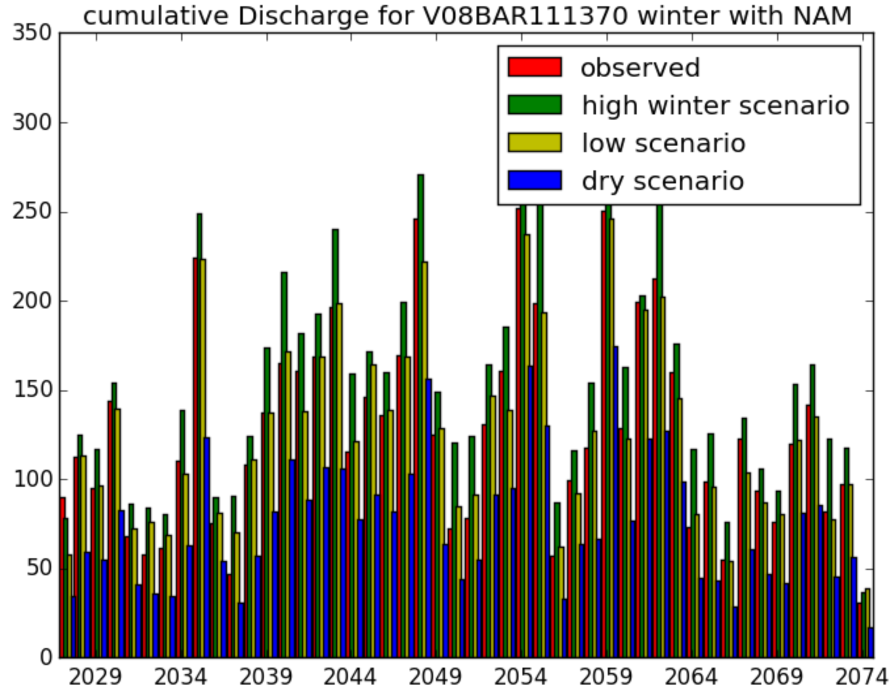
For the “high winter” and the “low” scenarios, cumulative flows often exceed the observed values, as reflected by the medians as well.

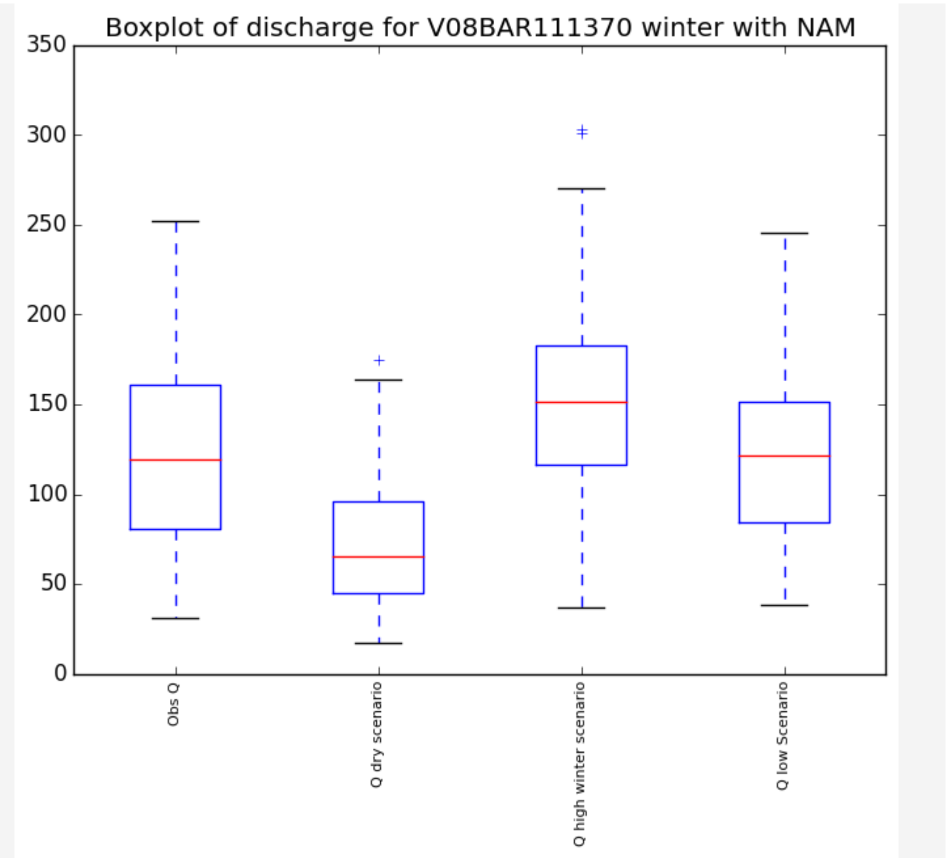
To see the seasonal effects of the perturbation more explicitly, we divided the series into summer (May to October) and winter periods (November to April).

Figure 24 – Comparison of the annual cumulative flow for the catchment V08BAR111370.
Top: Histogram ; Bottom: Box-plots

Cumulative flow comparison in winter (with NAM)

Flow is cumulated over the months from November to April, hereafter named "winter" period. The figures below show the case of the Barebeek V08BAR111370





Because seasonal variability plays an important role in the perturbation tool, winter flows are analyzed separately.

As reported in the literature of the perturbation tool, the "high winter" scenario reached the highest cumulative flow values in winter among all scenarios. Also, the distribution includes some extreme years, with very high flow in the winter months.

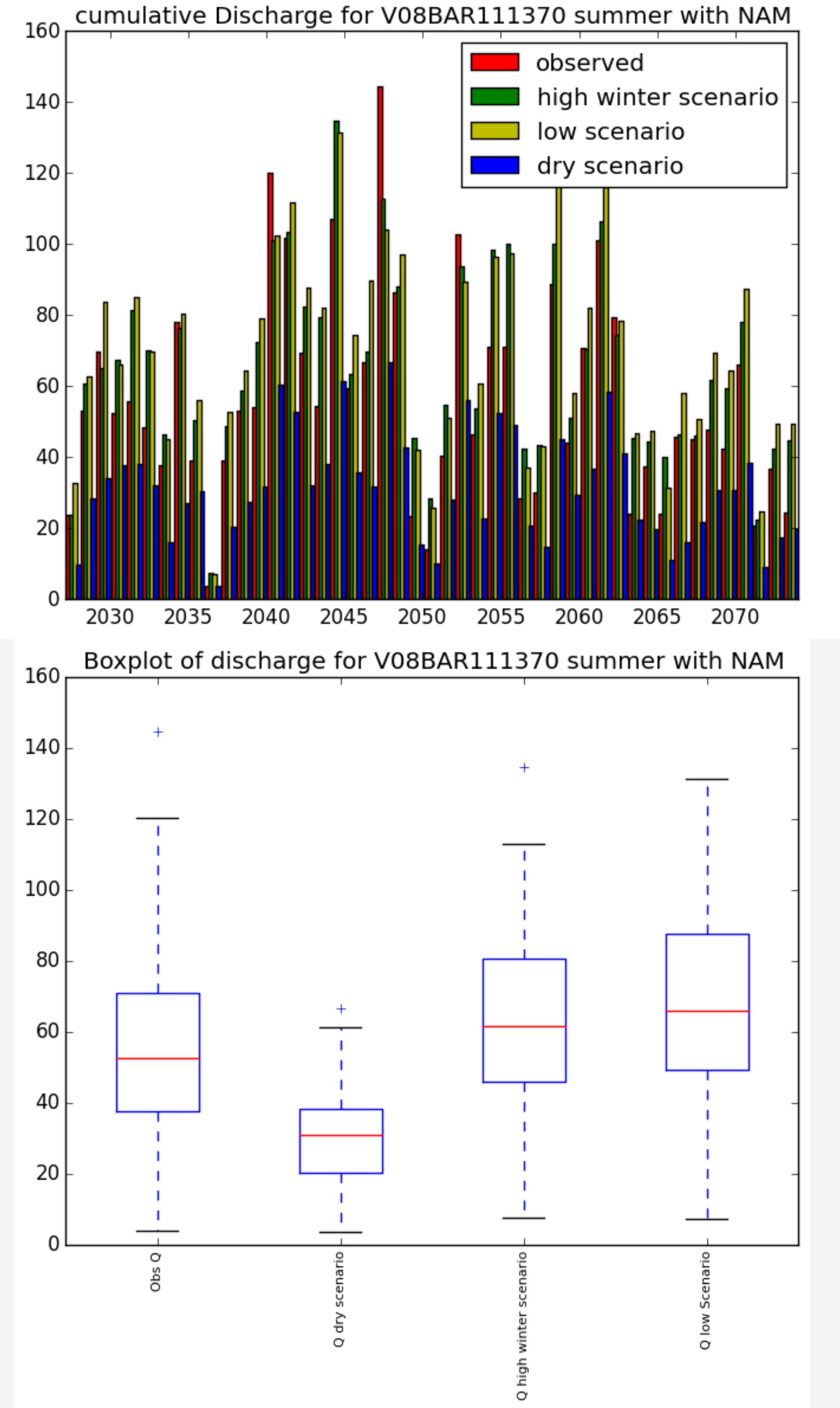
When it comes to the "low" scenario, winter cumulative flows are similar to the historical values, with a bit less variability.

The "dry" scenario naturally leads to the lowest winter cumulative flows, and lowest variability.

Figure 25 – Comparison of cumulative flow in the winter months, per year, for the catchment V08BAR111370.
Top: Histogram; Bottom: Box-plots

Cumulative flow comparison in summer (with NAM model)

Because seasonal variability plays an important role in the perturbation tool, summer flows are now compared. Flow is cumulated over the months from May to October, hereafter named "summer" period. The figures below show the case of the Barebeek V08BAR111370



On most years, the cumulative flow during the summer months is higher for the "low" scenario than for all other scenarios. This slightly surprising conclusion comes from the fact that, unlike the other scenarios, the evapotranspiration is taken low for this scenario, but with similar precipitation.

The median cumulative flow (in summer) for the "high winter" and the "low" scenarios is higher than the median simulated from historical unperturbed series. Also, the "low" scenario is wetter during the summer, with many years quite high in the distribution range. The likelihood of having summers drier than average looks similar however, between "high winter" and "low" scenarios, but possible lower than historical conditions.

In the context of planning for water shortages, only the "dry" scenario generates drastically drier summers.

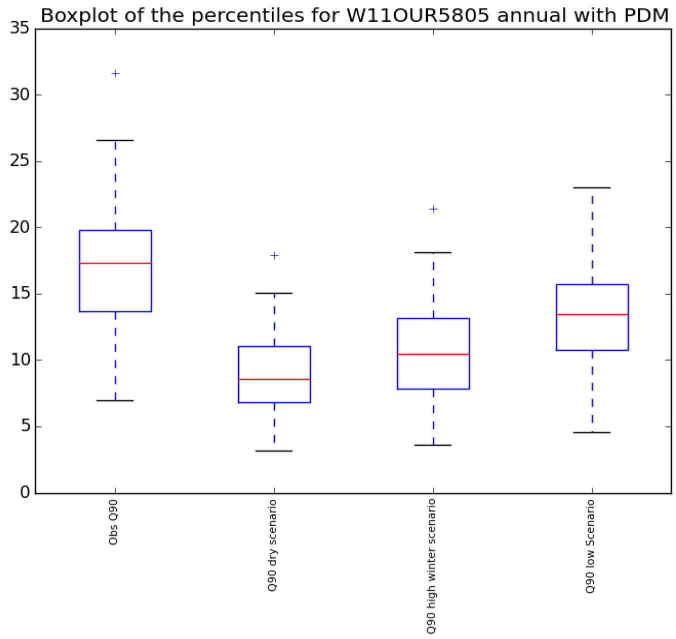
Figure 26 – Comparison of cumulative flow in the summer months, per year, for the catchment V08BAR111370.
Top: Histogram; Bottom: Box-plots

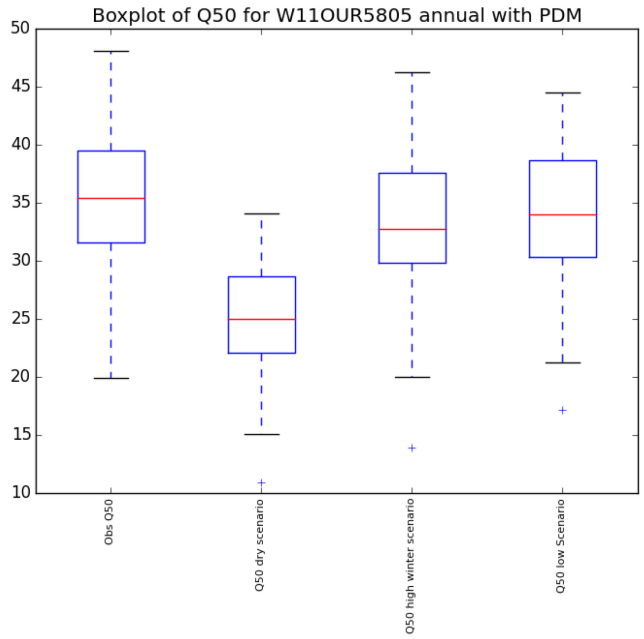
Percentiles comparison (with the PDM model)

The next step is to investigate the low flow statistical indicators calculated on daily series:

- exceedance "dry" percentile Q_{90} : the flow exceeded 90% of the time (equivalent to the non-exceedance percentile, or "wet" Q_{10})
- median Q_{50} : flow exceeded 50% of the time.

Q_{90} is an indicator for the low flow season and Q_{50} for the high flow season. These characteristic flows are calculated on each year and the distribution of all annual Q_{90} and Q_{50}, as well as their ration Q_{90}/Q_{50}, can then be represented as box-plots. The figures below show a comparison for the Barebeek catchment (V08BAR111370).





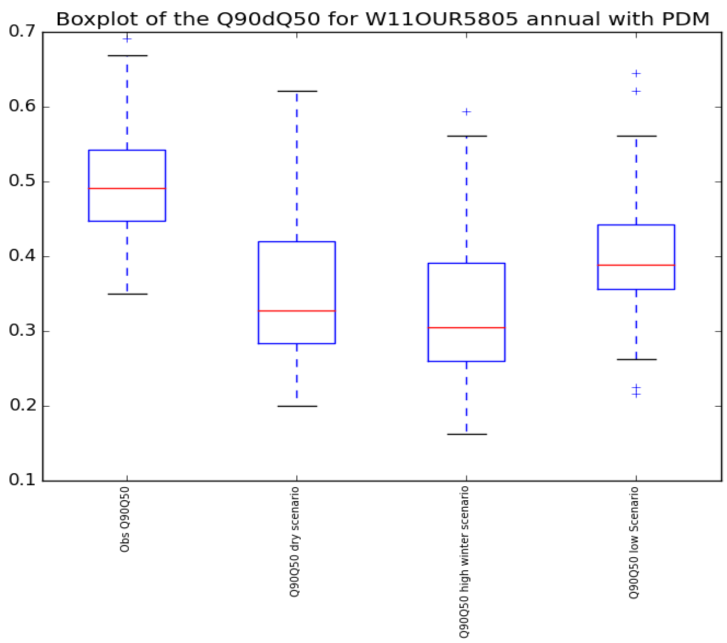


Figure 27 – Top left: Distribution of the percentile Q_{90} for the catchment W11OUR5805 using the PDM model.
Top right: Distribution of the percentile Q_{50} for the catchment W11OUR5805 using the PDM model.
Bottom: Distribution of the percentile Q_{90} /Q_{50} for the catchment W11OUR5805 using the PDM model.

On one hand, the “dry” scenario has the lowest indicators among all the scenarios. On the other hand, when we compare the “high winter” scenario with the “low” scenario, the first one has lower characteristic flows than the other. In fact, the median of annual Q_{90}’s for the “high winter” scenario is around 11 m³/d whereas it is about 14 m³/d for the l””ow” scenario. This means that the “high winter” scenario has drier conditions than the “low” scenario.

Cumulative volumes of total flow comparison (with the VHM model)

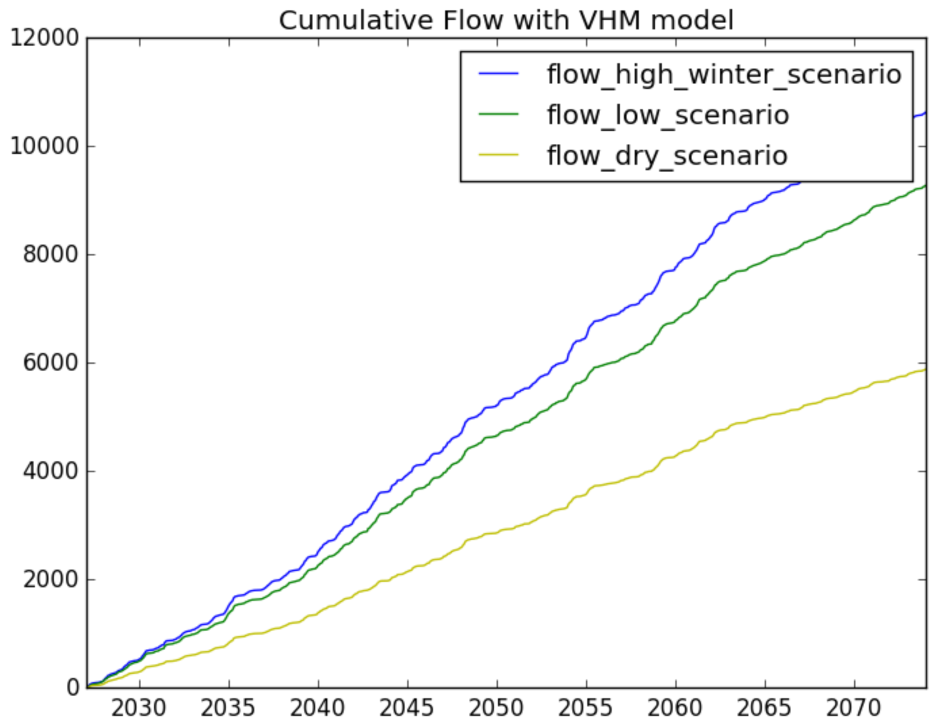


Figure 28 – Cumulative volumes of total low for V08BAR111370 with the VHM model

While the “high winter” scenario seem to present drier hydrological conditions in summer than the “low” scenario, the cumulative volumes of the flow also reach their highest level with the “high winter” scenario. It is thus a more contrasted scenario, with multiple flow peaks in the winter but also critical dry periods in the summer.

Cumulative frequency of the flow comparison (with PDM model)

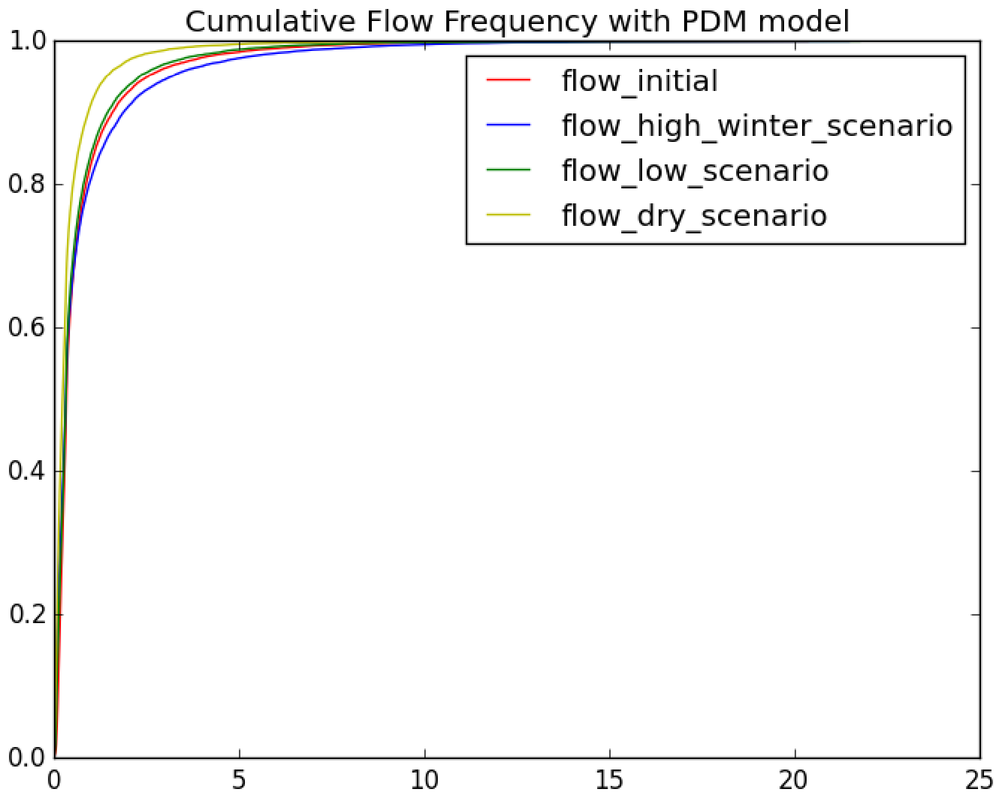


Figure 29 – Cumulative frequency of the flow for the catchment V01HAN488180 using the PDM model

When we compare the range of flows strictly between 0 and 1 m³/d, the cumulative frequency reaches 0.6 for all the main scenarios. It means that the lowest flow has almost the same probability to occur as all the scenarios since the curves are practically superimposed. Between the flow 5 m³/d and 10 m³/d, the curves of the “low” and the “dry” scenarios are above, which means that they have high probability of lower flow comparing to the “high winter” scenario.

