UNIVERSITY of TASMANIA

UNIVERSITY OF TWENTE.

Comparison of rainfall runoff models for the Florentine Catchment

Jan Gert Rinsema

25-11-2014

Comparison of rainfall runoff models for the Florentine Catchment

Author: Jan Gert Rinsema
S1129988
j.g.rinsema@student.utwente.nl
Bachelor Civil Engineering University of Twente (B-CiT)
Bachelor Thesis University of Tasmania

Supervisor University of Tasmania:
Prof. Stewart Franks

Supervisor University of Twente:
Dr. M.M. Mekonnen, MSc

Date: 25 – 11 – 2014

Picture front page: http://www.discovertasmania.com.au/attraction/lakemeadowbank

Summary

Entura is the engineering company of Hydro Tasmania which produces electricity in Tasmania, Australia. Most of the electricity is produced by hydro power plants. The lakes of the hydro power plant is filled with the rainfall from the area. To be able to us the power plant optimal, it is needed to predict the inflow in the lake. The prediction of the rainfall is done by rainfall runoff models employed by Entura. Historically four models are used: AWBM, GR4J, SYMHID and the VIC model. The choice of the model is based on experience and is not uniform for all the catchments and can lead to the wrong selection of a model for a catchment. This research tries to avoid the selection of not the best model by creating a stepwise model for the selection of a rainfall runoff model for a single catchment. This is done by selecting a rainfall runoff model for the Florentine catchment in Tasmania, Australia.

To be able to evaluate the rainfall runoff models after calibration, six evaluation methods have been evaluated. The Nash Sutcliffe coefficient, the Index of Agreement and the Coefficient of determination evaluate the shape of the simulated runoff, compared with the observed runoff. The NS coefficient evaluates the shape best compared with the observed runoff. The Nash Sutcliffe coefficient is the most important factor of the proposed evaluation method. The Nash Sutcliffe method evaluates the shape best for both low flow and peak flow. The RVE evaluates the difference between the simulated and observed volume best and is the second important element of the evaluation method which should be included because the volume of water is important for watermanagement challenges handled by Entura. The evaluation used is:

$$Ev = 0{,}75 \times NS \; + \; 0{,}25 \times \left(1 - \left|\frac{RVE}{100}\right|\right)$$

The calibration is done with treating the initial values with different methods. The initial value is treated as half full as parameter and is decreased with a startup time. The best results were achieved by treating the initial values as a parameter. Still it is advised to use a startup time because there is still the possibility of over- or under parameterization of the initial conditions. With a startup time this is reduced to minimum. When there is not enough data to include a startup time, the initial conditions can be treated as parameter in the Monte Carlo analysis.

To decrease the influence of the input error, the rainfall gages were calibrated by Monte Carlo analysis. The results are mixed and do not give convincing improvements in calibration and validation for the Florentine Catchment.

According to the comparison and evaluation of the AWBM, SYMHID and GR4J model the GR4J is the less sensitive model with the best output. The procedure which lead to the selection of the best result for the rainfall runoff model is recommended to Entura to Use in the future.

The stepwise model proposed is:

1. Check if the lumped model is sufficient for the catchment.
2. Create a Monte Carlo analysis for GR4J, AWBM and SYMHID with 10000 runs
    a. Use a startup time if possible
    b. Otherwise treat the initial conditions as parameter
3. Evaluate the parameter sets with the formula of Ev.
4. Validate the models with a Split sample test.
5. Compare the results of the models and select the best model

Preface

This bachelor thesis in civil engineering is not done only by myself. I'd like to tank some people for their support and effort they made to help me in completing the bachelor thesis. First I'd like to thank Prof. Stewart Franks for the opportunity to do the bachelor thesis at the University of Tasmania in Hobart Australia.

Second I'd like to tank Fiona Ling for her cooperation during the bachelor thesis. She gave the data which gave me the opportunity to do the bachelor thesis as planned.

At last I'd like to thank Mesfin Mekonnen from the University Twente. He supervised me during the period of my thesis. He gave me the freedom to do what I wanted to do for which I am thankful.

Contents

1. Introduction

The State of Tasmania is the smallest states of Australia. It has a population of 5134000 (Department of Treasury and Finance) people within an area of 68000 km^2 (Greenpeace, 2004). The state has currently 30 reservoirs in the state of which some of them produce electricity through hydroelectric power plants, as shown in Figure 1.2-1. The total amount of electricity produced is 7653 GWh in the period 1 July 2012 till 30 June 2013, and can be seen as average. The hydro electrical power plants have a total capacity of 2270 MW (Crean & Davy, 2013). The total output of electricity was 64,8 % of the total amount of electricity produced in the state of Tasmania. A net amount of 2040 GWh is exported to different states at the mainland of Australia, for example New South Wales and Victoria. In that same year, the total amount of consumption of electricity in the total state of Australia was 198,2 Terrawatt hours (Regulator). The reservoirs also contain water for irrigation, town water supply, aquaculture, recreational use, aquatic habitat and so on.

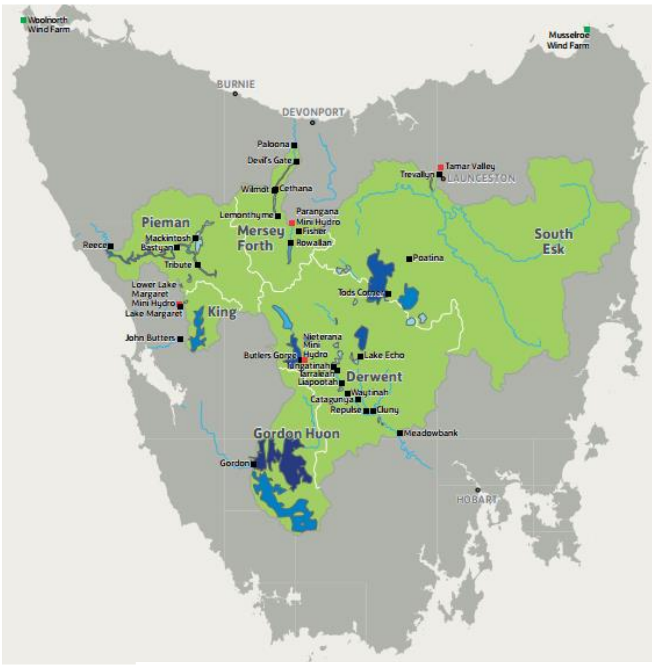


FIGURE 1.2-1 HYDRO ELECTRIC POWER PLANTS IN TASMANIA

The hydroelectric reservoirs are managed by Hydro Tasmania. Hydro Tasmania is Australia's leader in renewable energy in Australia and the development of renewable energy (HydroTasmania, About us, sd). To predict the level of the storage to create electricity, rainfall-runoff models are used. These models predict with a forecasted amount of rainfall the runoff for the catchment for which it is used.

Entura is the consultancy business of Hydro Tasmania. They calculate the runoff predictions for several catchments where hydro power plants are located. Entura has more than 100 years of experience with the water management of reservoirs. Based on that experience, they use 4 models to predict the runoff with forecasted rainfall. The four models employed are the GR4J model, AWBM, SYMHID and the VIC model. Only one model is used to predict the runoff for one catchment.

There is not a single procedure to select a model for a catchment which has to be calibrated by Entura. In some cases GR4J is used to model a catchment for example, but AWBM may give better results for runoff. The difference between the models is the way they predict the runoff given a certain rainfall. The number of parameters which has to be calibrated is different, so the level of detail is different. Some catchment have such a difficult natural structure, it is impossible to create the catchment with a detailed model (Perrin, Andréassian & Michel, 2003). A more simple structure will be able to predict the runoff of the catchment better. Next to the level of parameters, the mathematical basics for describing the process can also be different. Some models have a linear function for describing the storage where other models have non-linear functions to describe the storage (Wood, Lettenmaier, & Zartarian, 1992).

Better use of the models will be more important now Entura is more hived off Hydro Tasmania (Crean & Davy, 2013). This means Entura will also do other kind of water management projects and projects in other parts of the world in which not Hydro Tasmania is the client who will use the model but an external company will use the model for their purpose.

To be able to select the best model, a uniformed calibration and evaluation tool will be developed which can be followed by Entura. In the future Entura will pick the best model available for the catchment which helps them to model the catchment better and so the managers of the reservoirs will be able to operate the hydro power plant more efficient.

To create a simple and stepwise procedure for selecting a single model, the Florentine Catchment will be calibrated and evaluated for the models. The models selected will be the models which have the least similarities so the procedure can be used for all four models.

The Florentine River ends in The Meadowbank hydro storage is one of the hydroelectric power plants. The plant has a capacity of 40 MW and is used since 1967 (HydroTasmania, Meadowbank Power Station). The Florentine Catchment a large catchment which ends in the Meadowbank hydro storage to produce electricity. For the Florentine catchment, the AWBM is employed to predict the inflow.

1.2 Problem statement

There is not a uniform method for Entura to compare the four rainfall-runoff models employed. The selecting of the model is done based on experience. This can cause the choosing of the second best or even the worst model because there is more experience with one model than the other.

1.3 Research objective

The main objective of this research is to create a simple and stepwise evaluation procedure for GR4J, AWBM, SYMHD and the VIC models.

1.4 Research questions

With regard to the research objective, the following questions have been drafted:

Main question:
Which rainfall-runoff model does describe the rainfall – runoff relation best for the Florentine Catchment?

Sub questions:

1. What is the structure of the used models?
2. Which models are best to be created and evaluated?
3. Which evaluation method will be used for the evaluation of the selected models?
4. What is the accuracy of the selected models?
    a. Which model is calibrated best according to the evaluation methods?
    b. Which model describes the validation period best?
    c. Can the calibration/validation be improved by minimizing input errors?

2. Theoretical framework

The theoretical framework of the rainfall runoff models will be explained briefly in this chapter. First the basic theory of rainfall runoff models will be explained in chapter 2.1, and the theory of the AWBM model will be discussed in chapter 2.2, the GR4J model in chapter 2.3, the SYMHID model in chapter 2.4 and the VIC model in chapter 2.5

2.1 Rainfall Runoff basic theory

Before the models for calibration are selected, it is necessary to understand the basic principle of rainfall- runoff models. Models exist for many years and can be divided in four types of models (Jothiprakash & Magar, 2009):

- Empirical models
- Conceptual models
- Data driven models
- Physically based models

Empirical models are models which are based on experience of the natural behaviour of the catchment in the past like the Curve Number method (Shaw, Beven, Chappel, & Lamb, 2011). Conceptual based models have a conceptual idea of the behaviour of the soil and runoff in a catchment. They have to be calibrated for a single catchment. Data driven models are employed if there is not enough data available. The behaviour is estimated based on data collected from satellites and field work on the ground. Physically based models are based on the physical characteristics of the catchment. The physical characteristics focus on the more physical aspects of the catchment and try to estimate for each grid of the catchment. The majority of the rainfall runoff models cannot be categorized in one of the four model types above because the models contain more than one element from the four types. This leads to another type of model: the composed model. The models evaluated in this research are conceptual models, except for the VIC model which is a physically based model.

Conceptual based models have all conceptual ideas about the behaviour of the nature in the catchment but also have differences between them. The first difference is the difference between lumped and distributed models (Beven K. J., 2012, p. 16). Lumped models treat the catchment as a single unit with state variables which represent averages over the catchment area such as average storage in the saturated zone. Distributed models predict runoff that are distributed in space, with state variables which represent local averages of storage, flow depths or hydraulic potential. By discretising the catchment into a large number of elements or grid squares and solving the equations for the state variables associated with every element grid square the predicted runoff can be more specific according to the difference in physical aspects. Parameters values must be specified for every element in a disturbed model. This also increases the difficulty of the model, which is why the models are not used widely. Many variations of distributed and lumped models exist. One of the models most used is TOPMODEL (Beven & Kirkby, 1979). TOPMODEL calculate results for every grid. It has the feature that the predictions can be mapped back into space for comparison with any observations of the hydrological response of the catchment.