Conclusion

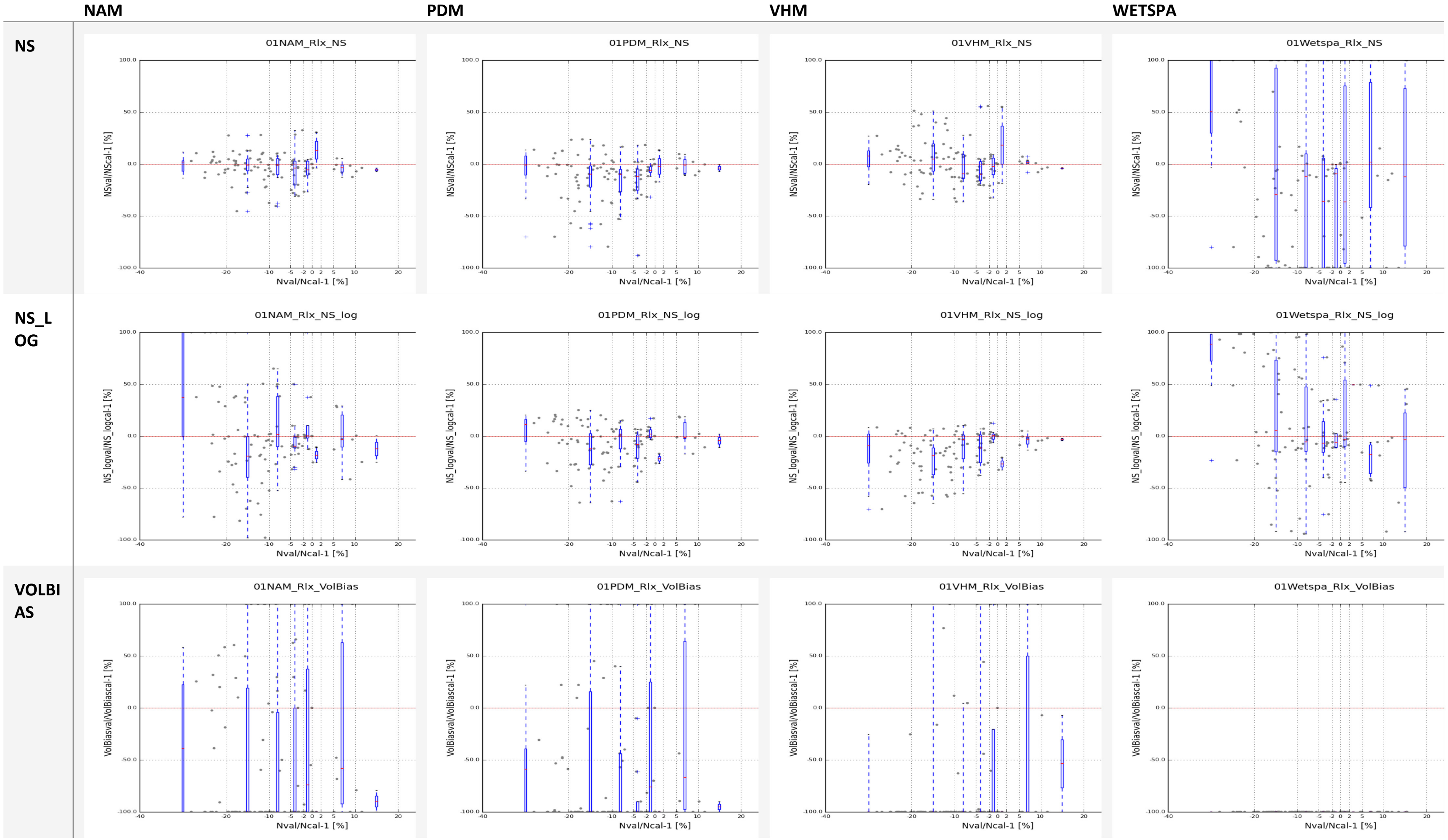
While the “dry” scenario provided the lowest flows among all other scenarios, it should not be used for actual planning because it does not reflect actual climate model downscaling for Belgium.

Now that we have an overall view of all the results, the “high winter” scenario appears to be the most critical for low flow analysis, especiall in the summer. This scenario provides indeed drier condition than the “low” scenario. The evapotranspiration is higher throughout the year and especially the winter which affects the remaining quantity of water. Plus, this scenario presents the driest summer which leads to lower values of the percentiles Q_{90} and Q_{50}. These parameters are generally in a lower range compared to the “low” scenario.

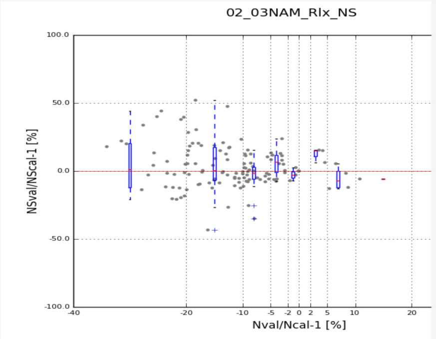
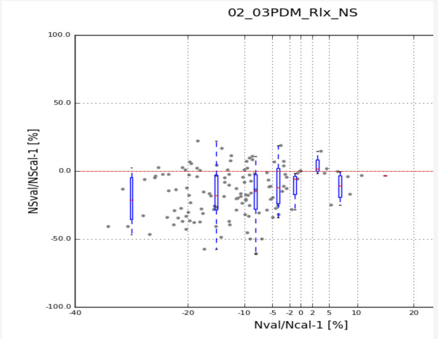
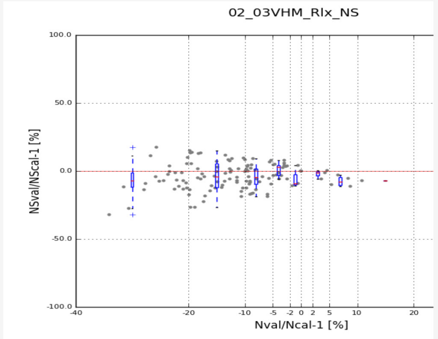
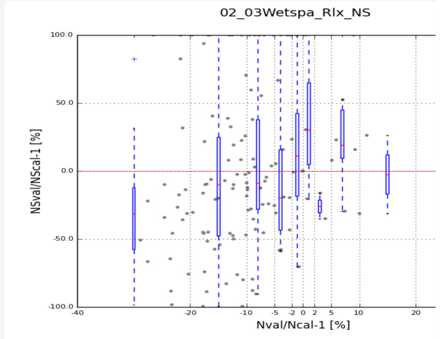
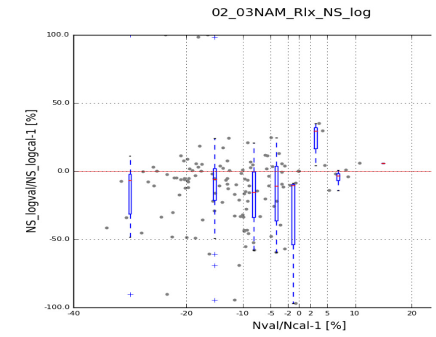
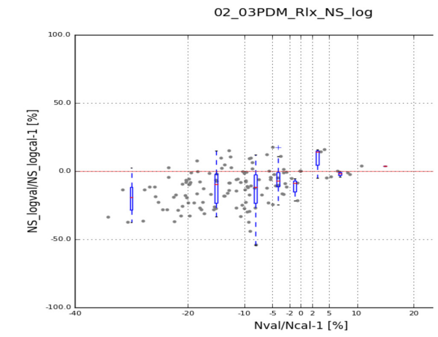
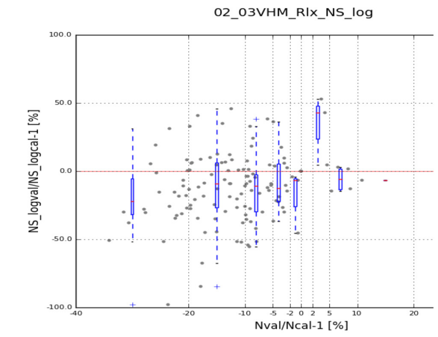
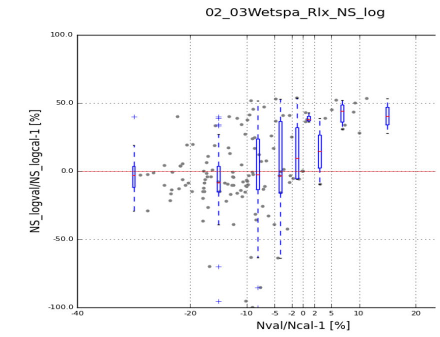
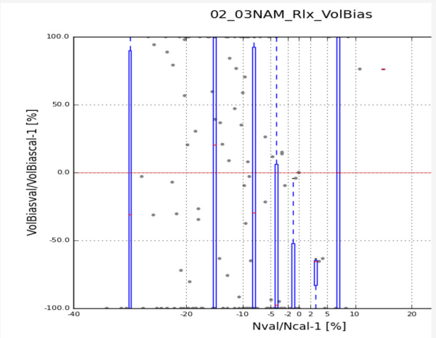
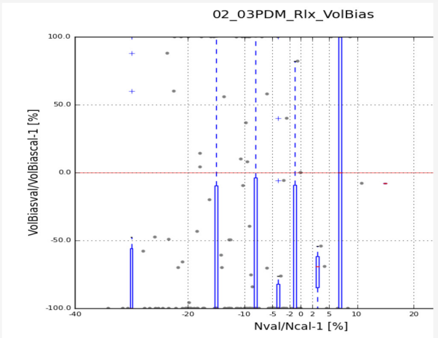
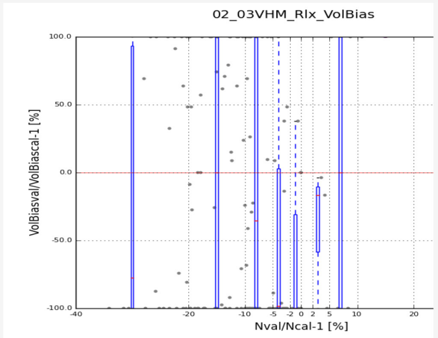
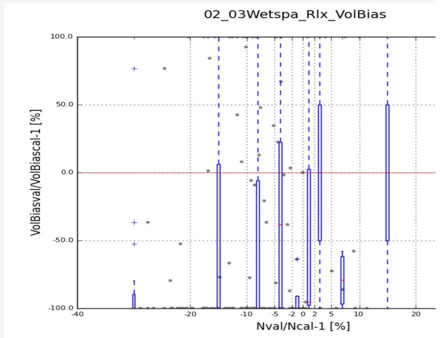
However, the cumulative frequency of the low flows is practically the same for all the scenarios in comparison with the high flows. For the cumulative flow, the “high winter” scenario is the highest in volumes and peaks. The perturbation tool was initially designed for extreme and wet (flooding) future climate and might not be the most appropriate for drought analysis. The contrast appears sufficient though with the “high winter” scenario.

Appendix 3 Differential Split Sample Test: dotty plots

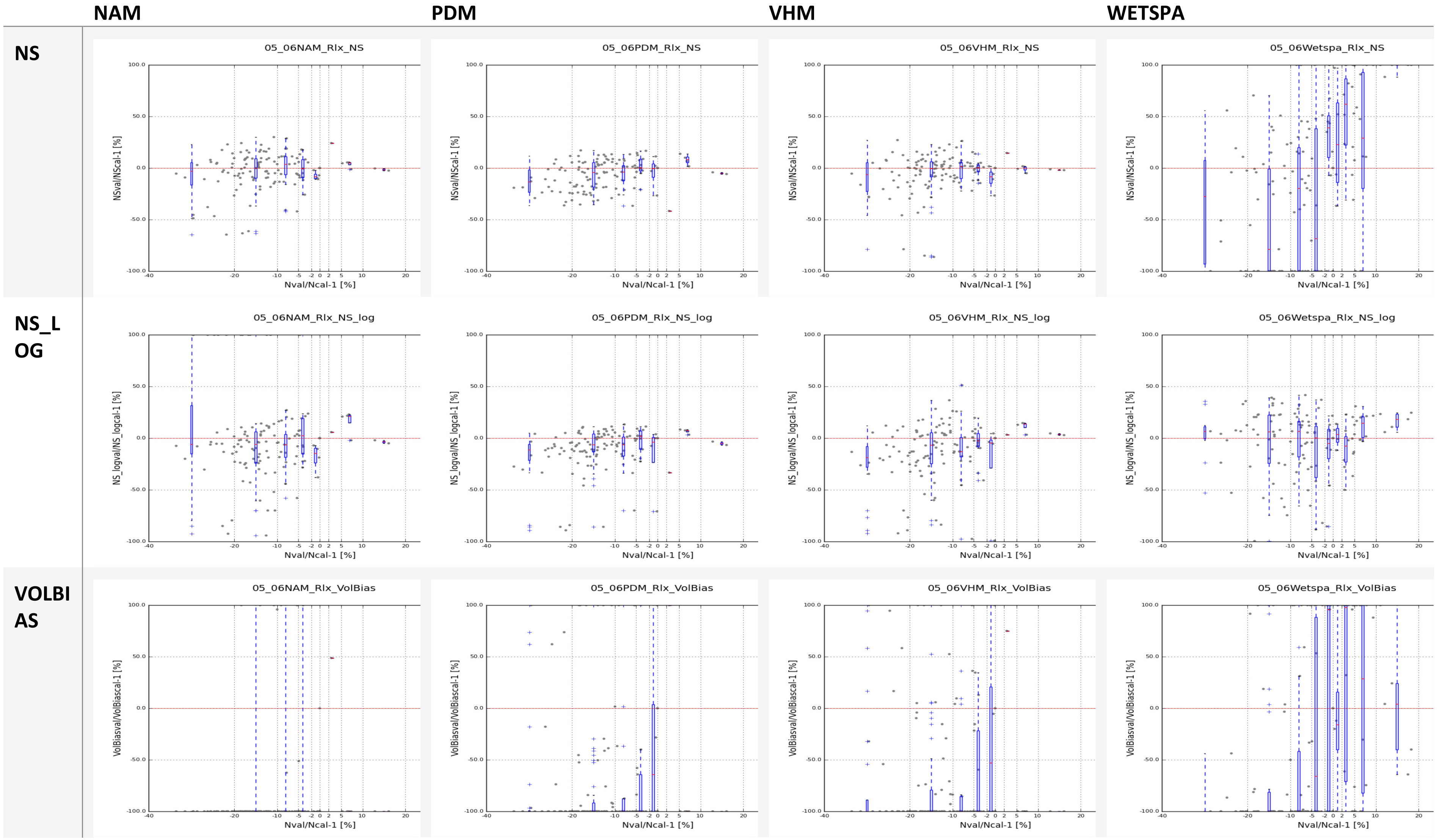
a) Model robustness evaluation for basin(s): [‘01’]



b) Model robustness evaluation for basin(s): ['02', '03']

	NAM	**PDM**	**VHM**	**WETSPA**
NS				
NS_LOG				
VOLBIAS				

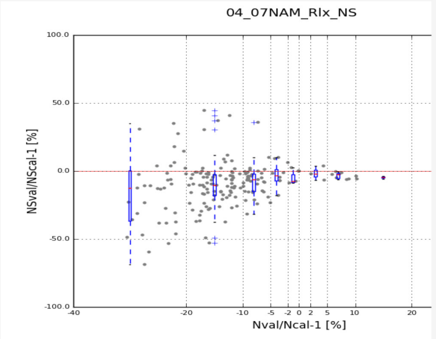
c) Model robustness evaluation for basin(s): ['05', '06']

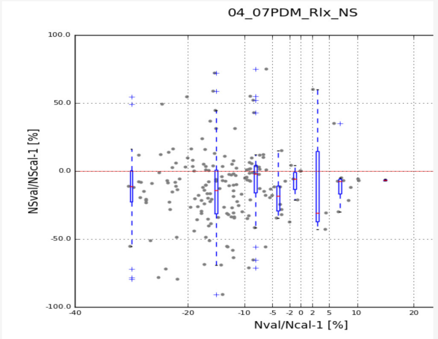


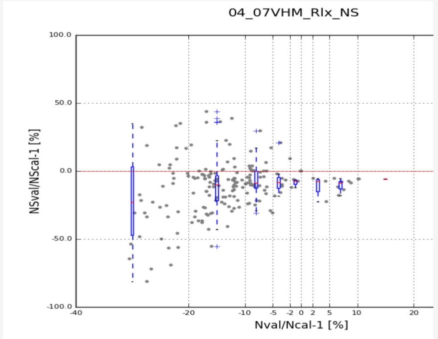
d) Model robustness evaluation for basin(s): ['04', '07']

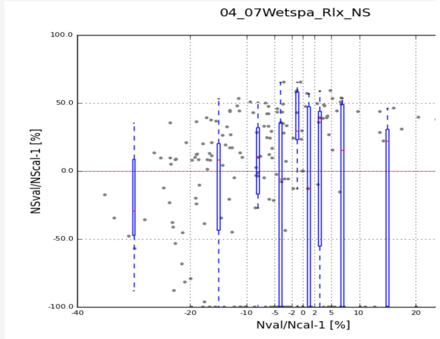
	NAM	PDM	VHM	WETSPA

NS

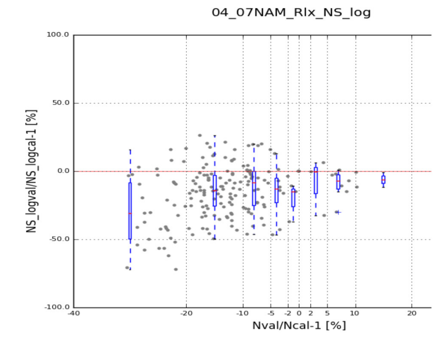


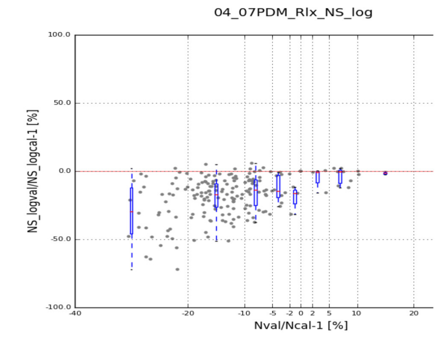


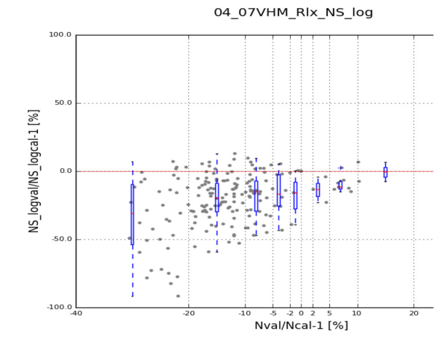


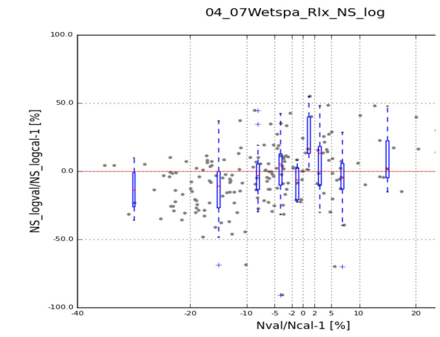


NS_LOG

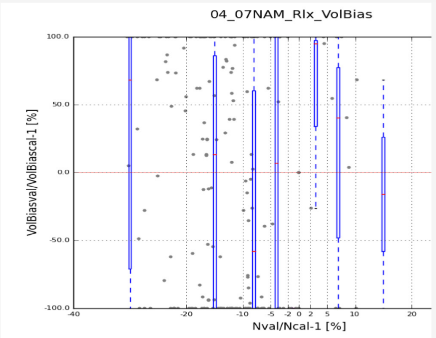


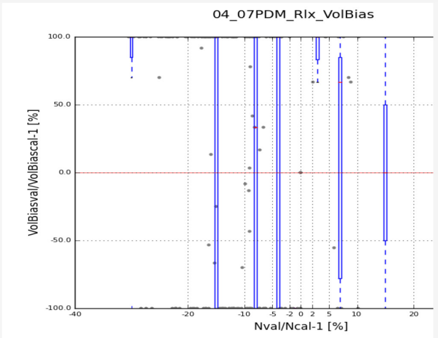


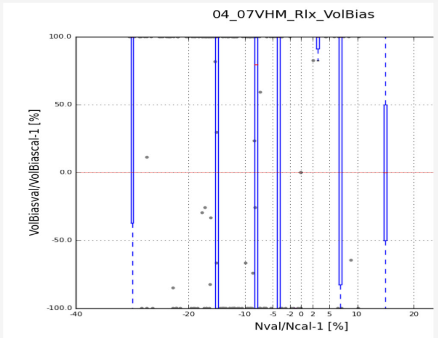


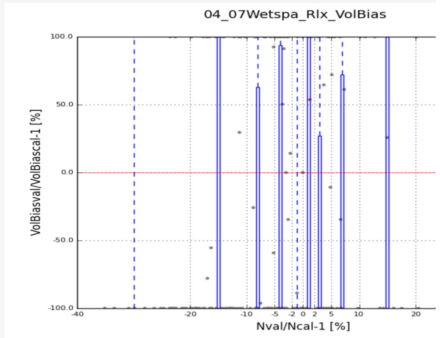


VOLBIAS

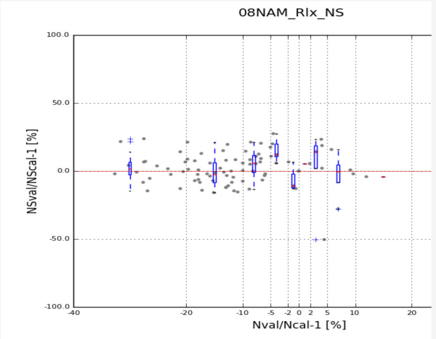
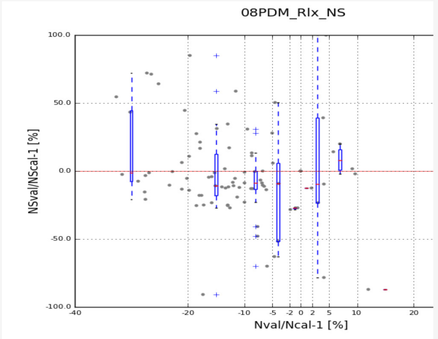
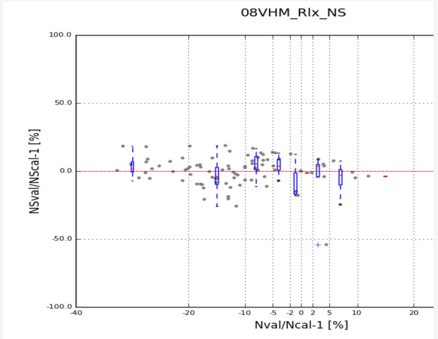
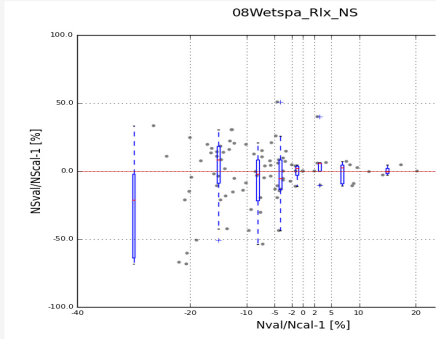
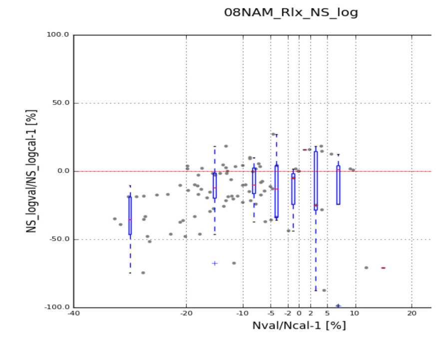
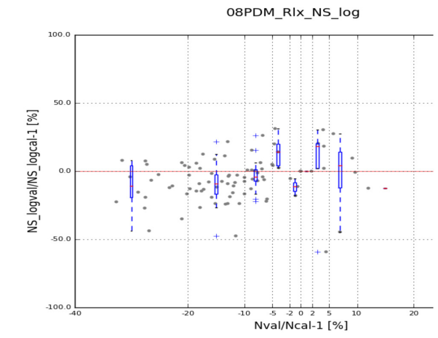
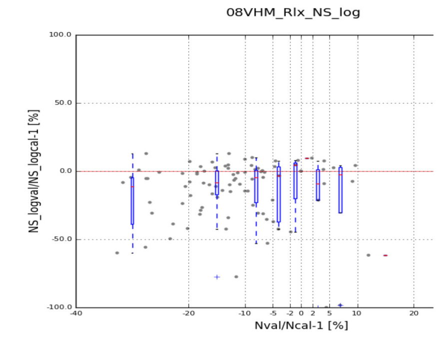
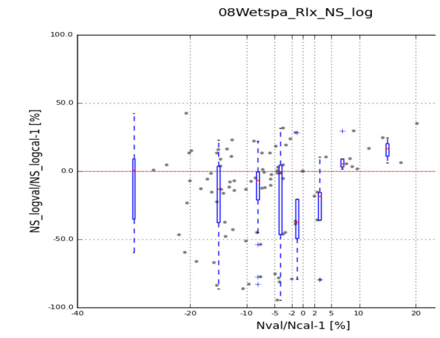
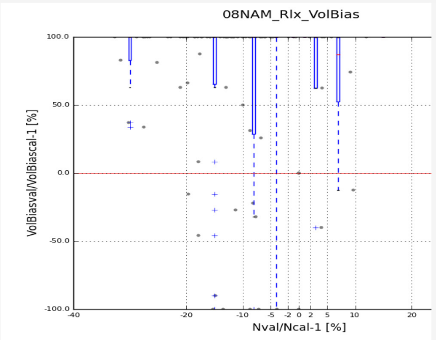
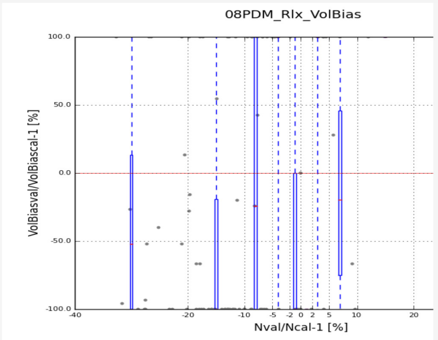
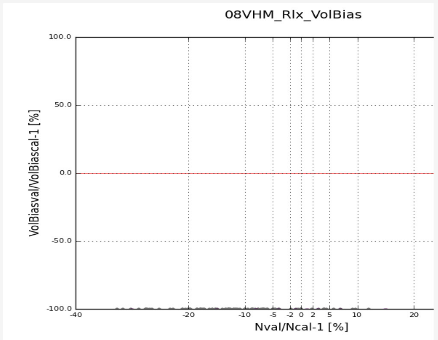
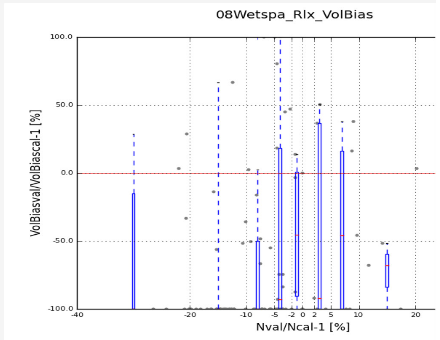




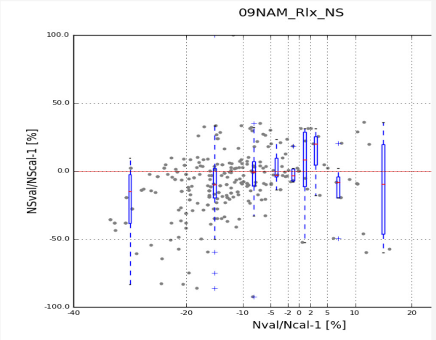
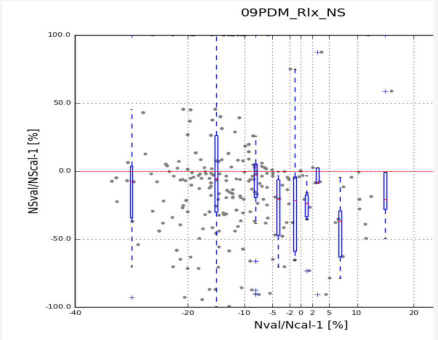
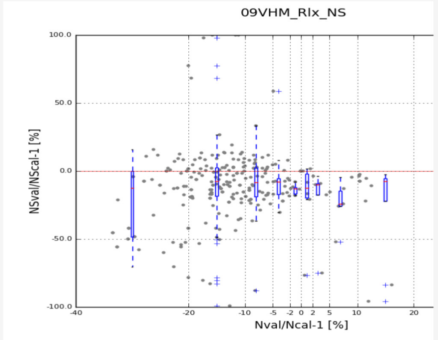
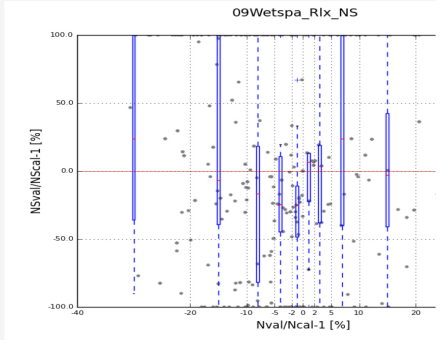
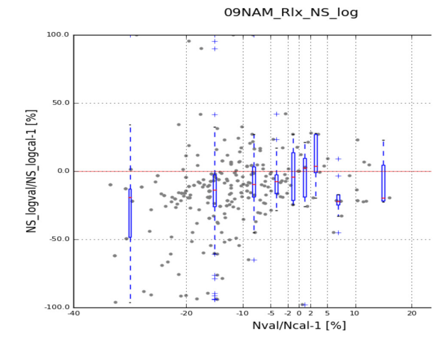
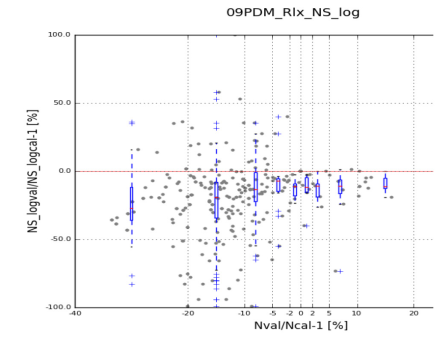
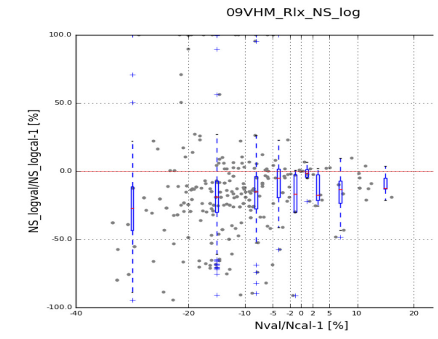
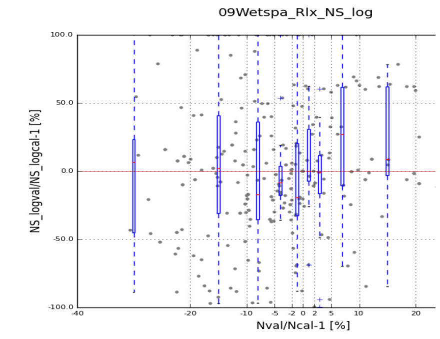
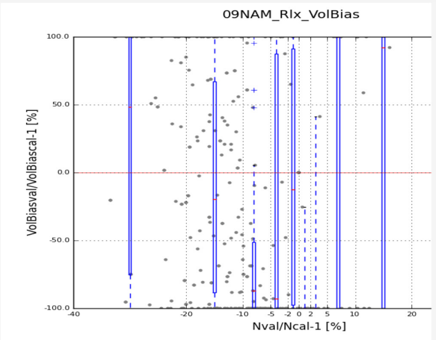
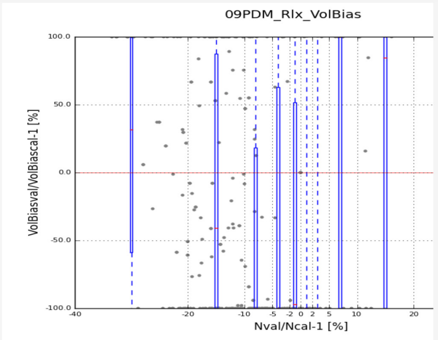
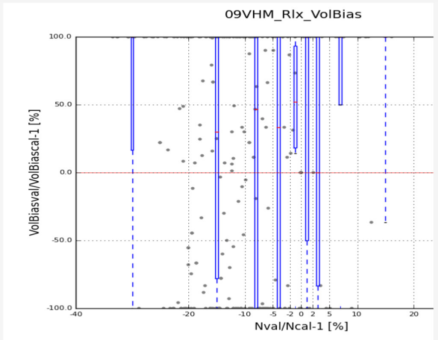
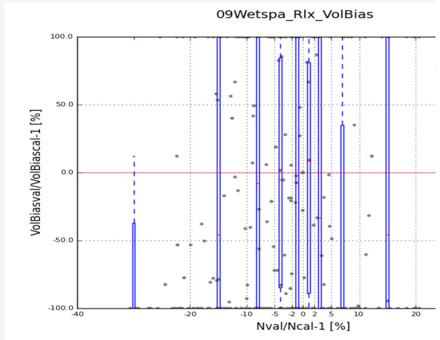




e) Model robustness evaluation for basin(s): ['08']

	NAM	**PDM**	**VHM**	**WETSPA**
NS				
NS_LOG				
VOLBIAS				

f) Model robustness evaluation for basin(s): [‘09’]

	NAM	PDM	VHM	WETSPA
NS				
NS_LOG				
VOLBIAS				

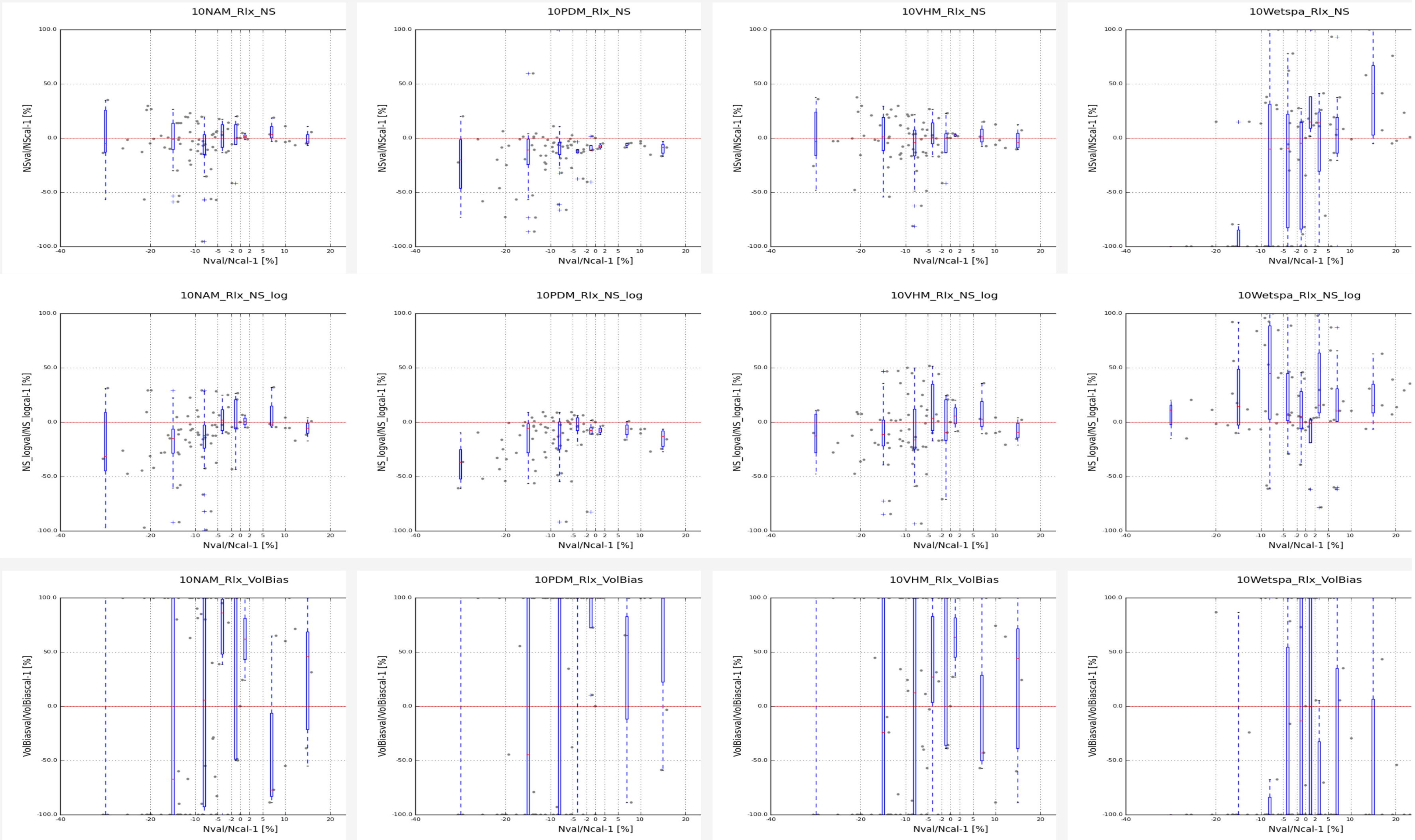
g) Model robustness evaluation for basin(s): [‘10’]