According to the difference between lumped and distributed, there is a difference between modelling the catchment as a whole or divide the catchment in sub areas. When the catchment is divided, the model will also be divided in a model with specific model parameters for each catchment. The rainfall in the divided catchments will runoff through a routing channel to a joint runoff channel. The flooding of the routing capacity will be described by a function or a unit hydrograph, which also can be describe both linear and non-linear. The advantage of using a routing channel is the ability to describe the runoff of the rainfall over several days. This model is usual used when the catchment is used in several semi-catchments, but can also be used to describe the runoff of the groundwater flow and surface flow and the time lag between the runoff of these flows. The runoff from each catchment will end in the routing channel, which in total will describe the runoff for the whole catchment.

Another difference is the difference between the storage capacities of the models is the storage inflow and outflow. The inflow and outflow can be described by a function which can be both be linear or non-linear. A linear model will create runoff according to a partial quantity of the store capacity. A nonlinear model will produce more runoff when the store capacity is more filled according to a determined function.

The last difference in model structure is the model in- and output. The choice has to be made between deterministic and stochastic output. Deterministic models permit only one outcome from a simulation with one set of inputs and parameter values. Stochastic models allow some randomness or uncertainty in the possible outcomes due to uncertainty in input variables, boundary conditions or model parameters.

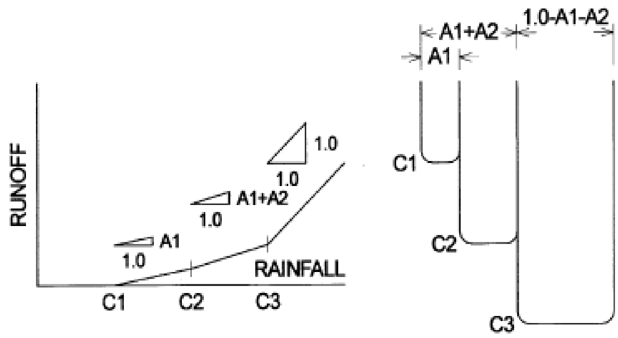
During calibration, several problems can occur. Because the used models are lumped models, the catchment characteristics will be set equal for the whole area, which is not very realistic. This can cause over- or under parameterization for the model parameters. Next to the characteristics of the model the measurements are limited. For most of the catchments only \ evapotranspiration, rainfall and the observed discharge is measured. This simplifies the natural processes but still can give good results. Also the processes in the soil are simplified by the models. All these problems are difficult to solve. Another input error is the dividing of the catchment in rainfall gages. The measurement points are not equal divided over the catchment. The weight of their measurement of the catchment can be determined in several methods. One method is the Thiessen polygon method (Shaw, Beven, Chappel, & Lamb, 2011, p. 168). In the Thiessen polygon method, the measurement points will be connected by lines so the area is divided into rainfall gages corresponding to a measurement point. To be able to apply this method, the exact boundaries of the catchment have to be known are estimated. This can also be a time spending process to draw and calculate the areas by hand or with the computer. This is a large effort for probably a small improvement. Because it is time spending mostly the average of the number of measurements is used. To try to improve the problem of the rainfall gages in the catchment, the rainfall gages also will be calibrated to see if there is any improvement.

2.2 AWBM

The Australian watershed based model (AWBM) is a lumped conceptual catchment water balance model introduced by Boughton (Boughton, 1993). The model calculates hourly runoff for design flood estimation and the daily version is used for water yield and management studies (Broughton, 2004). The daily version is used in this report.

The concept is based on saturation overland flow generation of runoff and is divided in the single bucket model with multi capacity models. When the AWBM is used as a single bucket model, it is used as a single lumped model which has one bucket for the whole catchment. For a small area it is possible to use the single bucket model, but for larger areas there is a risk of under parameterization. Because the whole catchment with all the differences is described in one parameter, it will become an average of the whole area. It is always more useful to use a multi- capacity models because it probably will give better results.

FIGURE 2.2-1 MULTI-CAPACITY STRUCTURE



The simplest form of the multi-capacity is a catchment with three capacities C1, C2 and C3 and partial area A1, A2 and A3 (A1+A2+A3 =1), as shown in figure Figure 2.2-1. The runoff begins after the smallest capacity is filled with rainfall. Until the rainfall is enough to fill the second capacity, the rainfall-runoff relationship from the smaller store will be a line with slope A1, vertical to 1.0 horizontal. If both capacities are filled, further rain will be runoff so the angle is 45 degrees. When the 45 degrees line is extended backwards the intersect with the x-axis is the average storage capacity Ave. Ave = C1*A1 + C2*A2 + C3*A3. If the catchment has more capacities, the line gets smoother and more like a curve. For the Florentine Catchment three capacities are used to subscribe the storage in the catchment.

Base flow (K_b) is usually modelled as the drainage from a single store. The recession of base flow is commonly linear when the logarithm is plotted against time. The fraction (1.0 - K_b) is the amount of water in the base flow store that is discharged every step in time. In the AWBM the base flow recharge is a fraction of the stream flow, the fraction is called the base flow index (BFI).

The structure of AWBM is shown in Figure 2.2-4. This can be used for both hourly and daily runoff. If there is any excess from any store, it becomes runoff divided between surface runoff and base flow. The P stands for rainfall and E for the Evaporation during the modelled time.