NAM **PDM** **VHM** **WETSPA**

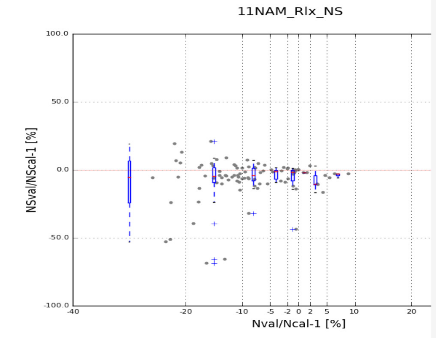
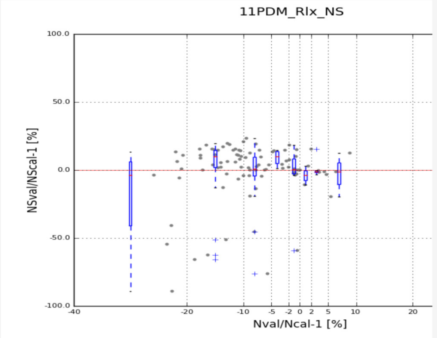
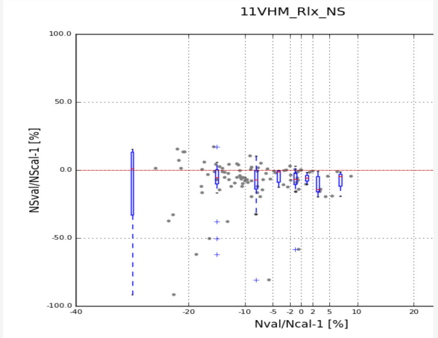
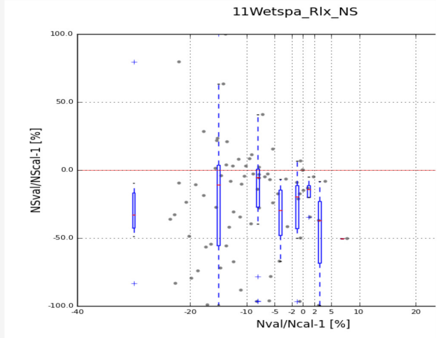
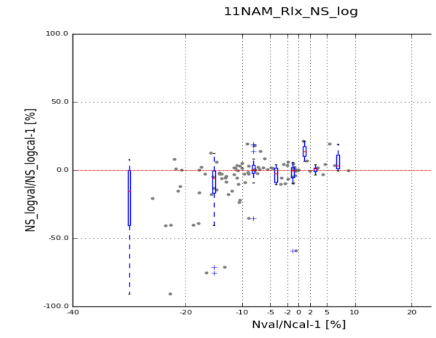
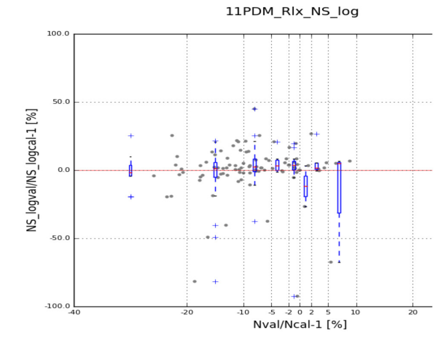
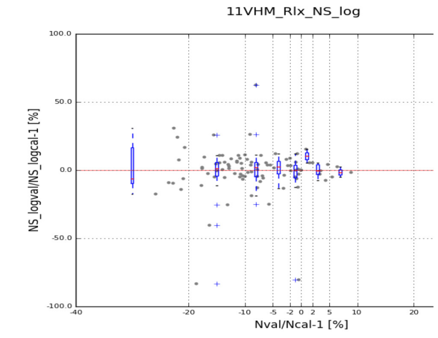
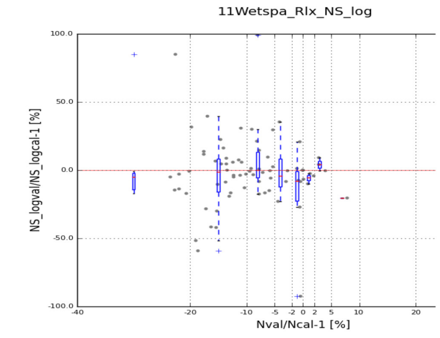
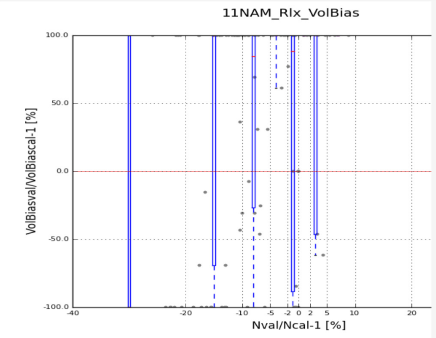
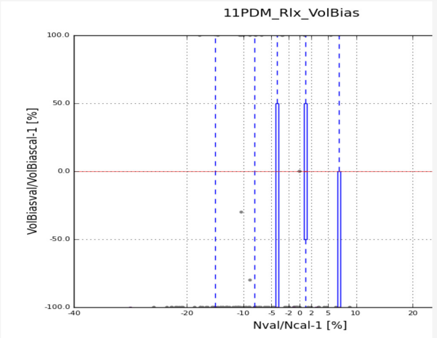
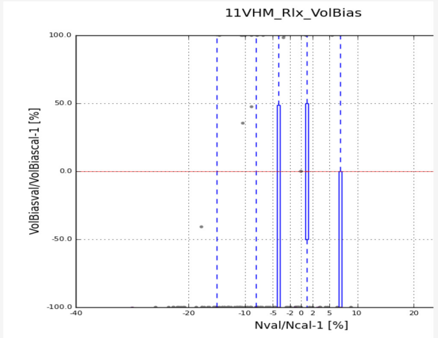
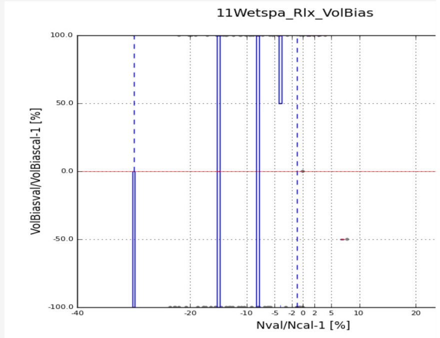
NS

NS_LOG

VOLBIAS



h) Model robustness evaluation for basin(s): [‘11’]

	NAM	PDM	VHM	WETSPA
NS				
NS_LOG				
VOLBIAS				

Appendix 4 Differential Split Sample Test: Box Plots

Note: under parenthesis are indicated the number of catchments analyzed in the basin

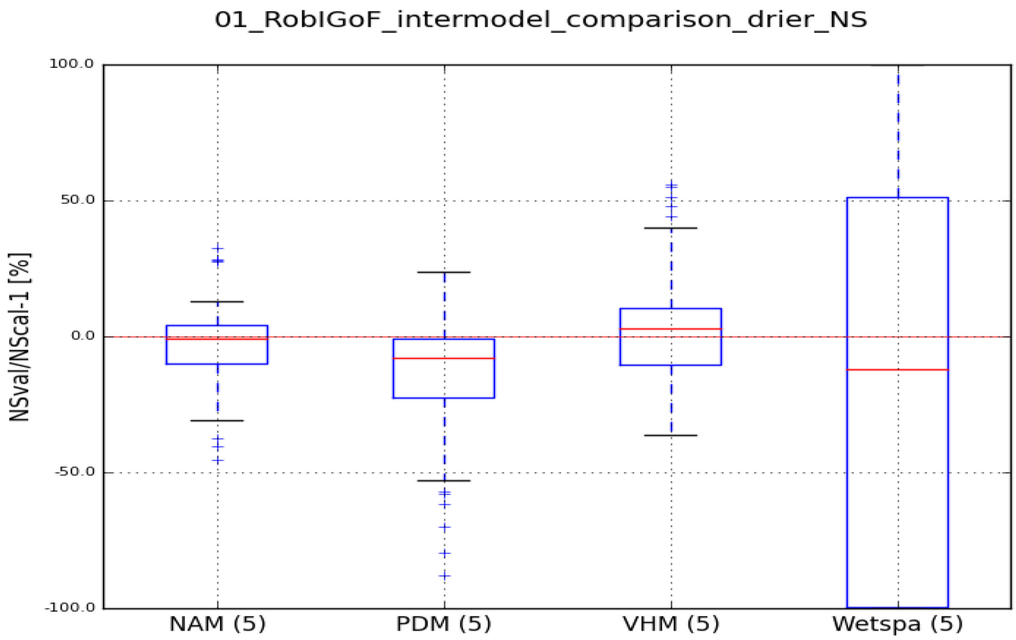


Figure 1 – Relative loss of NS performance against relative drier climate conditions
(IJzer)

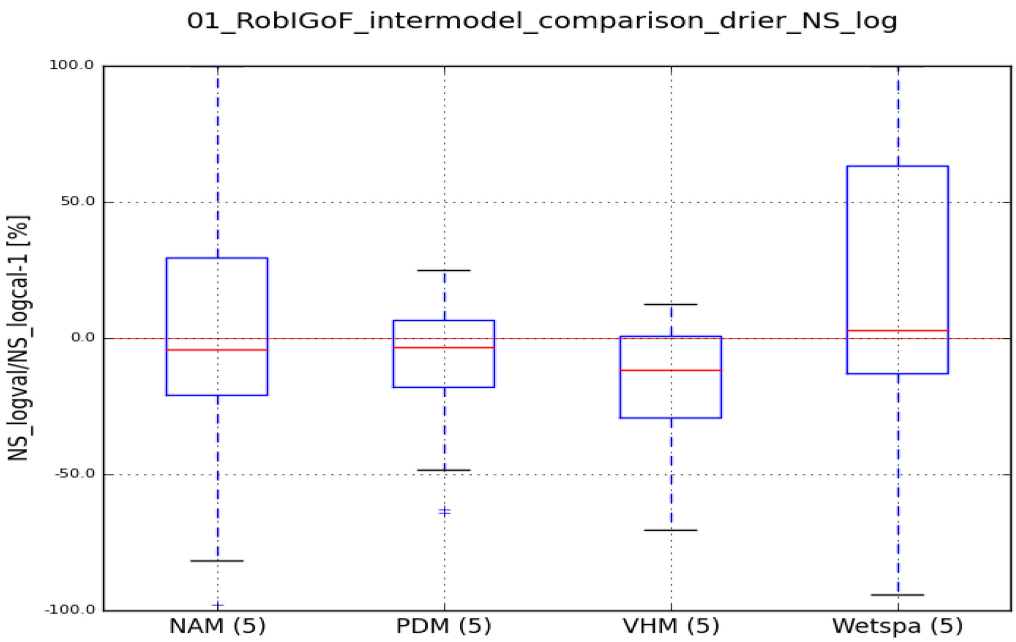


Figure 2 – Relative loss of NS_log performance against relative drier climate conditions
(basin)

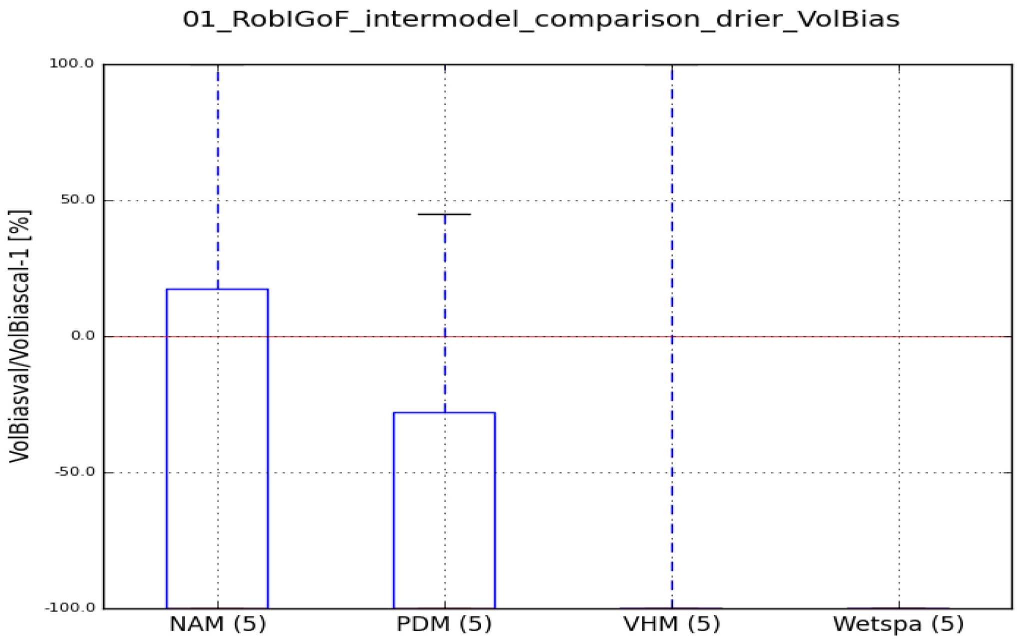


Figure 3 – Relative loss of VolBias performance against relative drier climate conditions (basin)

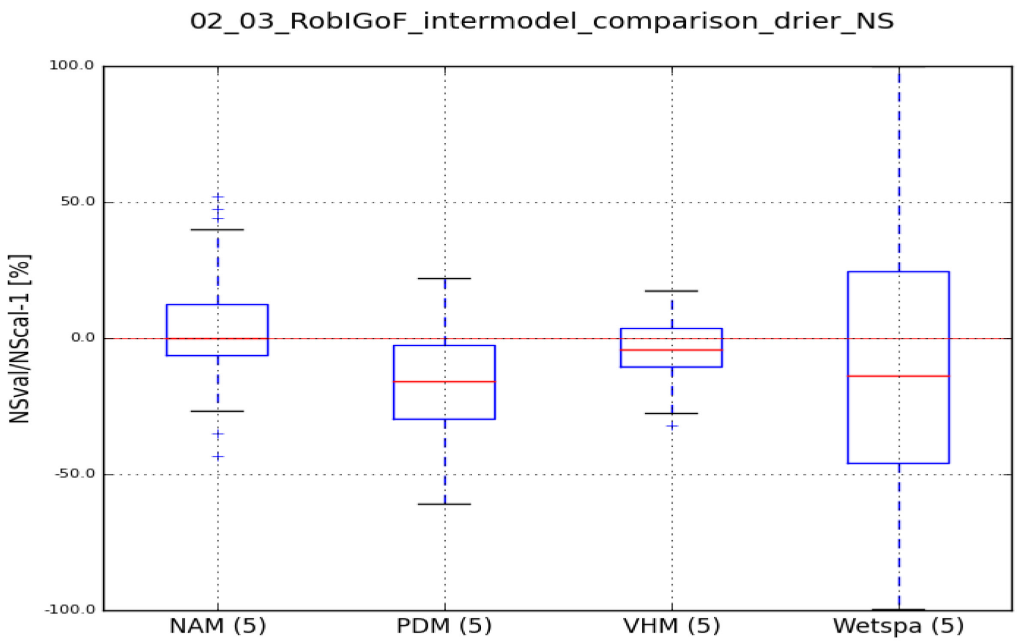


Figure 4 – Relative loss of NS performance against relative drier climate conditions (basin)

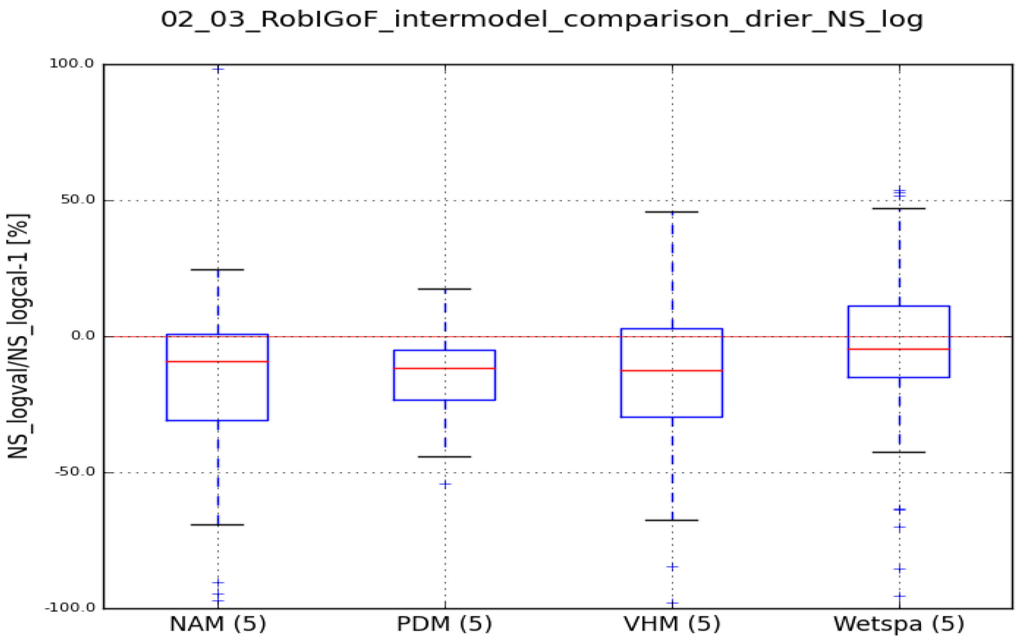


Figure 5 – Relative loss of NS_log performance against relative drier climate conditions (basin)

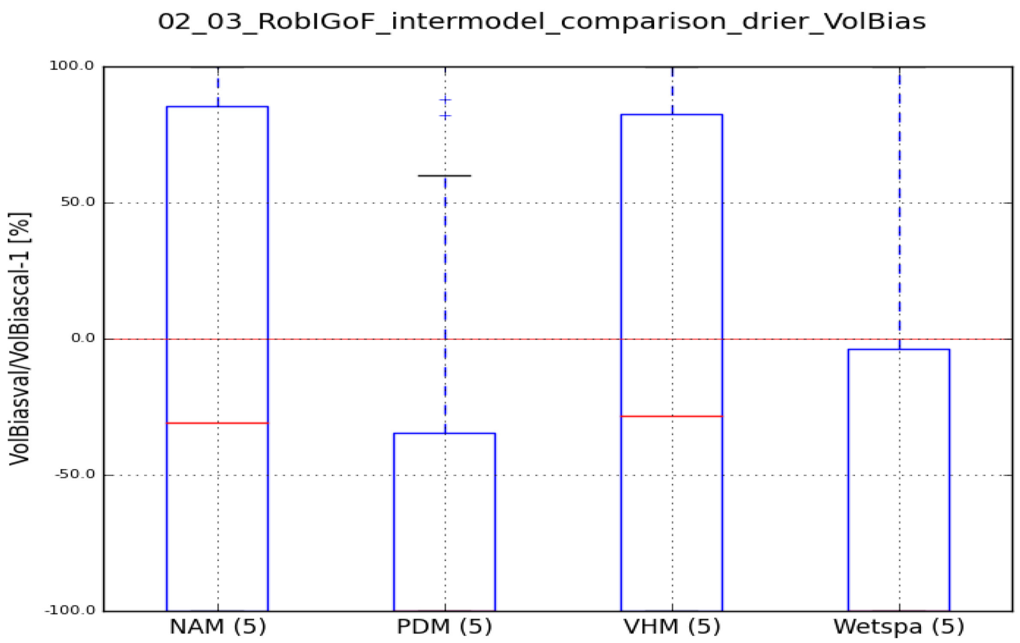


Figure 6 – Relative loss of VolBias performance against relative drier climate conditions (basin)

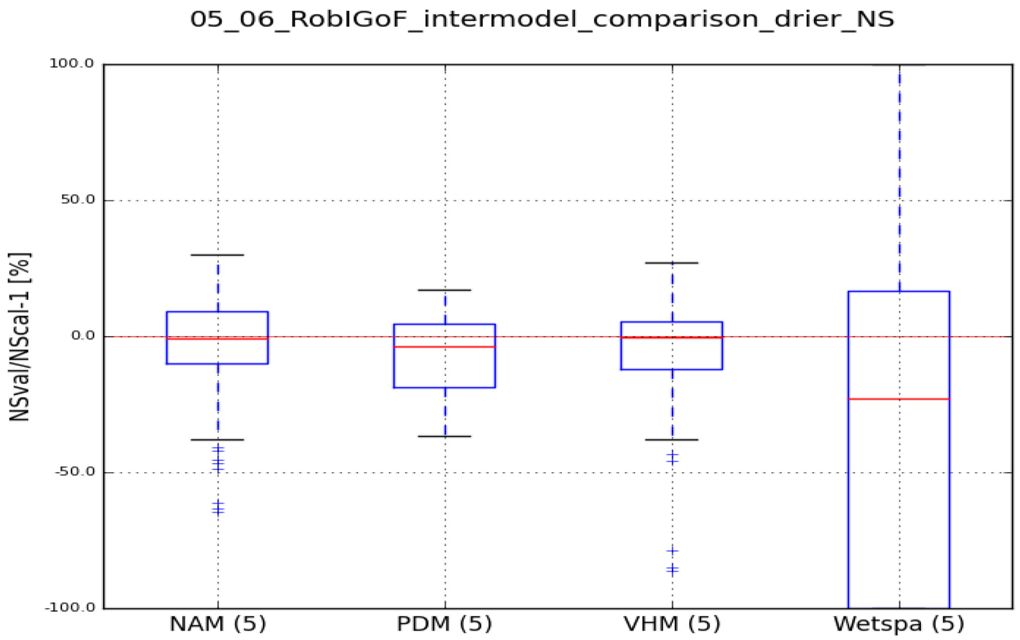


Figure 7 – Relative loss of NS performance against relative drier climate conditions (basin)

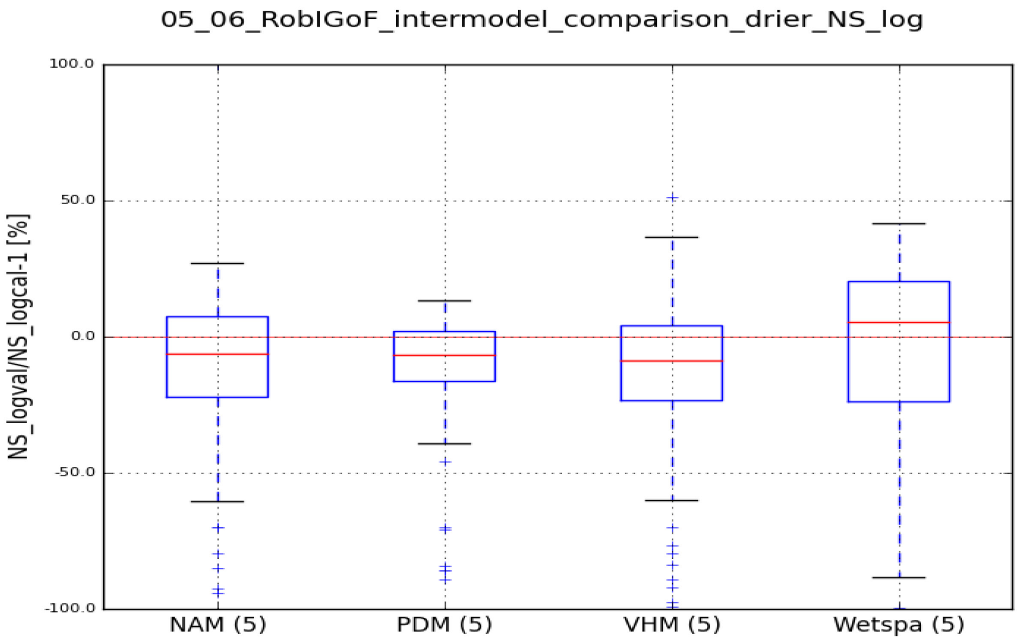


Figure 8 – Relative loss of NS_log performance against relative drier climate conditions (basin)

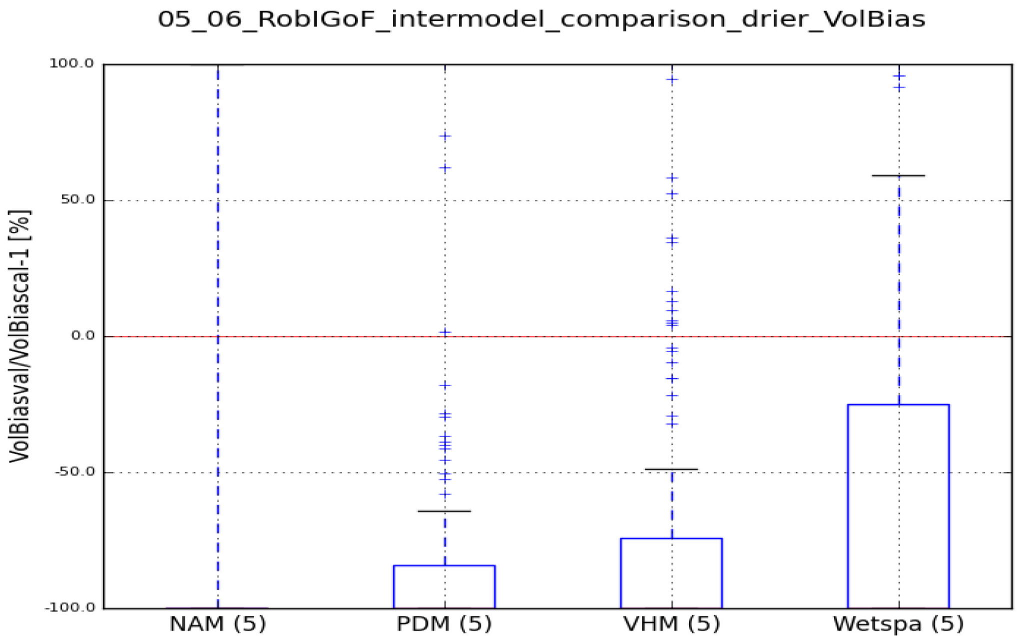


Figure 9 – Relative loss of VolBias performance against relative drier climate conditions (basin)

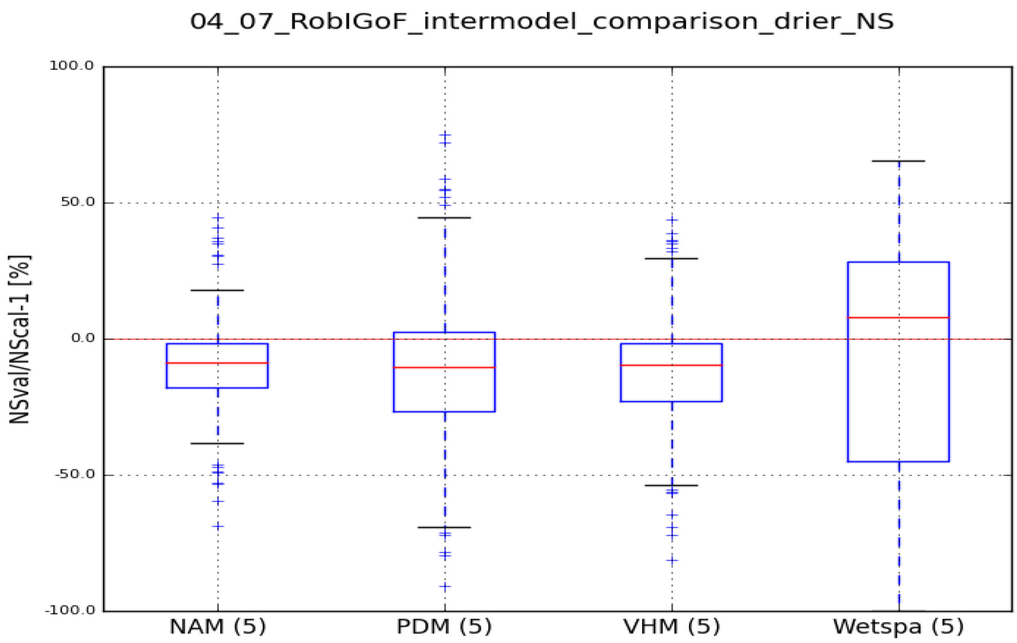


Figure 10 – Relative loss of NS performance against relative drier climate conditions (basin)

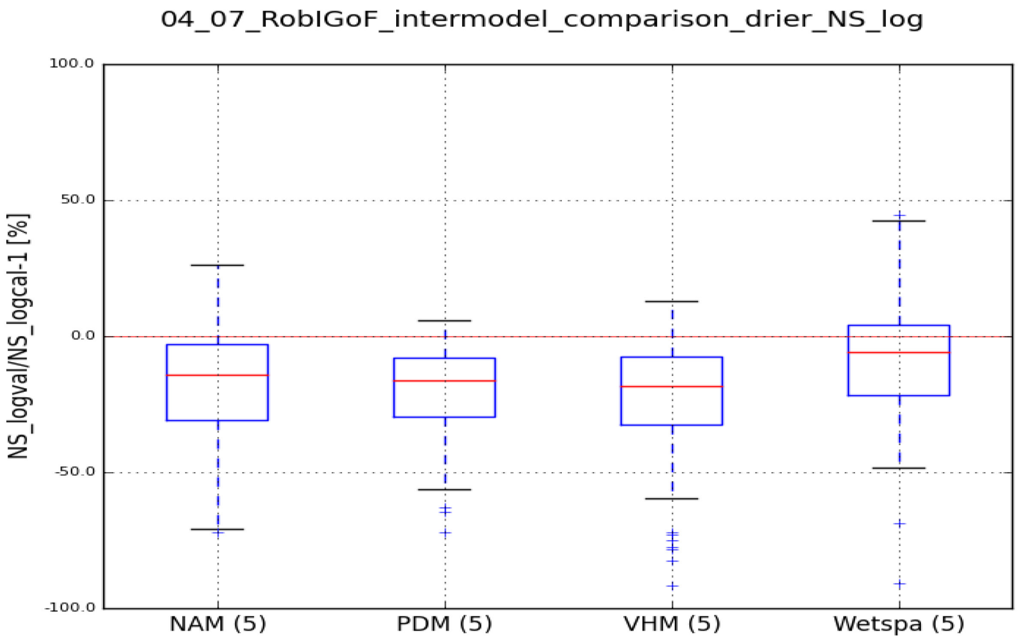


Figure 11 – Relative loss of NS_log performance against relative drier climate conditions (basin)

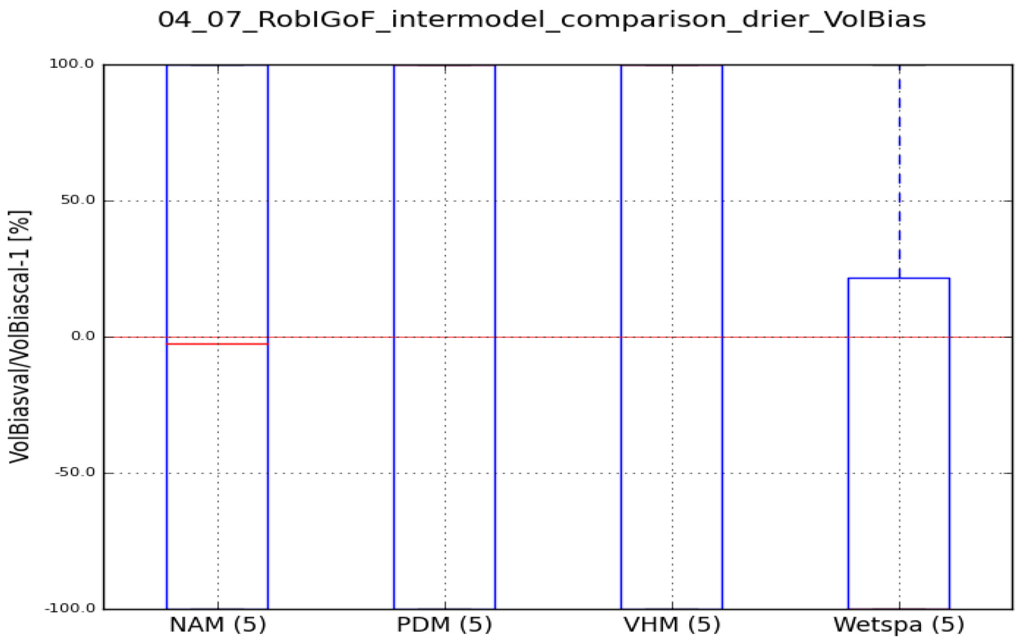


Figure 12 – Relative loss of VolBias performance against relative drier climate conditions (basin)

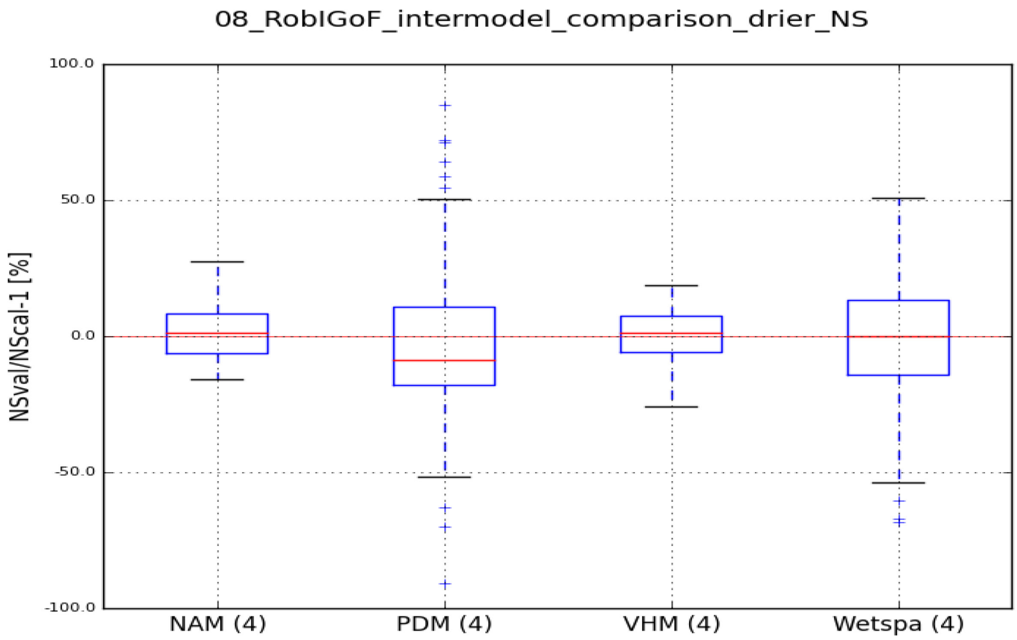


Figure 13 – Relative loss of NS performance against relative drier climate conditions (basin)

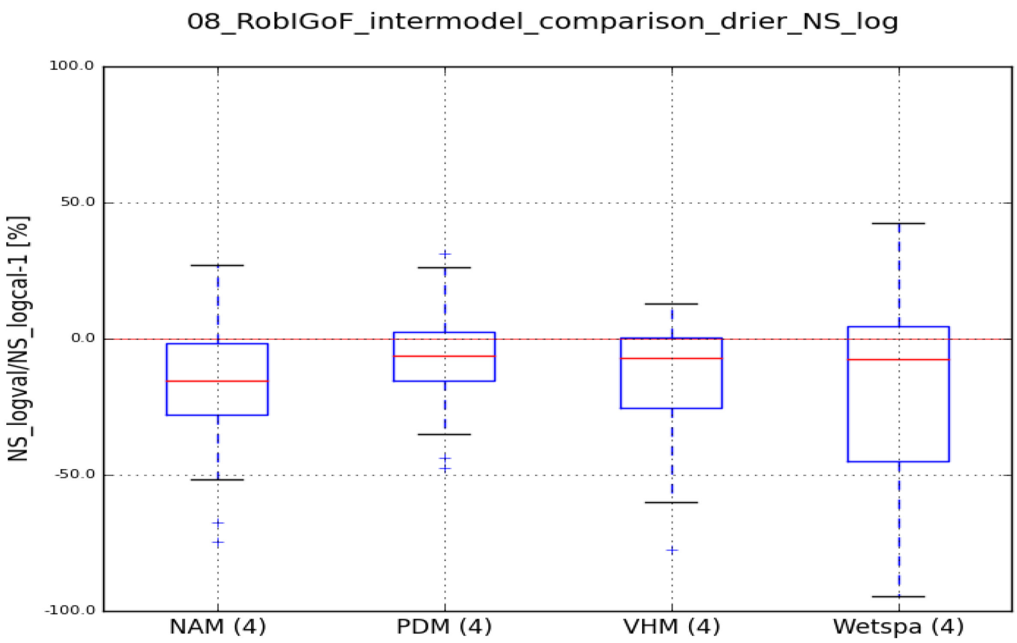


Figure 14 – Relative loss of NS_log performance against relative drier climate conditions (basin)

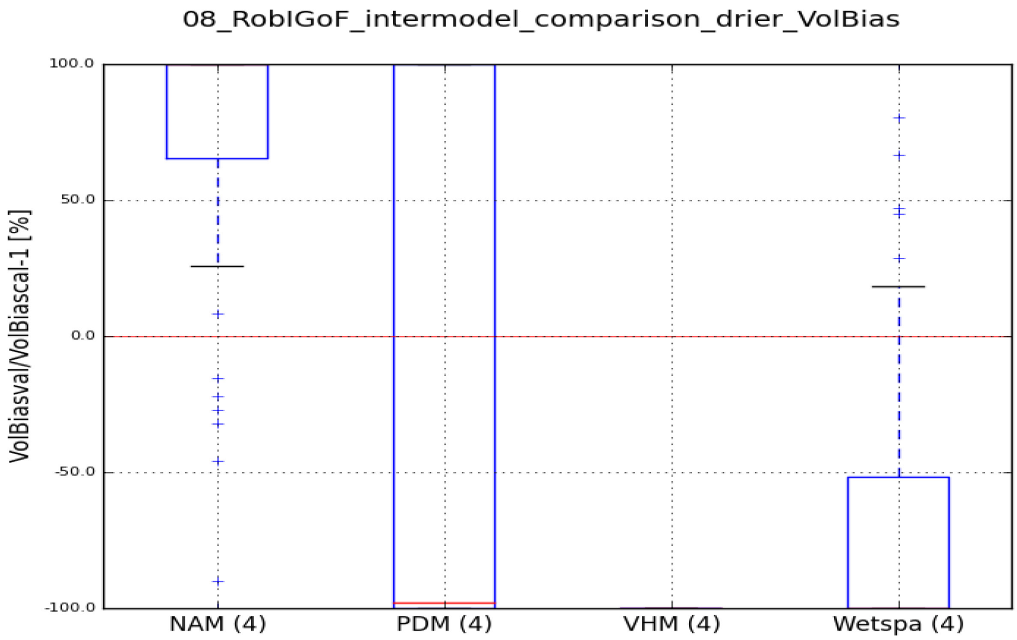


Figure 15 – Relative loss of VolBias performance against relative drier climate conditions (basin)

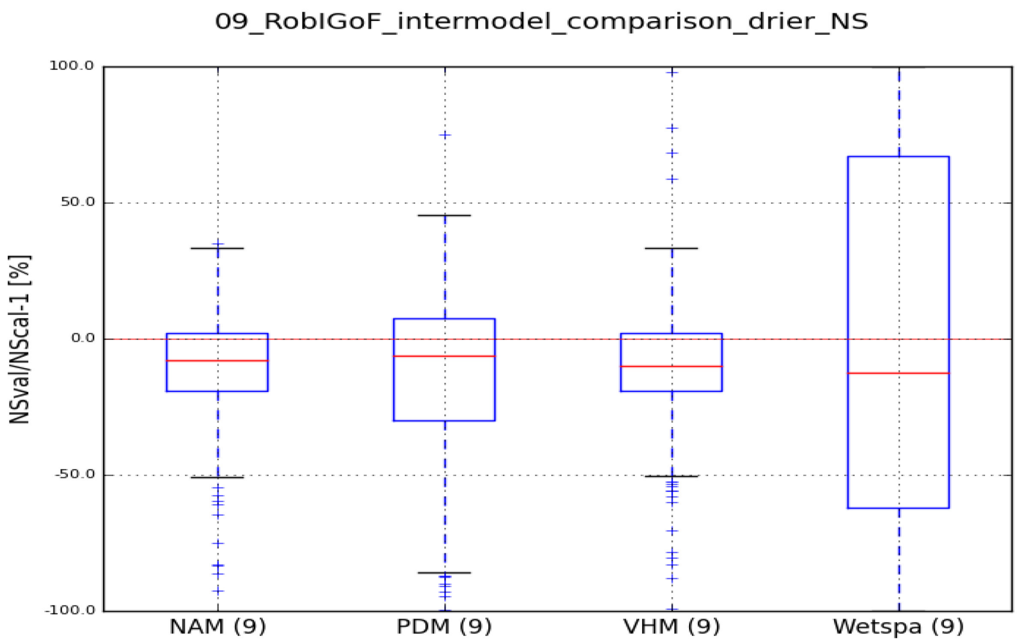


Figure 16 – Relative loss of NS performance against relative drier climate conditions (basin)