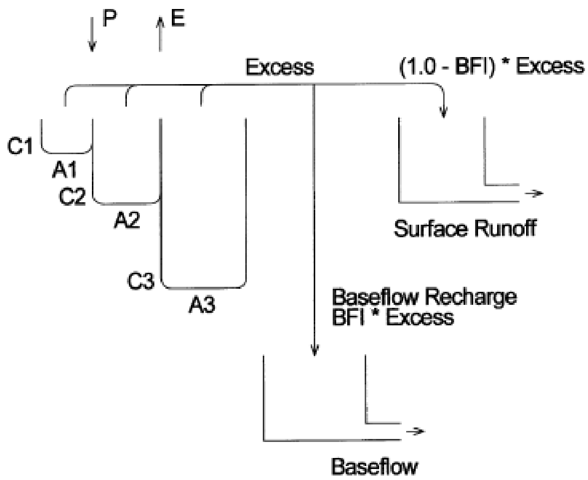


FIGURE 2.2-4 STRUCTURE AWBM

The mathematical equations for the AWBM are described in appendix 9.2.1.

The AWBM is a semi non-linear model. The curve for the runoff (depending on the three capacities) divided in three linear parts, as shown in Figure 2.2-1. The model has no routing or hydrograph included by Boughton in the single catchment version. If the larger catchment is divided in sub-catchments, it is necessary to use routing. The parameters which have to be optimized are the base- and surface flow fraction, the base flow index and the storage capacity. Boughton has provided a fixed solution for the length of the partial areas in a three bucket capacity store A1, A2 and A3. With the fixed partial area, Boughton also suggested a fixed partition of the capacity's C1, C2 and C3. The volume of the capacity is depending of the average capacity of the three capacities Ave = C1*A1 + C2*A2 + C3*A3. The optimised average and limits of the parameters are given in table 2.2-1.

TABLE 2.2-1 LIMITS AWBM

	LIMITS	OPTIMISED AVERAGE
C_1		7.5
C_2		76
C_3		152
C_{AVE}		100
A_1	0 – 1	0.134
A_2	0 – 1	0.433
A_3	0 – 1	0.433
K_B	0 – 1	
K_S	0 – 1	
BFI	0 – 1	

2.3 GR4J model

The second model used by Entura is the GR4J model (which stands for modele du *G*énie *R*ural a 4 parameters *J*ournalier). This model is developed in France by Charles Perrin et. all and introduced in 2003 (Perrin, Andréassian & Michel, 2003). The introduced GR4J model is an improvement of the earlier introduced GR3J model, first introduced by Edijatno et. all. (Edijatno, Oliveira Nascimento de, Yang, Makhlouf, & Michel, 1999). The number of degrees of freedom in terms of parameters is a difficult optimization problem. Too few will prevent the model from being sufficiently flexible, too many will lead to problems of parameter definitions and model robustness. Also more mathematical functions in the model means more parameters. The GR4J has four free parameters instead of 3 free parameters in the GR3J model. The basis of the GR4J model is a lumped model.

The model description is shown in Figure 2.3-1. The rainfall input is given by P and the evapotranspiration is given by E. P is an estimation over the whole catchment, while E is a long-time average, so this can be repeated every year. The next step is to determine either a net of rainfall or a net evapotranspiration. This operation is computed as if there were an interception storage of zero capacity. A part of the rainfall will end in the production store. Percolation will infiltrate from the production store into the ground and will be described by Pr. Pr will reach the routing function wherein Pr is divided in two flow components according to a fixed split. UH1 will be 90% of the described percolation by a unit hydrograph and then a routing store. The first Unit Hydrograph slow the water down and will runoff the largest part of the water in the last day of the hydrograph. The remaining 10% will be routed by another unit hydrograph which is twice the length of UH1. The time lag until the peak is described by parameter x_4 and is described in days. The water will increase until the peak after x_4 days. After the peak the water outflow will decrease until it is zero again. With UH1 and UH2 the model is able to simulate the time lag between rainfall and the resulting stream flow peak. The rainfall of UH1 is divided over x_4 days, the rainfall of UH2 is divided over $2x_4$ days, see Figure 2.3-2.

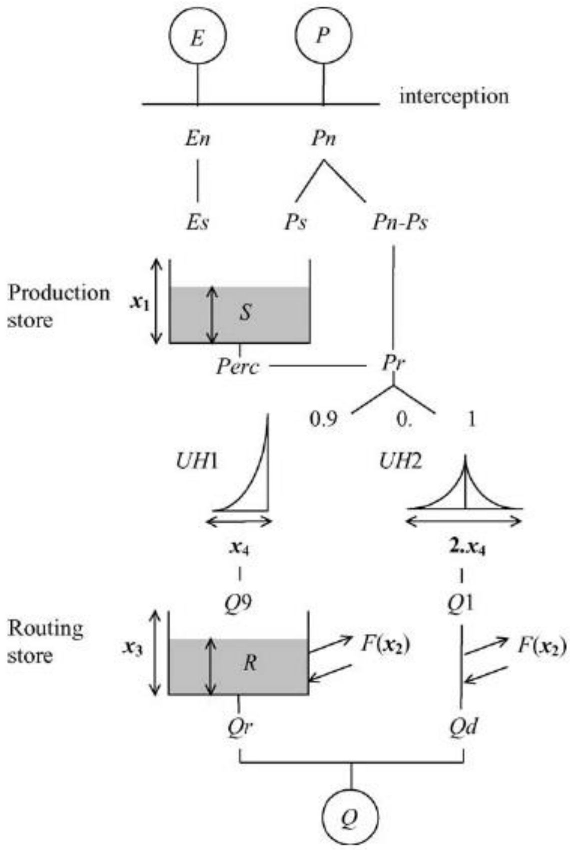


FIGURE 2.3-1 STRUCTURE GR4J

There is also groundwater exchange possible in the routing store. The groundwater exchange (depending of variable x_2) can either extract or add water to the routing store.