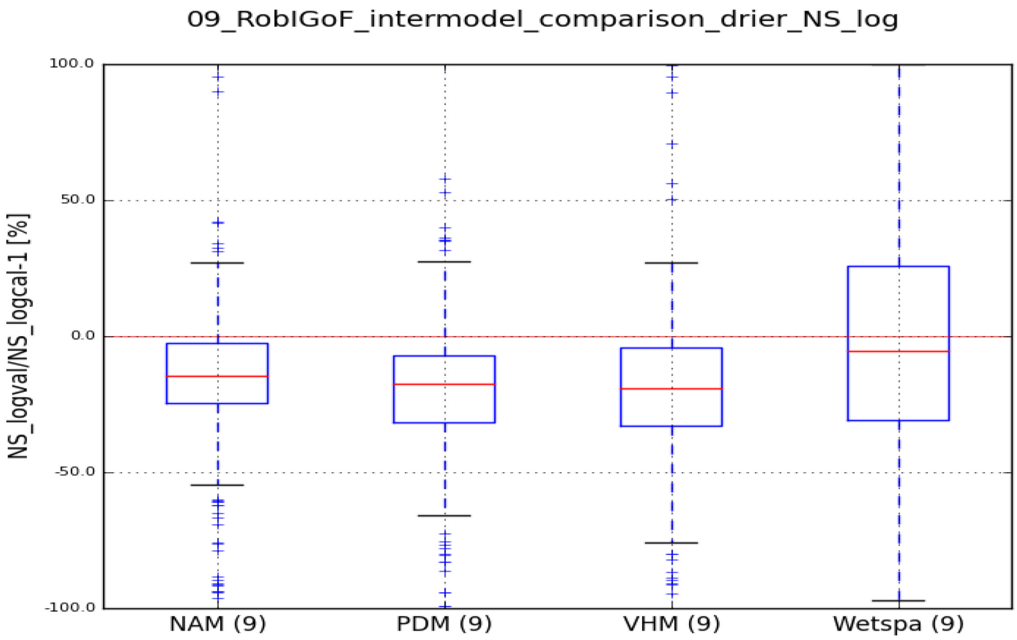


Figure 17 – Relative loss of NS_log performance against relative drier climate conditions (basin)

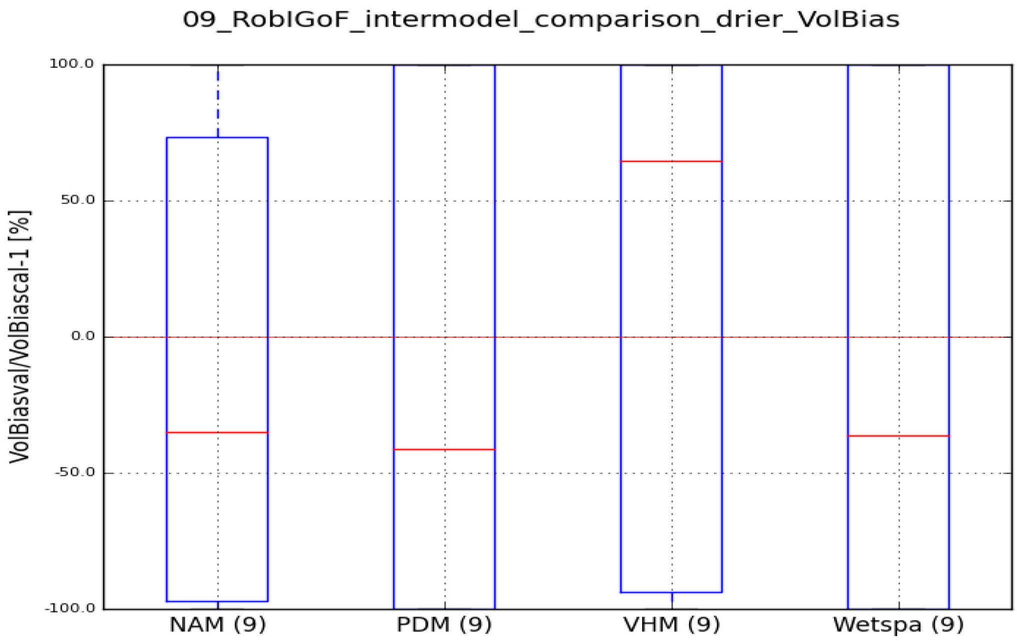


Figure 18 – Relative loss of VolBias performance against relative drier climate conditions (basin)

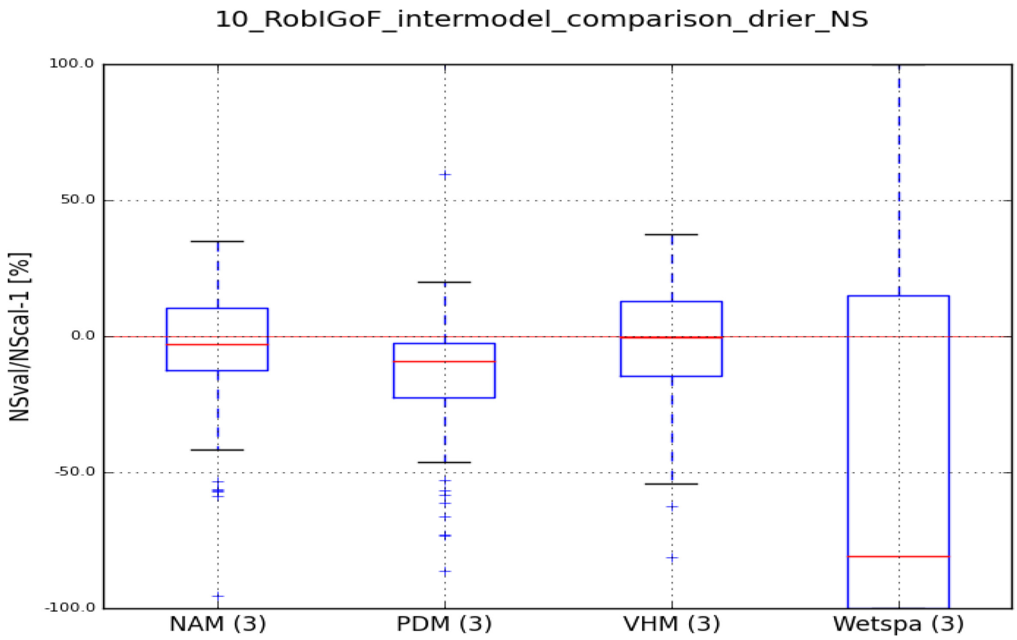


Figure 19 – Relative loss of NS performance against relative drier climate conditions (basin)

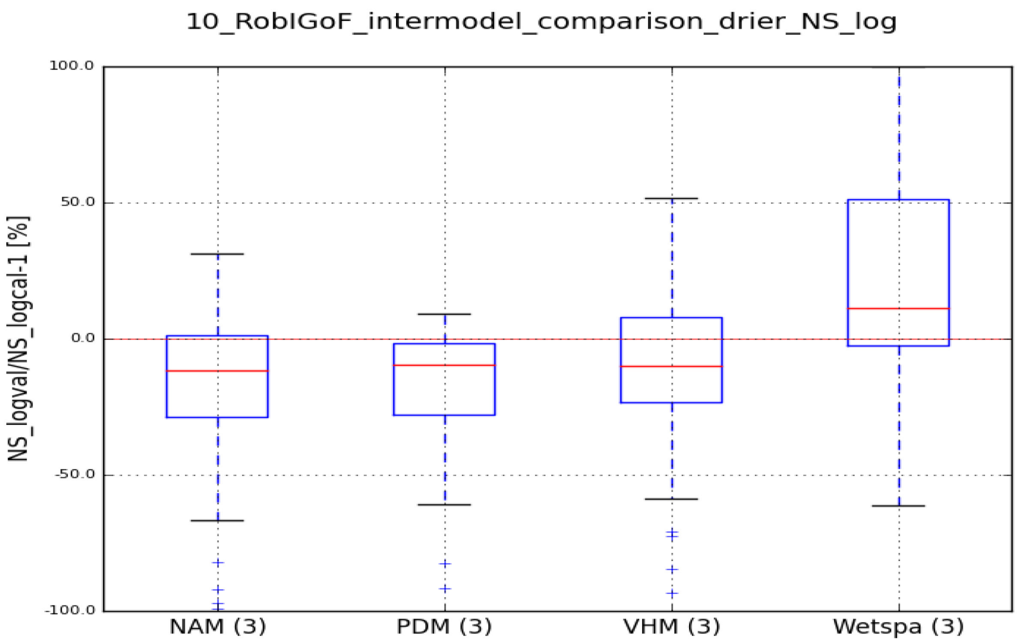


Figure 20 – Relative loss of NS_log performance against relative drier climate conditions (basin)

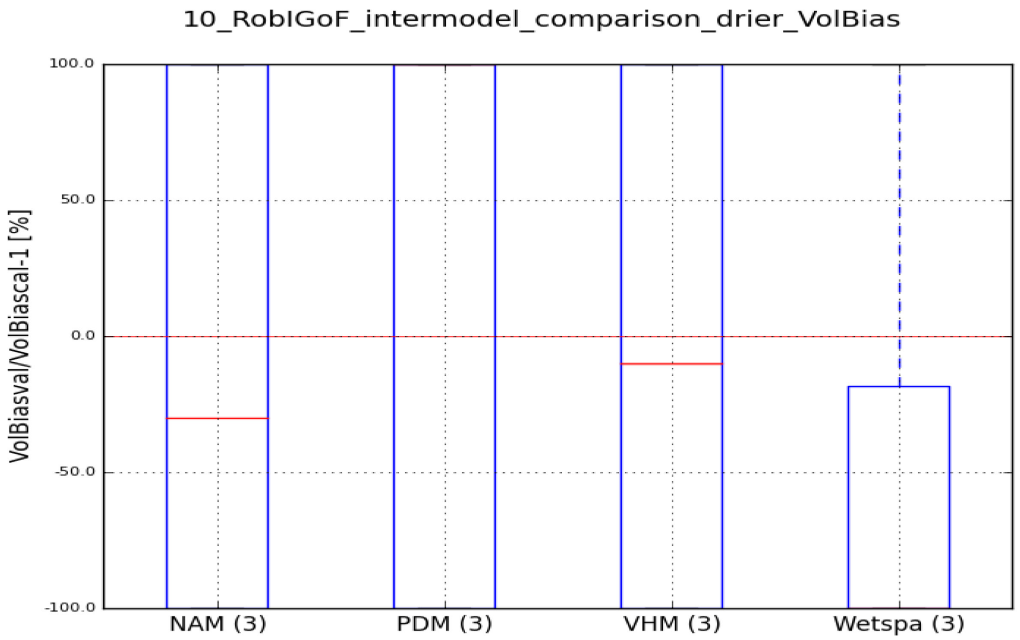


Figure 21 – Relative loss of VolBias performance against relative drier climate conditions (basin)

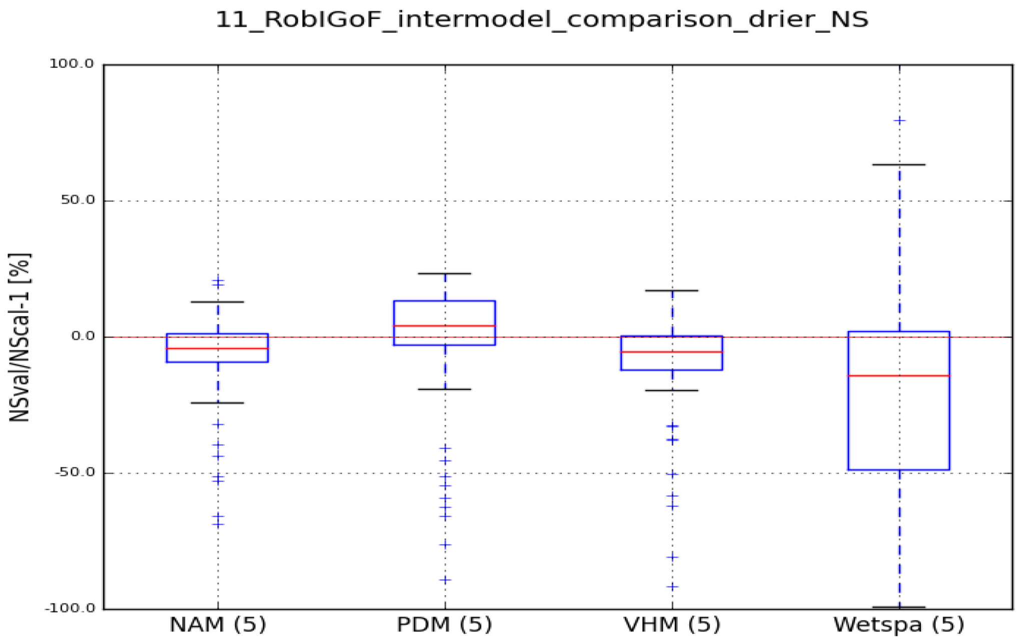


Figure 22 – Relative loss of NS performance against relative drier climate conditions (basin)

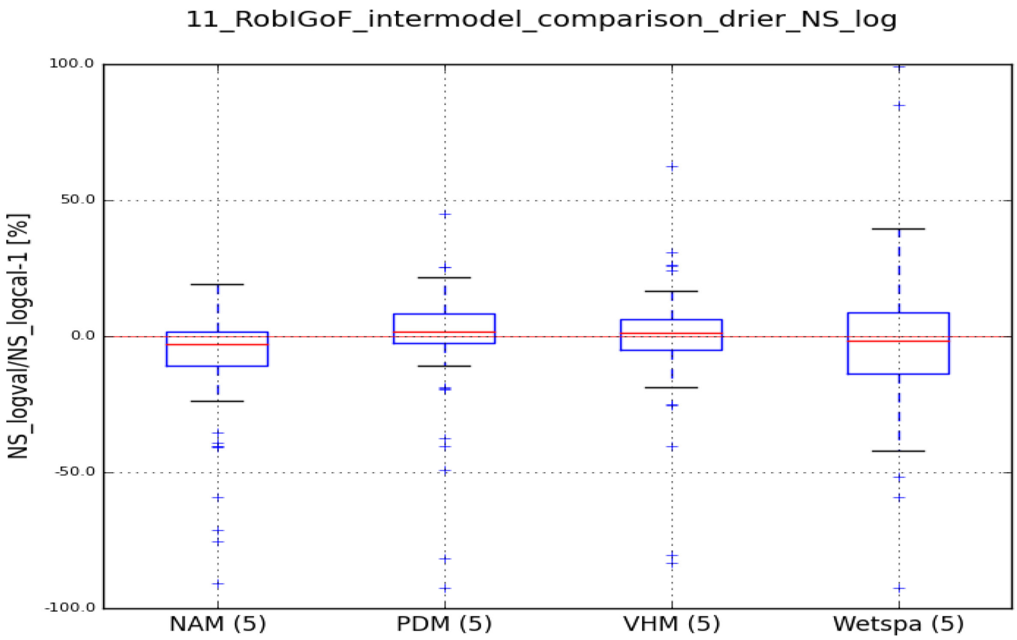


Figure 23 – Relative loss of NS_log performance against relative drier climate conditions (basin)

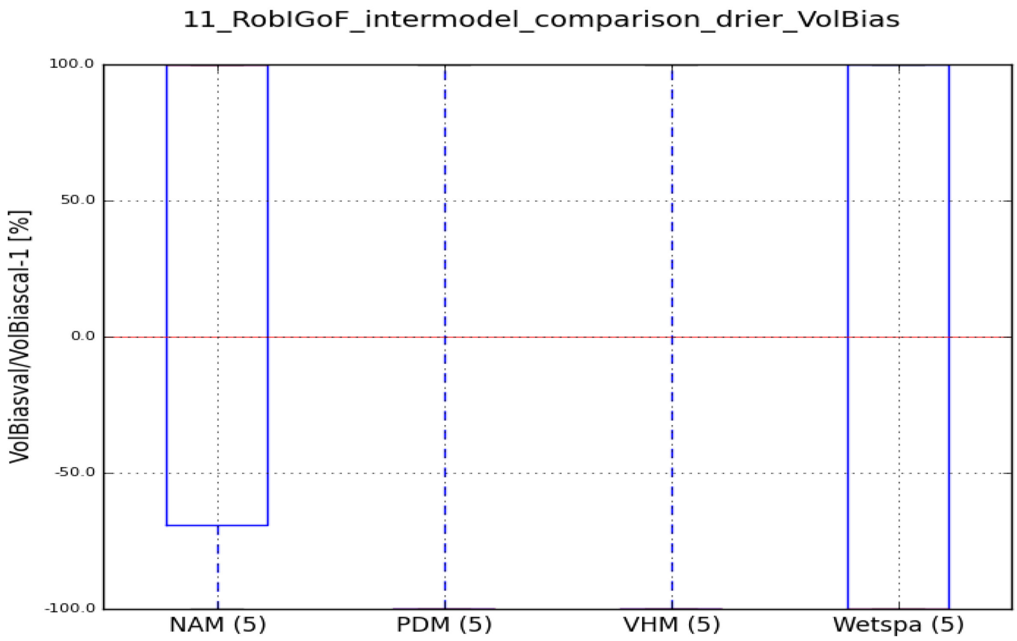


Figure 24 – Relative loss of VolBias performance against relative drier climate conditions (basin)

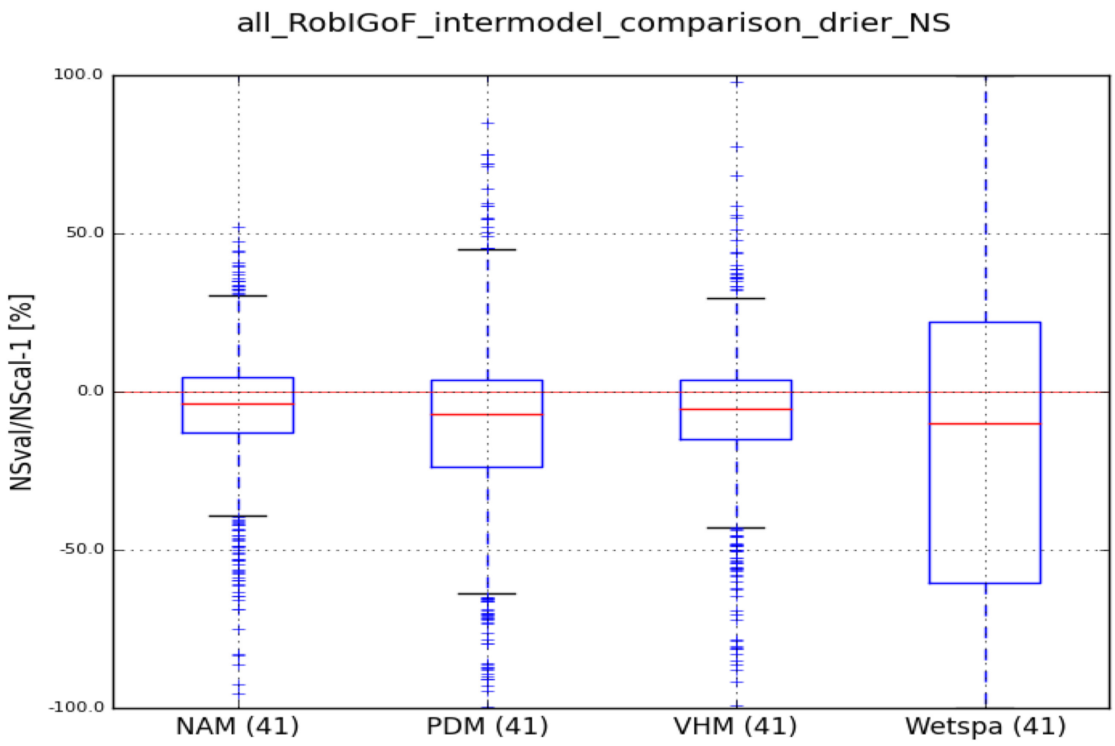


Figure 25 – Relative loss of NS performance against relative drier climate conditions

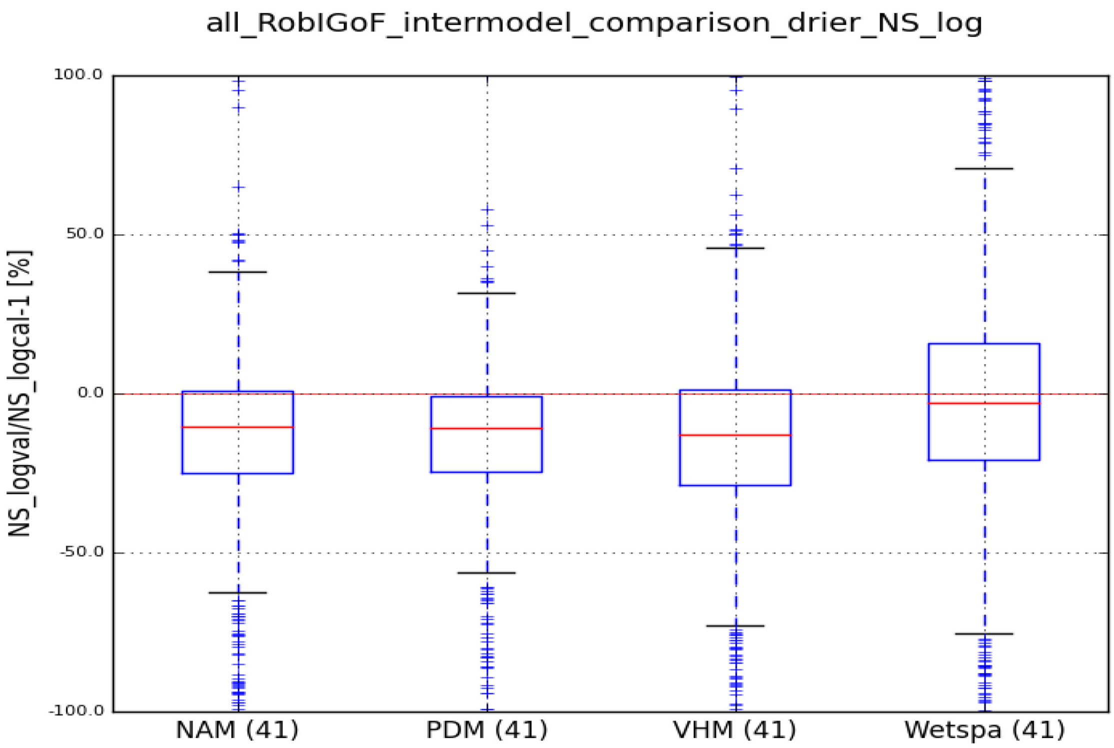


Figure 26 – Relative loss of NS_Log performance against relative drier climate conditions

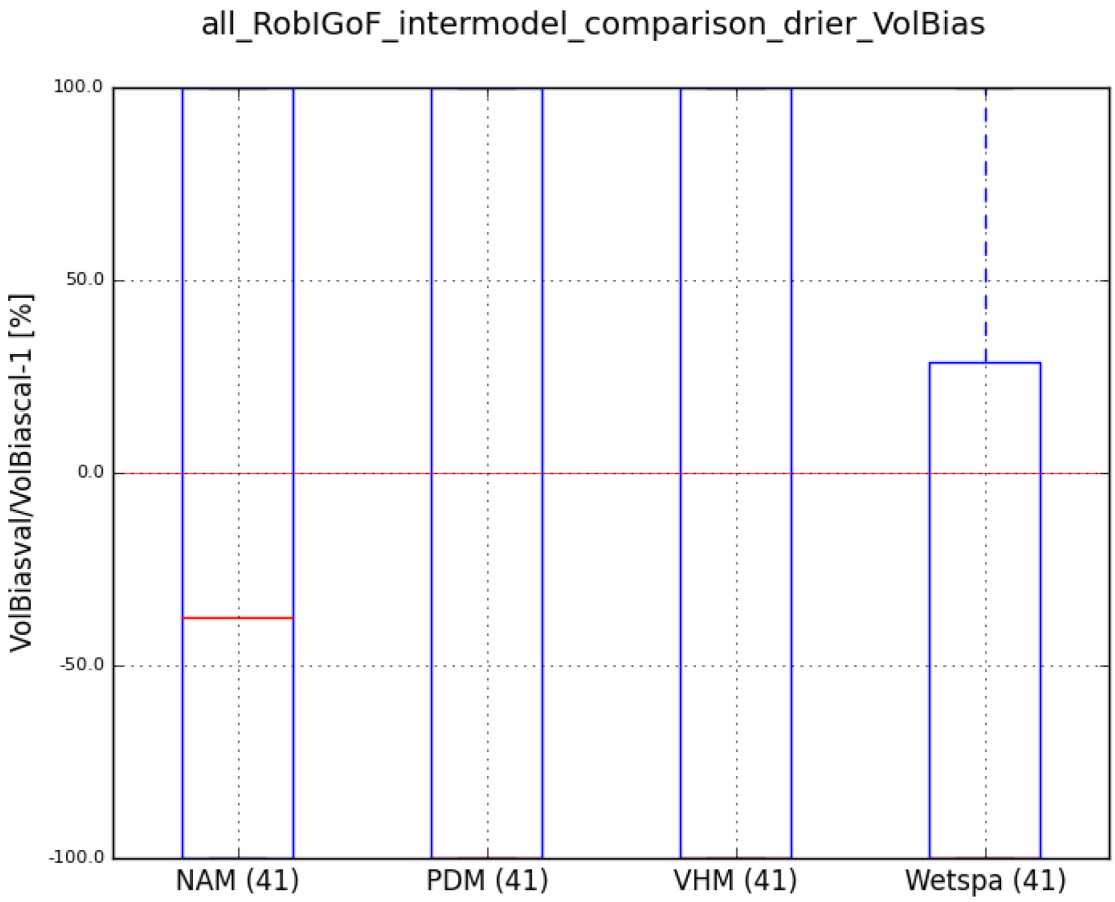


Figure 27 – Relative loss of Volume Bias performance against relative drier climate conditions

Appendix 5 Reports of the relative variation on projected climate and hydrological conditions methodology

Zie info: R:\00_162-ModWatbeschik\9_Rapportering\02-tweede-bestek\DO4\DO4_CCA_Final_Reports.zip

Appendix 6 The table for the different medians of the evaluation incators for all the catchments

Modelname	Model	Re.Ch Annual Cumulative Rainfall (%)	Re.Ch Annual Summer Cumulative Rainfall	Re.Ch Annual Cumulative Q	Re.Ch Summer Cumulative Q	Re.Ch(Vol Q)/ Re.Ch(Vol Rainfall)	Re.Ch (Q90/Q50)	Re.Ch Q90	Re.Ch Q50
F05LEI386999	NAM	-6.713	-19.240	-0.082	-0.224	1.093	-0.272	-0.392	-0.104
	PDM			-0.088	-0.248	1.113	-0.339	-0.413	-0.114
	VHM			-0.077	-0.253	1.061	-0.468	-0.557	-0.128
	WETSPA			-0.115	-0.247	1.228	-0.420	-0.486	-0.157
F06BOS325001	NAM	11.783	-8.779	0.184	0.172	2.087	-0.144	0.006	0.244
	PDM			-0.014	-0.131	-0.039	-0.180	-0.228	-0.051
	VHM			0.068	-0.057	0.655	-0.202	-0.175	0.032
	WETSPA			-0.146	-0.175	-1.069	-0.225	-0.337	-0.197
F11MAA8702	NAM	-14.053	-19.714	-0.149	-0.259	1.181	-0.365	-0.479	-0.198
	PDM			-0.244	-0.339	1.779	-0.150	-0.382	-0.257
	VHM			-0.157	-0.242	1.254	-0.208	-0.376	-0.232
	WETSPA			-0.163	-0.296	1.114	-0.424	-0.554	-0.237
V01HAN488180	NAM	5.735	-4.984	0.118	-0.209	1.886	-0.204	-0.278	-0.074
	PDM			0.131	-0.219	2.090	-0.446	-0.485	-0.052
	VHM			0.149	-0.134	2.379	-0.451	-0.430	0.059
	WETSPA			0.086	-0.141	1.338	-0.543	-0.547	-0.006
F01IJZ468000	NAM	4.900	-6.185	0.173	-0.077	2.501	-0.486	-0.424	0.111
	PDM			0.090	-0.146	1.542	-0.410	-0.426	-0.032
	VHM			0.165	-0.092	1.899	-0.316	-0.306	0.055
	WETSPA			0.070	-0.077	1.581	-0.369	-0.378	-0.005
V01IEP495080	NAM	5.603	-5.592	0.040	-0.108	1.432	-0.099	-0.077	0.056
	PDM			0.034	-0.187	1.093	-0.286	-0.327	-0.062
	VHM			0.103	-0.103	1.774	-0.194	-0.191	-0.007
	WETSPA			-0.001	-0.159	0.465	-0.342	-0.369	-0.046
V01KEM492060	NAM	5.524	-4.771	0.161	-0.099	2.638	-0.320	-0.303	0.061
	PDM			0.135	-0.216	2.122	-0.421	-0.448	-0.047
	VHM			0.159	-0.107	2.472	-0.465	-0.443	0.058
	WETSPA			0.093	-0.128	1.608	-0.624	-0.611	0.013
V01MAR496120	NAM	5.671	-6.106	0.197	-0.098	2.988	-0.330	-0.269	0.085
	PDM			0.145	-0.236	2.291	-0.507	-0.534	-0.070
	VHM			0.153	-0.182	2.584	-0.401	-0.442	-0.077
	WETSPA			0.177	-0.117	2.592	-0.525	-0.491	0.068
V01POP491030	NAM	5.452	-3.678	0.204	-0.063	2.784	-0.355	-0.275	0.130
	PDM			0.083	-0.233	1.752	-0.231	-0.300	-0.093
	VHM			0.214	-0.035	3.167	-0.430	-0.372	0.078
	WETSPA			0.135	-0.100	1.955	-0.239	-0.224	0.032

Modelname	Model	Re.Ch Annual Cumulative Rainfall (%)	Re.Ch Annual Summer Cumulative Rainfall	Re.Ch Annual Cumulative Q	Re.Ch Summer Cumulative Q	Re.Ch(Vol Q)/ Re.Ch(Vol Rainfall)	Re.Ch (Q90/Q50)	Re.Ch Q90	Re.Ch Q50
V01SSV499140	NAM	5.339	-6.320	0.124	-0.207	1.989	0.104	0.098	0.000
	PDM			0.102	-0.248	2.062	-0.347	-0.400	-0.092
	VHM			0.142	-0.125	2.455	-0.457	-0.442	0.060
	WETSPA			0.128	-0.122	2.100	-0.293	-0.327	-0.053
V02EDE442120	NAM	5.889	-7.413	0.070	-0.391	1.573	-0.106	-0.268	-0.222
	PDM			0.141	-0.295	2.142	-0.430	-0.491	-0.071
	VHM			0.148	-0.215	2.312	-0.385	-0.424	-0.041
	WETSPA			0.147	-0.124	2.470	-0.544	-0.541	0.043
V02HER426010	NAM	5.931	-6.750	-0.012	-0.261	0.751	0.082	-0.156	-0.213
	PDM			0.110	-0.249	1.994	-0.361	-0.441	-0.076
	VHM			0.122	-0.131	2.400	-0.233	-0.273	-0.031
	WETSPA			0.167	-0.158	2.456	-0.434	-0.455	-0.022
V02KER422030	NAM	5.935	-4.961	0.166	-0.101	2.694	-0.271	-0.281	0.094
	PDM			0.124	-0.240	1.879	-0.327	-0.383	-0.074
	VHM			0.136	-0.254	2.219	-0.403	-0.467	-0.129
	WETSPA			0.112	-0.142	1.721	-0.466	-0.459	0.042
V02RIV425020	NAM	6.466	-4.369	0.059	-0.139	1.415	-0.230	-0.259	0.012
	PDM			0.135	-0.221	2.306	-0.405	-0.471	-0.052
	VHM			0.150	-0.115	2.401	-0.230	-0.233	-0.007
	WETSPA			0.090	-0.118	1.477	-0.198	-0.240	-0.049
V03POE446000	NAM	5.166	-7.479	0.123	-0.212	2.469	-0.171	-0.325	-0.040
	PDM			0.127	-0.268	2.202	-0.473	-0.523	-0.055
	VHM			0.139	-0.164	2.402	-0.326	-0.380	0.025
	WETSPA			0.182	-0.077	2.856	-0.505	-0.476	0.053
V04MOL036110	NAM	5.928	-5.498	-0.002	-0.177	0.868	-0.189	-0.257	-0.096
	PDM			0.118	-0.233	1.808	-0.373	-0.404	-0.059
	VHM			0.083	-0.187	1.656	-0.355	-0.380	-0.083
	WETSPA			0.114	-0.131	1.808	-0.440	-0.435	0.000
V04MOM037100	NAM	6.081	-2.634	0.030	-0.242	1.013	-0.421	-0.460	-0.067
	PDM			0.080	-0.217	1.434	-0.339	-0.389	-0.070
	VHM			0.134	-0.077	1.786	-0.331	-0.307	0.091
	WETSPA			0.124	-0.119	1.602	-0.407	-0.376	0.050
V05HEU403210	NAM	5.421	-5.528	0.164	-0.176	2.780	-0.410	-0.423	0.069
	PDM			0.139	-0.242	2.359	-0.507	-0.567	-0.053
	VHM			0.161	-0.159	2.501	-0.440	-0.418	0.027
	WETSPA			0.128	-0.077	2.083	-0.362	-0.319	0.058
V05MAN401230	NAM	5.275	-4.897	0.021	-0.260	1.350	0.118	0.020	-0.153
	PDM			0.121	-0.226	2.096	-0.373	-0.412	-0.089
	VHM			0.112	-0.148	2.076	-0.433	-0.416	0.028
	WETSPA			0.044	-0.136	0.970	-0.424	-0.418	0.022

Modelname	Model	Re.Ch Annual Cumulative Rainfall (%)	Re.Ch Annual Summer Cumulative Rainfall	Re.Ch Annual Cumulative Q	Re.Ch Summer Cumulative Q	Re.Ch(Vol Q)/ Re.Ch(Vol Rainfall)	Re.Ch (Q90/Q50)	Re.Ch Q90	Re.Ch Q50
V06MAA347160	NAM	5.593	-6.513	-0.065	-0.132	0.846	0.121	-0.044	-0.166
	PDM			0.067	-0.246	2.155	-0.307	-0.393	-0.096
	VHM			0.138	-0.127	2.639	-0.276	-0.250	0.046
	WETSPA			0.154	-0.192	2.211	-0.495	-0.513	-0.050
V06ZWA342190	NAM	5.638	-3.908	0.120	-0.072	2.245	-0.243	-0.203	0.082
	PDM			0.056	-0.209	1.284	-0.301	-0.334	-0.044
	VHM			0.097	-0.147	1.622	-0.484	-0.467	0.041
	WETSPA			0.102	-0.133	1.687	-0.276	-0.282	-0.005
V07BEL285070	NAM	6.076	-3.651	0.012	-0.116	0.586	-0.016	-0.065	-0.053
	PDM			0.027	-0.214	1.054	-0.235	-0.305	-0.086
	VHM			0.106	-0.180	1.802	-0.353	-0.399	-0.059
	WETSPA			0.110	-0.208	1.803	-0.384	-0.413	-0.040
V07MAR289015	NAM	6.046	-5.624	0.031	-0.246	1.171	-0.372	-0.419	-0.055
	PDM			0.102	-0.249	1.862	-0.302	-0.390	-0.114
	VHM			0.152	-0.089	2.369	-0.288	-0.236	0.050
	WETSPA			0.106	-0.189	1.786	-0.279	-0.339	-0.070
V07MOE282100	NAM	5.230	-6.120	0.159	-0.080	2.690	-0.221	-0.181	0.126
	PDM			0.097	-0.243	1.706	-0.345	-0.411	-0.074
	VHM			0.162	-0.070	2.567	-0.265	-0.185	0.121
	WETSPA			0.136	-0.172	2.268	-0.464	-0.495	-0.008
V07MOG288020	NAM	6.329	-3.066	-0.099	-0.352	-0.707	-0.130	-0.300	-0.239
	PDM			0.151	-0.254	2.582	-0.366	-0.449	-0.119
	VHM			0.096	-0.146	1.958	-0.420	-0.458	-0.056
	WETSPA			0.169	-0.191	2.807	-0.474	-0.483	-0.019
V08BAR111370	NAM	5.150	-4.705	0.020	-0.185	0.809	-0.266	-0.314	-0.036
	PDM			0.017	-0.194	0.905	-0.259	-0.314	-0.070
	VHM			0.039	-0.145	1.120	-0.274	-0.298	-0.012
	WETSPA			0.006	-0.168	0.681	-0.381	-0.394	-0.008
V08DIJ093400	NAM	22.566	-27.680	0.053	-0.109	0.085	-0.287	-0.249	0.017
	PDM			0.034	-0.103	0.104	-0.150	-0.170	-0.023
	VHM			0.124	-0.094	0.365	-0.103	-0.068	0.073
	WETSPA			0.048	-0.076	0.099	-0.284	-0.272	0.045
V08ZUU233100	NAM	4.720	-8.075	0.011	-0.162	0.652	-0.215	-0.252	0.006
	PDM			0.055	-0.253	1.359	-0.255	-0.362	-0.140
	VHM			0.126	-0.160	1.939	-0.425	-0.451	-0.041
	WETSPA			0.146	-0.146	2.101	-0.466	-0.467	0.021
V09DEM136000	NAM	3.754	-5.074	-0.021	-0.163	0.710	-0.131	-0.161	-0.028
	PDM			-0.052	-0.185	-0.076	-0.190	-0.243	-0.062
	VHM			0.023	-0.116	0.885	-0.056	-0.013	0.023
	WETSPA			-0.159	-0.362	-1.372	-0.402	-0.572	-0.287

Modelname	Model	Re.Ch Annual Cumulative Rainfall (%)	Re.Ch Annual Summer Cumulative Rainfall	Re.Ch Annual Cumulative Q	Re.Ch Summer Cumulative Q	Re.Ch(Vol Q)/ Re.Ch(Vol Rainfall)	Re.Ch (Q90/Q50)	Re.Ch Q90	Re.Ch Q50
V09GET152080	NAM	3.834	-9.796	-0.085	-0.217	-0.458	-0.196	-0.310	-0.096
	PDM			-0.072	-0.181	-0.607	-0.201	-0.280	-0.096
	VHM			0.085	-0.054	1.571	-0.163	-0.152	0.051
	WETSPA			0.009	-0.067	0.622	-0.139	-0.111	0.017
V09HER163010	NAM	3.327	-7.707	-0.107	-0.174	-0.379	-0.125	-0.163	-0.105
	PDM			-0.065	-0.199	-0.332	-0.209	-0.287	-0.106
	VHM			0.061	-0.131	1.459	-0.246	-0.242	0.025
	WETSPA			0.083	-0.172	1.843	-0.312	-0.323	-0.016
V09HUL147150	NAM	4.925	-8.019	0.082	-0.006	1.718	-0.036	0.033	0.078
	PDM			-0.020	-0.129	0.311	-0.145	-0.177	-0.025
	VHM			0.062	-0.148	1.254	-0.378	-0.322	0.061
	WETSPA			-0.017	-0.143	0.140	-0.235	-0.274	-0.040
V09MAN161040	NAM	4.892	-5.403	0.017	-0.132	1.149	-0.244	-0.247	0.016
	PDM			-0.039	-0.161	0.321	-0.217	-0.260	-0.040
	VHM			0.081	-0.110	1.538	-0.233	-0.205	0.055
	WETSPA			0.041	-0.115	0.976	-0.314	-0.318	0.047
V09MOT144270	NAM	4.514	-7.870	0.030	-0.091	0.897	-0.014	-0.025	-0.023
	PDM			0.005	-0.189	0.582	-0.265	-0.316	-0.060
	VHM			0.079	-0.149	1.712	-0.191	-0.222	-0.024
	WETSPA			0.070	-0.129	1.300	-0.452	-0.460	0.014
V09VEL145100	NAM	3.939	-10.039	-0.069	-0.215	0.403	-0.201	-0.312	-0.117
	PDM			0.035	-0.229	1.327	-0.296	-0.354	-0.101
	VHM			0.075	-0.147	1.821	-0.233	-0.295	-0.074
	WETSPA			0.079	-0.146	1.508	-0.483	-0.484	0.009
V09WIN141310	NAM	4.452	-7.477	-0.135	-0.329	-0.708	-0.241	-0.417	-0.228
	PDM			-0.062	-0.205	-0.090	-0.160	-0.261	-0.109
	VHM			0.083	-0.158	1.770	-0.306	-0.318	-0.022
	WETSPA			0.137	-0.149	2.155	-0.504	-0.517	0.011
V09ZWA148120	NAM	5.236	-7.727	-0.025	-0.157	0.635	-0.094	-0.144	-0.045
	PDM			-0.061	-0.160	-0.269	-0.155	-0.212	-0.054
	VHM			0.072	-0.139	1.415	-0.126	-0.135	0.007
	WETSPA			0.074	-0.133	1.357	-0.391	-0.378	0.035
V10GNE076999	NAM	5.623	-11.882	0.008	-0.086	0.426	-0.169	-0.158	0.010
	PDM			-0.042	-0.115	-0.539	-0.131	-0.170	-0.041
	VHM			0.034	-0.044	0.644	-0.097	-0.083	0.014
	WETSPA			-0.013	-0.096	-0.021	-0.226	-0.240	-0.016
V10KNE052000	NAM	5.073	-6.165	0.049	-0.172	1.607	-0.308	-0.329	-0.015
	PDM			0.010	-0.213	0.819	-0.258	-0.315	-0.063
	VHM			0.077	-0.139	1.625	-0.344	-0.325	0.045
	WETSPA			0.092	-0.160	1.494	-0.422	-0.410	0.031