The four free parameters which have to be optimized are:
X_1: maximum capacity of the production store (mm)
X_2: groundwater exchange coefficient (mm)
X_3: one day ahead maximum capacity of the routing store (mm)
X_4: time base of unit hydrograph (days)

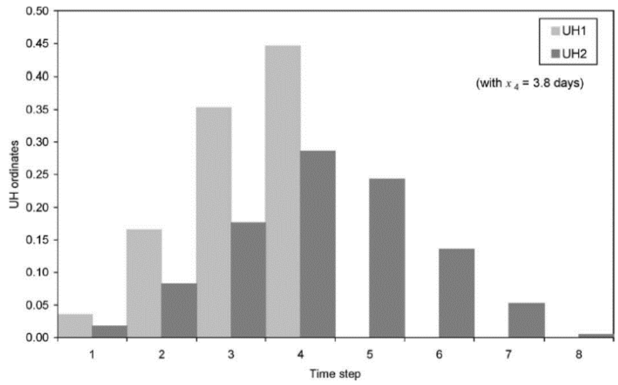


FIGURE 2.3-2 EXAMPLE FOR UH1 AND UH2

The model is a nonlinear model with routing. Both production- and routing store are non-linear, based on a function which is described in appendix 9.2.2.

To be able to calibrate the model, it is necessary to know the interval of the values. The 80% of confidence interval is shown in table 2.3-1.

TABLE 2.3-1 PARAMETER LIMITS OF GR4J

	MEAN VALUE	80 % OF CONFIDENCE INTERVAL
X1 (MM)	350	100 - 1200
X2 (MM)	0	-5 to 3
X3 (MM)	90	20 - 300
X4 (DAYS)	1.7	1.1 - 2.9

2.4 SYMHID

The third rainfall-runoff model used is SYMHID and introduced by Chiew et. all (Chiew, Peel, & Westernd, 2002). SYMHYD is a lumped conceptual daily rainfall-runoff and is one of the most common used rainfall-runoff models in Australia.

The SYMHID model can be used with 5 or 7 ‘free’ parameters which have to be calibrated. The difference between the 5 and 7 parameter version of the model is the including of the infiltration loss described by parameters COEFF and SQ. The model used in this research is the 7 parameter SYMHID model like Figure 2.4-1.

The rainfall falls in the inception store of the model. A part of the rainfall will evaporate while another part of the water will flow to store F. The amount of runoff INR will be the lesser of the maximum store capacity and the evapotranspiration. From store F IRUN will runoff to the open water channel. From the surface store S SMF will flow to the soil moisture store and SRUN will flow to the open water channel to runoff. From the soil moisture store discharge will flow from the soil moisture to the groundwater store. From the groundwater store base flow will runoff in an open water channel. The stores F and S are imaginary and describe together with the excess runoff from the infiltration and saturation. The runoff will be the total of the surface flow IRUN, interflow SRUN and the base flow BAS. The calculation of every step is shown in figure Figure 2.4-1 Structure SYMHID and also in appendix 9.2.3.

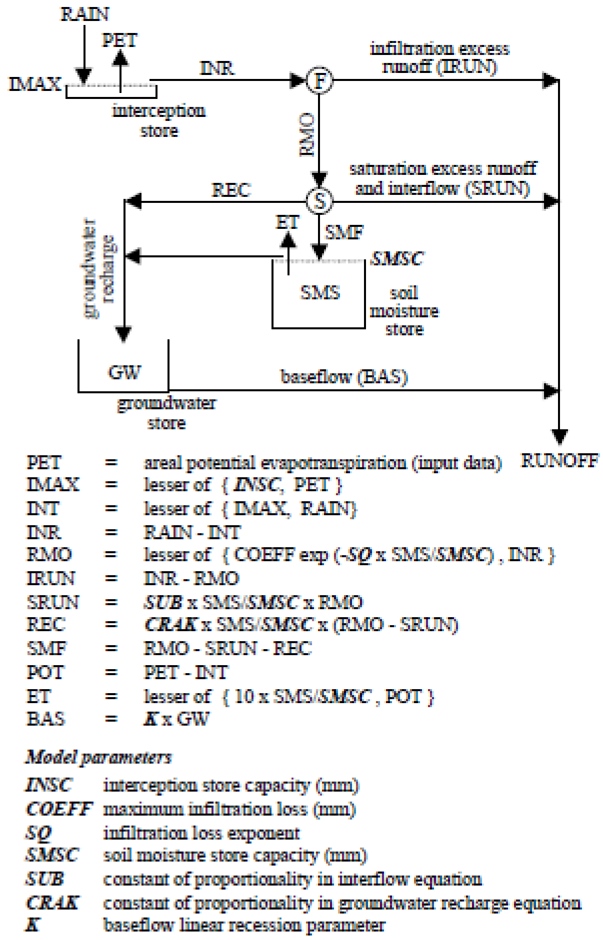


FIGURE 2.4-1 STRUCTURE SYMHID

The model is linear, because every time step (days for daily runoff, hours of hourly runoff) the saturation excess runoff and interflow will be calculated by the percentage of the soil moisture capacity filled. The function of the soil moisture capacity set out against the maximum capacity of the soil moisture store.

Table 2.4-1 contains the limits for the ‘free’ parameters used in in the SYMHID model. For the infiltration loss exponent and the maximum infiltration loss coefficient the limits are not known. Only the optimised average is known for these parameters. These parameters will in first instance not be optimized, but will have their optimised average value.

	LIMITS	OPTIMISED AVERAGE	10^{TH} – 90^{TH} PERCENTILES OF OPTIMISED PARAMETERS VALUES
INSC (MM)	0.5 – 5.0	3.2	0.9 – 5.0
SMSC (MM)	50 – 500	195	100 – 400
SUB	0 – 1	0.31	0.19 – 0.62
CRAK	0 – 1	0.11	0 – 0.66
K	0.003 – 0.3	0.12	0.02 – 0.3
SUB		1.5	
COEFF		200	

TABLE 2.4-1 LIMITS OF SYMHID

2.5 Vic Model

The Vic (variable infiltration capacity) model is the last of 4 models which is evaluated. The VIC model is a non-linear data physically driven model and calculates the rainfall runoff for every grid (Xu, Lettenmaier, Wood, & Burges, 1994). The Vic model is different from the more conceptual previous models, because a lot of data has to be determined by fieldwork. The model assumes that infiltration capacities and therefore runoff generation and evapotranspiration vary within an area due to vegetation in topography, soil and vegetation. The three parameters for the VIC model are the variable evapotranspiration, variable infiltration capacity and the base flow recession coefficient. In this case we describe a VIC model with two layers, as shown in Figure 2.5-1.