Modelname	Model	Re.Ch Annual Cumulative Rainfall (%)	Re.Ch Annual Summer Cumulative Rainfall	Re.Ch Annual Cumulative Q	Re.Ch Summer Cumulative Q	Re.Ch(Vol Q)/ Re.Ch(Vol Rainfall)	Re.Ch (Q90/Q50)	Re.Ch Q90	Re.Ch Q50
V10MOP062140	NAM	4.840	-7.561	-0.102	-0.337	-0.502	-0.537	-0.622	-0.167
V10MOP062140	PDM	4.840	-7.561	-0.010	-0.223	0.716	-0.236	-0.310	-0.105
V10MOP062140	VHM	4.840	-7.561	0.087	-0.191	1.762	-0.503	-0.515	0.013
V10MOP062140	WETSPA	4.840	-7.561	0.087	-0.134	1.518	-0.436	-0.418	0.015
V10WIM082050	NAM	4.717	-7.537	-0.088	-0.245	-1.124	-0.252	-0.327	-0.108
V10WIM082050	PDM	4.717	-7.537	-0.023	-0.215	0.838	-0.212	-0.300	-0.095
V10WIM082050	VHM	4.717	-7.537	0.103	-0.129	2.012	-0.227	-0.273	-0.005
V10WIM082050	WETSPA	4.717	-7.537	0.141	-0.120	2.358	-0.520	-0.506	0.053
W06RHOL54100	NAM	5.924	-4.594	-0.109	-0.257	-0.616	-0.118	-0.251	-0.248
W06RHOL54100	PDM	5.924	-4.594	0.024	-0.251	1.162	-0.266	-0.350	-0.127
W06RHOL54100	VHM	5.924	-4.594	0.097	-0.141	1.814	-0.267	-0.278	0.017
W06RHOL54100	WETSPA	5.924	-4.594	0.229	-0.173	3.439	-0.536	-0.539	-0.025
W07DENLES004	NAM	6.148	-11.046	-0.082	-0.191	-0.902	-0.025	-0.191	-0.201
W07DENLES004	PDM	6.148	-11.046	-0.011	-0.166	0.375	-0.157	-0.243	-0.096
W07DENLES004	VHM	6.148	-11.046	0.070	-0.062	1.325	-0.149	-0.141	0.012
W07DENLES004	WETSPA	6.148	-11.046	0.110	-0.109	1.694	-0.338	-0.353	-0.033
W08SAMRON000	NAM	5.486	-5.941	-0.062	-0.213	0.220	-0.267	-0.326	-0.081
W08SAMRON000	PDM	5.486	-5.941	-0.001	-0.202	0.721	-0.233	-0.304	-0.082
W08SAMRON000	VHM	5.486	-5.941	0.095	-0.116	1.584	-0.182	-0.179	-0.004
W08SAMRON000	WETSPA	5.486	-5.941	0.083	-0.117	1.438	-0.269	-0.268	0.008
W08SENRON010	NAM	4.348	-9.150	-0.057	-0.188	0.315	-0.253	-0.311	-0.063
W08SENRON010	PDM	4.348	-9.150	-0.039	-0.218	0.305	-0.206	-0.292	-0.107
W08SENRON010	VHM	4.348	-9.150	0.075	-0.110	1.574	-0.218	-0.217	0.028
W08SENRON010	WETSPA	4.348	-9.150	0.077	-0.245	1.487	-0.370	-0.444	-0.110
W08SENTUB030	NAM	4.347	-7.134	0.061	-0.122	1.858	-0.270	-0.248	-0.010
W08SENTUB030	PDM	4.347	-7.134	0.044	-0.237	1.429	-0.253	-0.358	-0.123
W08SENTUB030	VHM	4.347	-7.134	0.159	-0.045	2.601	-0.133	-0.106	0.046
W08SENTUB030	WETSPA	4.347	-7.134	0.171	-0.128	2.919	-0.478	-0.489	0.014
W11BER551010	NAM	4.451	-6.684	0.094	-0.180	1.695	-0.305	-0.309	-0.018
W11BER551010	PDM	4.451	-6.684	0.066	-0.218	1.372	-0.293	-0.335	-0.058
W11BER551010	VHM	4.451	-6.684	0.080	-0.187	1.568	-0.369	-0.390	-0.029
W11BER551010	WETSPA	4.451	-6.684	0.092	-0.141	1.571	-0.383	-0.379	0.002
W11HOY5990	NAM	5.322	-3.507	0.331	0.139	4.424	-0.276	-0.069	0.294
W11HOY5990	PDM	5.322	-3.507	0.212	-0.070	3.838	-0.201	-0.193	0.038
W11HOY5990	VHM	5.322	-3.507	0.400	0.157	5.653	-0.266	-0.020	0.359
W11HOY5990	WETSPA	5.322	-3.507	0.123	-0.094	1.748	-0.349	-0.332	0.028
W11MAAPROF	NAM	7.417	-3.449	0.173	-0.077	2.216	-0.486	-0.424	0.111
W11MAAPROF	PDM	7.417	-3.449	-0.244	-0.339	-2.750	-0.150	-0.382	-0.257
W11MAAPROF	VHM	7.417	-3.449	0.195	0.041	2.496	-0.247	-0.111	0.129
W11MAAPROF	WETSPA	7.417	-3.449	0.090	-0.196	1.090	-0.479	-0.478	0.033

Modelname	Model	Re.Ch Annual Cumulative Rainfall (%)	Re.Ch Annual Summer Cumulative Rainfall	Re.Ch Annual Cumulative Q	Re.Ch Summer Cumulative Q	Re.Ch(Vol Q)/ Re.Ch(Vol Rainfall)	Re.Ch (Q90/Q50)	Re.Ch Q90	Re.Ch Q50
W11MEH5820	NAM	3.931	-6.631	-0.075	-0.167	-0.261	-0.239	-0.281	-0.090
	PDM			0.065	-0.218	1.600	-0.230	-0.292	-0.094
	VHM			0.065	-0.139	1.308	-0.230	-0.204	0.057
	WETSPA			0.194	-0.112	3.184	-0.356	-0.347	0.040
W11OUR5805	NAM	6.172	-4.972	0.140	-0.061	2.034	-0.274	-0.246	0.091
	PDM			0.109	-0.230	1.675	-0.379	-0.412	-0.038
	VHM			0.148	-0.096	2.031	-0.431	-0.412	0.050
	WETSPA			0.093	-0.128	1.435	-0.493	-0.494	-0.005
W11SAM7319	NAM	4.879	-8.135	0.041	-0.164	1.064	-0.249	-0.241	0.008
	PDM			0.048	-0.215	1.298	-0.234	-0.316	-0.117
	VHM			0.131	-0.083	2.189	-0.072	-0.056	0.020
	WETSPA			0.139	-0.096	2.415	-0.379	-0.389	0.056

Appendix 7 The model efficiency coefficient Table

Modelname	Area	NS				NS_log				ReError				NAM		PDM		VHM		WETSPA	
		NAM	PDM	VHM	WETSPA	NAM	PDM	VHM	WETSPA	NAM	PDM	VHM	WETSPA	Start Calib	End Calib	Start Calib	End Calib	Start Calib	End Calib	Start Calib	End Calib
F01IJZ468000	393.007	0.682	0.681	0.605	0.328	0.641	0.69	0.522	0.398	0.009	-0.022	-0.032	-0.04	200501010000	201312310000	200501010000	201312310000	200501010000	201312310000	200501010000	199612310000
V01HAN488180	78.559	0.789	0.619	0.703	0.669	0.777	0.784	0.735	0.67	0.002	0.006	0.02	0.006	200101010000	201312310000	200101010000	201312310000	200101010000	201312310000	200801010000	201301310000
V01MAR496120	76.137	0.676	0.683	0.673	-0.052	0.672	0.678	0.606	0.401	0.012	-0.039	-0.02	0.007	200101010000	201312310000	200101010000	201312310000	200101010000	201312310000	200201010000	200812310000
V01POP491030	84.868	0.617	0.587	0.584	0.077	0.423	0.629	0.618	0.326	0.052	-0.022	-0.015	-0.053	200101010000	201312310000	200101010000	200812310000	200101010000	201312310000	200801010000	201312310000
V01SSV499140	16.095	0.684	0.644	0.607	0.453	-0.067	0.682	0.619	0.288	-0.206	-0.036	0.007	-0.031	199301010000	200506300000	199301010000	200506300000	199306300000	200506300000	199901010000	200412310000
V02EDE442120	45.489	0.809	0.726	0.722	0.622	0.549	0.709	0.634	0.49	0.014	0	0.009	-0.051	200101010000	201312310000	200101010000	201312310000	200101010000	201312310000	200905010000	201312310000
V02HER426010	77.272	0.683	0.578	0.653	0.217	0.614	0.76	0.549	0.599	-0.01	0.056	-0.002	-0.051	200101010000	201312310000	200101010000	201312310000	200101010000	201312310000	200001010000	201312310000
V02KER422030	62.719	0.605	0.606	0.691	0.594	0.664	0.706	0.626	0.54	0.007	-0.015	0.006	-0.061	199501010000	200712310000	199501010000	200712310000	199501010000	200712310000	200301010000	200712310000
V02RIV425020	63.98	0.587	0.702	0.649	-0.058	0.627	0.82	0.616	0.648	-0.031	-0.023	-0.015	-0.022	200101010000	201312310000	200101010000	201312310000	200101010000	201312310000	200301010000	201312310000
V03POE446000	106.837	0.722	0.71	0.724	0.503	0.696	0.705	0.645	0.774	-0.009	-0.003	-0.005	-0.012	199301010000	200912310000	199301010000	200912310000	199301010000	200912310000	199501010000	200012310000
V04MOL036110	32.562	0.579	0.626	0.559	-0.03	0.676	0.787	0.668	0.344	0.005	0.004	-0.005	0.013	200101010000	201312310000	200101010000	201312310000	200101010000	201312310000	200801010000	201312310000
V04MOM037100	67.301	0.761	0.402	0.67	0.59	0.731	0.755	0.653	0.705	0.009	0.003	0.011	0.008	200101010000	201312310000	200101010000	201312310000	200101010000	201312310000	200901010000	201312310000
V07BEL285070	88.642	0.6	0.52	0.707	0.618	0.631	0.722	0.699	0.758	-0.014	0.004	0.011	0.004	200101010000	201312310000	200101010000	201312310000	200101010000	201312310000	200001010000	200412310000
V07MAR289015	173.909	0.751	0.585	0.688	0.438	0.62	0.704	0.532	0.565	-0.011	-0.026	0.004	-0.017	200101010000	201312310000	200101010000	201312310000	200101010000	201312310000	200901010000	201212310000
V07MOE282100	46.367	0.532	0.607	0.511	0.44	0.737	0.669	0.709	0.688	-0.025	-0.063	-0.005	-0.011	199701010000	200912310000	199701010000	200912310000	199701010000	200912310000	200501010000	200912310000
F05LEI386999	2981.78	0.778	0.8	0.794	0.457	0.731	0.784	0.803	0.607	-0.015	-0.02	-0.013	-0.018	200101010000	201312310000	200101010000	201312310000	200101010000	201312310000	200301010000	201312310000
V05HEU403210	91.912	0.764	0.758	0.761	0.193	0.673	0.712	0.662	0.629	-0.036	-0.045	-0.029	-0.058	200101010000	201312310000	200101010000	201312310000	200101010000	201312310000	200201010000	200712310000
V05MAN401230	258.442	0.688	0.812	0.748	0.329	-0.023	0.831	0.757	0.508	-0.298	0.001	0.005	-0.05	198301010000	199512310000	198301010000	199512310000	198301010000	199512310000	198801010000	199212310000
F06BOS325001	5217.586	0.665	0.582	0.718	0.467	0.612	0.672	0.626	0.601	-0.046	-0.048	-0.045	-0.026	200201010000	201312310000	200201010000	201312310000	200201010000	201312310000	200501010000	201312310000
V06ZWA342190	112.118	0.647	0.693	0.489	0.469	0.68	0.712	0.617	0.647	0.003	-0.033	-0.01	-0.028	200001010000	201212310000	200001010000	201212310000	200001010000	201212310000	200801010000	201212310000
V08BAR111370	70.08	0.761	0.717	0.742	0.562	0.801	0.802	0.754	0.725	-0.001	0.014	0.015	0.025	199701020000	200411040000	199701020000	200411040000	199701020000	200411040000	199801010000	200312310000
V08DIJ093400	861.413	-3.412	0.61	0.606	0.329	-0.470	0.618	0.157	0.394	-0.971	0.025	0.017	-0.011	201301010000	201504080000	201301010000	201504080000	201301010000	201504080000	197401010000	198912310000
V08ZUU233100	64.771	0.621	0.487	0.627	0.548	0.571	0.637	0.542	0.592	-0.031	0.009	-0.005	-0.055	200101010000	201312310000	200101010000	201312310000	200101010000	201312310000	199601010000	201212310000
W08SAMRON000	134.097	0.635	0.438	0.693	0.739	0.551	0.696	0.684	0.616	-0.098	0.056	0	-0.005	199801010000	201012310000	199801010000	201012310000	199801010000	201012310000	200501010000	201012310000
W08SENTUB030	215.911	0.652	0.744	0.67	0.711	0.684	0.768	0.621	0.581	-0.007	-0.021	-0.002	-0.015	200101010000	201312310000	200101010000	201312310000	200101010000	201312310000	200701010000	201112310000
V09DEM136000	255.882	0.686	0.636	0.12	-0.477	0.647	0.732	-0.106	0.454	0.014	-0.001	0.407	0.078	199801010000	201012310000	199801010000	201012310000	199801010000	201012310000	200501010000	201012310000

Modelname	Area	NS				NS_log				ReError			
		NAM	PDM	VHM	WETSPA	NAM	PDM	VHM	WETSPA	NAM	PDM	VHM	WETSPA
V09GET152080	800.395	0.514	0.672	0.712	0.274	0.459	0.604	0.627	0.276	-0.179	0.007	-0.004	-0.066
V09HER163010	274.602	0.462	0.669	0.605	0.459	0.122	0.643	0.54	0.695	-0.318	-0.02	-0.003	-0.035
V09MAN161040	103.081	0.627	0.597	0.666	0.577	0.505	0.593	0.616	0.359	-0.005	0.003	-0.001	-0.01
V09MOT144270	33.59	0.685	0.621	0.645	0.45	0.519	0.572	0.507	0.426	0.07	0.013	0.025	0.06
V09VEL145100	96.801	0.675	0.547	0.724	0.493	0.708	0.747	0.658	0.562	0.004	0.032	0.006	-0.005
V09WIN141310	64.739	0.738	0.522	0.677	0.25	0.667	0.507	0.58	0.603	-0.075	0.043	0.017	-0.035
V09ZWA148120	96.515	0.604	0.711	0.637	0.556	0.481	0.563	0.481	0.396	0.058	0.028	0.034	0.036
V10GNE076999	359.885	0.705	0.755	0.714	0.276	0.651	0.738	0.679	0.464	-0.016	-0.033	-0.012	-0.006
V10KNE052000	584.669	0.799	0.78	0.756	0.568	0.808	0.792	0.72	0.715	-0.007	-0.004	0	-0.003
V10MOP062140	77.319	0.591	0.451	0.614	0.037	0.578	0.668	0.659	0.45	-0.049	0.026	0.014	-0.033
F11MAA8702	10132	0.802	0.675	0.695	0.427	0.81	0.625	0.761	0.747	-0.02	-0.039	-0.059	0.001
W11BER551010	128	0.62	0.633	0.623	0.555	0.624	0.66	0.645	0.581	-0.007	-0.006	0.003	-0.005
W11MAAPROF	12585	0.738	0.738	0.72	0.667	0.66	0.66	0.776	0.622	-0.032	-0.032	-0.006	-0.016
W11MEH5820	355.8	0.565	0.61	0.604	0.308	0.396	0.765	0.574	0.463	-0.123	-0.016	-0.027	0.098
W11OUR5805	3621	0.778	0.836	0.706	0.391	0.782	0.847	0.803	0.6	0.006	-0.003	0.006	-0.004
W11SAM7319	2669	0.699	0.804	0.728	0.594	0.699	0.774	0.636	0.614	0.049	0	-0.005	-0.007
Total Good		33	30	37	6	28	38	31	18	37	42	41	42
Total Medium		8	12	4	26	10	4	9	22	3	0	0	0
Total Poor		1	0	1	10	4	0	2	2	2	0	1	0

Modelname	NAM		PDM		VHM		WETSPA	
	Start Calib	End Calib	Start Calib	End Calib	Start Calib	End Calib	Start Calib	End Calib
V09GET152080	200101010000	201312310000	200101010000	201312310000	200101010000	201312310000	200001010000	201312310000
V09HER163010	200301010000	201312310000	200301010000	201312310000	200301010000	201312310000	200801010000	201312310000
V09MAN161040	199801010000	201012310000	199801010000	201012310000	199801010000	201012310000	200501010000	200912310000
V09MOT144270	199701010000	200712310000	199701010000	200712310000	199701010000	200712310000	200301010000	**201012310000**
V09VEL145100	200101010000	201312310000	200101010000	201312310000	200101010000	201312310000	200501010000	200912310000
V09WIN141310	200101010000	201312310000	200101010000	201312310000	200101010000	201312310000	200801010000	201312310000
V09ZWA148120	200101010000	200712310000	200101010000	200712310000	200101010000	200712310000	200201010000	200712310000
V10GNE076999	200301010000	201312310000	**200301010000**	201312310000	**200301010000**	201312310000	200801010000	201312310000
V10KNE052000	200101010000	201312310000	200101010000	201312310000	200101010000	201312310000	200801010000	201312310000
V10MOP062140	199701010000	201312310000	199701010000	201312310000	199701010000	201312310000	200501010000	201312310000
F11MAA8702	200201010000	201312310000	200201010000	201312310000	200101010000	201312310000	200701010000	201312310000
W11BER551010	199401010000	200612310000	199401010000	200612310000	199401010000	200612310000	200001010000	200412310000
W11MAAPROF	200201010000	201312310000	200201010000	201312310000	200101010000	201312310000	200701010000	201312310000
W11MEH5820	200401010000	201312310000	200401010000	201312310000	200401010000	201312310000	200101010000	200801310000
W11OUR5805	200101010000	201312310000	200101010000	201312310000	200101010000	201312310000	200301010000	201212310000
W11SAM7319	200701010000	201212310000	200701010000	201212310000	200701010000	201212310000	200701010000	201212310000

DEPARTMENT **MOBILITY & PUBLIC WORKS**
Flanders hydraulics Research

Berchemlei 115, 2140 Antwerp
T +32 (0)3 224 60 35
F +32 (0)3 224 60 36
waterbouwkundiglabo@vlaanderen.be
www.flandershydraulicsresearch.be