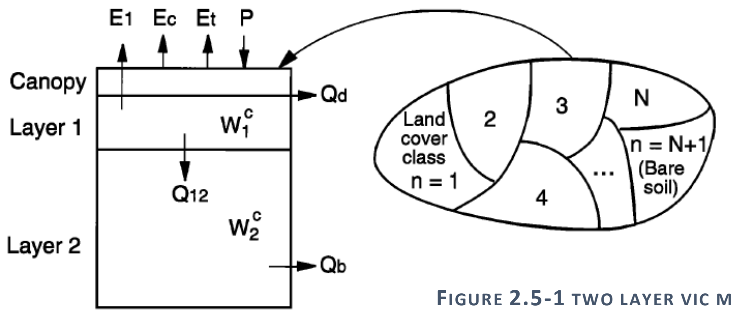


FIGURE 2.5-1 TWO LAYER VIC MODEL

In the model the surface is described in N+1 land cover types where n represent different types of vegetation, which will be specified by the leaf area index (LAI), canopy resistance and relative fraction of roots in each of the two soil layers. The number of vegetation's will always be less than 10 because this is determined by Xu et all. The evapotranspiration is dependent of the vegetation on the soil layer, together with the canopy resistance, aerodynamic resistance to transfer of water and architectural resistance. Beneath the canopy there are two different soil layers, the upper zone (layer 1) and the lower zone (layer 2). The lower layer can only respond to rainfall if the upper layer is saturated. Three types of evapotranspiration are considered in the model: evapotranspiration from canopy layer, from vegetation and from the soil layer. The total evapotranspiration is the sum of the three sorts of evapotranspiration (Gao, et al., 2009).

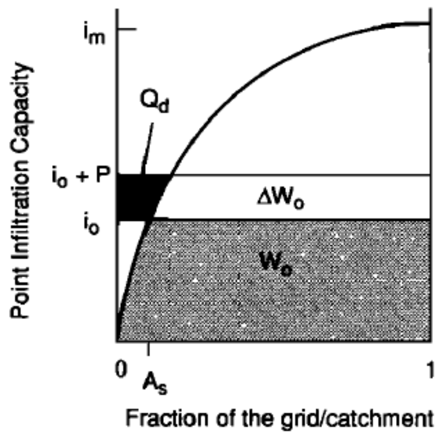


FIGURE 2.5-2 INFILTRATION CAPACITY

Figure 2.5-2 shows the variable infiltration capacity curve with the maximum infiltration capacity i_m. The maximum infiltration capacity means the depth of water that can be stored in the soil column for the incremental area fraction of dA. The shape of the curve is determined by curve variable B, which is dependent of the (semi-) catchment. A_s is the fraction of the capacities which is filled (saturated). W_0 is the depth of the soil moisture capacity of one grid, W_c is the maximum soil moisture storage over the area. The fraction A_s will generate the runoff Q_d because this area is already saturated. The infiltration capacity of the saturated area is i_0. The amount of water P which cannot infiltrate in the grid because the grid is saturated will eventually runoff.

The last variable is the base flow recession coefficient k_b. The variable infiltration coefficient is a nonlinear parameter. Based on large areas and dry periods the base flow can be considered a linear flow based on recession coefficient k_b. The value of the recession coefficient is between 0 and 1. The model uses a linear routing model to simulate the runoff of the whole catchment. Because the Vic Model is mostly data-driven no limits are determined by the Xu et. all. Calibration for the Vic model without having the data will be very hard to do.

3. Research Method

The methods will be explained which will lead to the answers for the research questions in this chapter. First the catchment data will be discussed in chapter 3.1, the evaluation methods to evaluate rainfall runoff models is discussed in chapter 3.2 and the method for the model comparison is discussed in chapter 3.3.

3.1 Catchment data

The rainfall data used to compare the rainfall runoff models is rainfall data from the Florentine Catchment. The data is given by Entura, who control the measurement points in the Florentine Catchment for some years. The rainfall is registered at 6 observation points: Dee lagoon, Ouse River, Mount Tim Shea, Florentine Valley, Clarence River and at the Misery Platea. The runoff is registered at two points: Above Derwint River and at the bridge at Eleven Road. A map of the catchment with the observation points is set in chapter 9.1. The observation points Misery Platea, Florentine Valley and Mount Tim Shea will be used for the Florentine Catchment, because they are located in the Florentine Valley. The other measurement points will be ignored for the calibration and evaluation because they only give general information about the rainfall in the area of the Florentine Catchment, but not of the catchment.

The data for the observation points differs in starting date. Next to the different starting dates a lot of data is missing. The missing data differs in time for each observation point. Eventually only the years 2003, 2011 and 2012 have data all year available. For calibration the year 2011 is used and for the validation the year 2012 is used. Several calibration methods are used for initializing the start values of the stores in the model. For the startup method (see chapter 3.3.2) data is used from 20/8/2010 until 31/12/2010.

Evapotranspiration of the catchment is not provided. To create representative output in the catchment, the evapotranspiration of is provided by the Bureau of Metrology of Australia (Meteorology, 2005). The evapotranspiration used is the monthly average of the catchment.

3.2 Evaluation methods

Evaluation is compare the simulated outflow with the observed outflow. To give interpretation to the output given by the calibration method's, evaluation methods have to be chosen. The evaluation methods are listed in this subchapter 3.2.

3.2.1 Nash Sutcliffe coefficient

The Nash and Sutcliffe coefficient gives the difference of the simulated and observed outflow as fraction of the observed outflow (Nash & Sutcliffe, 1970).

The NS coefficient is calculated by:

$$NS = 1 - \frac{\sum_{i=1}^{i=N}[Q_S(i) - Q_O(i)]^2}{[Q_O(i) - \overline{Q_O}]^2} \qquad \text{(Eq. 3.2-1)}$$

Where Q_s of the i^{th} element is the simulated runoff in time and Q_o is the observed runoff at the same point in time. A NS of 1 is a perfect simulation because the simulated and observed runoff is the same. The NS should be as close to 1 as possible. The range for possible NS coefficients is between negative infinity till 1.

3.2.2 Relative volume error

The relative volume error (Biondi, Freni, Iacobellis, Mascaro, & Montanari, 2012) is a method to calculate the percentage of the absolute error between the simulated and observed error.

The RVE coefficient is calculated by:

$$RVE = 100 \times \frac{\sum_{i=1}^{i=N}[Q_s(i) - Q_o(i)]}{\sum_{i=1}^{i=N}[Q_o(i)]} \qquad \text{(Eq. 3.2-2)}$$

Where Q_s and Q_o is the same as in the NS coefficient. A RVE between -5 and 5 is an acceptable simulation, because het volume of the runoff is 95% the same (-5 is 5% volume less simulated than observed, 5 is 5% volume more simulated than observed). A RVE of 0 is a perfect simulation for the volume of the runoff during a given period because the volume of the simulated and observed period are the same.

3.2.3 Peak flow

The peak flow evaluation method will evaluate the peak flow of the simulated runoff and compares the simulated flow with the observed runoff (Madsen, 2000).

The peak flow is calculated by:

$$PF = \frac{1}{M_p}\sum_{j=1}^{M_p}\left[\sum_{i=1}^{n_j}[Q_s(i) - Q_s(\theta)]^2\right]^{1/2} \qquad \text{(Eq. 3.2-3)}$$

Whereby M_p is the number of peak flow events. The peak flow appears to be a flow above a given threshold discharge $Q_s(\theta)$. The value of this threshold can differ for each simulated catchment but is fixed for one calibration and validation period. The value of the threshold is chosen by plotting the runoff and set the threshold manual.

3.2.4 Base flow

The peak flow evaluation method will evaluate the low flow of the simulated runoff and compares the simulated flow with the observed runoff (Madsen, 2000).

The low flow is calculated by:

$$BF = \frac{1}{M_l}\sum_{j=1}^{M_l}[[Q_s(i) - Q_s(\theta)]^2]^{1/2} \qquad \text{(Eq. 3.2-4)}$$

Where M_l is the number of low flow events.

3.2.5 Index of agreement

The index of agreement gives the simulated outflow as a fraction of the Mean square error of the simulated outflow and the absolute error of the observed outflow itself. (Biondi, Freni, Iacobellis, Mascaro, & Montanari, 2012)

The index of agreement is calculated by:

$$D = 1 - \frac{\sum_{t=1}^{N}(Q_s(i) - Q_o(i))^2}{\sum_{t=1}^{N}(|Q_s(i) - \overline{Q_o}| + |Q_o(i) - \overline{Q_o}|)^2} \qquad \text{(Eq. 3.2-5)}$$

The value of the index of agreement is between 0 and 1.

3.2.6 Coefficient of determination

The coefficient of determination is based on the statistical principle of correlation (Biondi, Freni, Iacobellis, Mascaro, & Montanari, 2012). The coefficient of determination measures the correlation between the observed and simulated outflow per time step.

The coefficient of determination is calculated by:

$$R^2 = \left\{ \frac{\sum_{t=1}^{N}(Q_O(i)-\overline{Q_O})(Q_S(i)-\overline{Q_S})}{\left[\sum_{t=1}^{N}(Q_O(i)-\overline{Q_O})^{1/2}\right]^2\left[\sum_{t=1}^{N}(Q_S(i)-\overline{Q_S})^{1/2}\right]^2} \right\}^2 \qquad \text{(Eq. 3.2-6)}$$

The value of the coefficient determination is between 0.0 and 1.0.

3.3 Model comparison

The models described in chapter 2.2 to chapter 2.5 have to be evaluated in performance. The simplest way to evaluate is to code the models in MatLab and evaluate the performance of the model with evaluation methods. As described in chapter 3.2 the evaluation methods use the observed runoff as parameter for the evaluation of the performance.

First the models will be scripted and tested if they work accurate. Before the models can be tested, the parameters have to be estimated. The average for the parameter is given in chapter 2.2 to chapter 2.5 and is sufficient for the accuracy test. The next step in the check is to compare the simulated runoff visual with the observed runoff of the catchment. If the difference is too much, the script will be checked again with the equations set out in appendix 9.2. If the simulated runoff does not differ much in contours with the observed runoff, the model will be tested by a water balance. In the water balance, in each step the volume of water in the model is verified. The model include several stores which contain water and the water level changes every step. The total amount of water can't change other than because of the rainfall, evapotranspiration, runoff or a model parameter (like parameter x_2 in GR4J). If the total water level in the model does change by another course, the model has to be adjusted in a way the model can't change anymore other than rainfall, evapotranspiration and runoff or model parameter.

3.3.1 Comparison of evaluation methods

The evaluation methods described in chapter 3.2 are not new methods. There all used in previous studies, but never have been directly compared to each other. The studies that introduce the evaluation method compare their new model with an already existing one or with the graphical method. To be able to pick the best method or maybe combine evaluation methods, all evaluation methods will be compared with each other with a model used by Entura and be examined for the purpose of Entura which is the prediction of the runoff into the storage.

The evaluation methods will be compared with the GR4J. It is not necessary to test all the evaluation methods with all the models because the evaluation methods evaluate the simulated runoff. It doesn't matter if the simulated runoff is produced by a linear or non-linear model. The GR4J is calibrated by Monte Carlo with the routing- and moisture store half full (see 3.3.2).

The threshold of the low- and high flow evaluation has to be set manual for each calibration/validation period. The mean discharge is 11 m^3/s with high peaks and lower peaks. To be able to evaluate all the peaks the threshold for peak flow is set on 20 m^3/s. The threshold for low flow is set on 6 m^3/s because the below the threshold the low flow only occurs is dryer periods.

There are several methods for the evaluation method. If one evaluation method is superior to the other evaluation methods, the single evaluation method will be chosen. A weighted method with different evaluation methods is also a possibility (Madsen, 2000). A weight is given to a certain methods and the total result will give the best parameter set for the calibration.

If more than one evaluation methods is chosen, the evaluation methods will be combined with equation 3.3-1.

$$Ev = \mathrm{w}_1 \times Eval_1 \; + \cdots + \; w_n \times Eval_n \qquad \text{(Eq. 3.3-1)}$$

Where the outcome Ev will give the best parameter set, w_1 is the weighting factor for evaluation method 1 ($Eval_1$) and w_n the weighting factor of evaluation method n. The value of Ev depends on the individual values of the evaluation methods. The weight has to be determined based on the output of the evaluation method, and so the evaluation methods give the same size of output. When the best parameter will give 1 for the best simulation for the first method and a 3 for the best simulation for the second output, the result of the second method will dominate the result, so the result for the second model also has to be one, or the result for the first model has to be 3.

The evaluation methods will finally be examined by plotting the simulated versus the observed runoff. Every evaluation method will probably chose another best parameter set. If the simulated runoff is perfect, all dots should be on a straight line with an angle with the x-axis of 45 degrees so y=x. The best evaluation method is the method which chooses the best parameter set. If two parameter sets are close to each other, the total simulated volume will be included. Entura uses the rainfall runoff models to predict the rainfall into the dams for hydro electrical engineer or agriculture, so the volume is an important factor for Entura to use a model.

3.3.2 Calibration of the models

If the model is tested and approved by the plot and the water balance, the model can be calibrated. The calibration is done by Monte Carlo analysis. The Monte Carlo gives every parameter a random value between given limits. The limits given to the Monte Carlo are the limits per model given in chapter 2. Every run of the Monte Carlo will create another parameter set, so at the end of the calibration 10000 parameter sets are created.

The Monte Carlo analysis can be done with multiple methods. The difference is the method of determine the start values of the model. Every models has stores to store water. In the AWBM the initial conditions of the initial store, the groundwater store and the surface store have to be estimated. In the GR4J the routing store (x_3) and the initial store (x_1) have to be estimated and in the SYMHID the initial capacities of the Soil moisture store (SMS) and the groundwater storage (GW) have to be estimated. These stores can't be empty at the start of the modeling time, but the exact amount of water in the stores at the beginning of the modeling time is unknown. To solve this problem, 3 options will be explored and compared. First the start value will be set as the half of the capacity of the store. The stores will be fill half full, regardless to the value of the capacity store. The second method will fill the store random between 0 and 100%. The start value of the model is treated as a model parameter in this method. In the last method a startup time will be used. A fixed period of time will be used to run the model, which means the start values of the parameters will be produced by the model self. The evaluation of the model is not influenced by the startup time because the evaluation starts by the beginning of the calibration period, so after the startup time.

3.3.3 Comparison of models

The calibration of the model is done based on the evaluation method described in chapter 3.3.1. The models are more complex and therefore can't be described in one single number. The sensitivity of the models give an indication how the model reacts when circumstances are different.

Scatterplots of the model parameter against the evaluation method are created to evaluate the sensitivity of the model. The scatterplots give information about the optimum of the parameter value. To compare the results of the scatterplots, cumulative plots are created for the parameter value.

The cumulative plots give the probability of a parameter value in the Monte Carlo analysis. A cumulative plot shows the probability of a value for the model parameter. This probability is based on the output of the best method for the calibration of the initial parameter values. The number of runs is set out in the cumulative plot against the parameter value. A horizontal line in the plot gives a higher probability for the value in calibration. When the parameter doesn't have a preference for a value, the cumulative plot will be a linear line.

Confidence intervals of the models are created. The confidence gives an idea of the uncertainty of the model. If the uncertainty bound is low and the observed runoff is within the uncertainty bound, the model gives output with low uncertainty.

At last validation is done by split sample test (Klemes, 1986). In a split sample test, recorded data is split into groups. One group is used for calibration and the second group is used for validation. The data is from the same catchment but from a different point in time so the accuracy of the model is tested on new data to see if the calibrated parameters are really that good. The evaluation method for calibration is the same as for validation.

3.3.4 Calibration of the rainfall gage

The rainfall runoff models have a lot of errors which effect the results of the model. One of the errors is the input error. The rainfall used to predict the runoff is observed by several observation points in the catchment. For the Florentine Catchment, 3 observation points are used for the 435 km^2 large catchment. The exact area for the observation point to cover is unknown. Assumptions can be made by several methods about the area covered, like the Thiessen polygon method. In the first calibration, the average of the rainfall is set as rainfall input. This result is probably not the best possible result because the average of the rainfall is regardless to location of the observation point. But the result is good enough to have a good impression of the models to calibrate the runoff for the rainfall given.

After calibration with the average of the rainfall, the weights of the rainfall gages will be set by a Monte Carlo analysis. The Monte Carlo analysis will create a set with 10000 random weights of the three rainfall observation points. The total volume of the rainfall in the average is 1290. In the Monte Carlo analysis the total volume can vary and the total area can also vary and don't have to end up to one. This doesn't mean the total area can become larger, but it does mean the total volume of rainfall can differ more than it could be when the average area had to end up to one. It also means an observation point can be dominant over the other observation points. This means an observation point with higher peaks than average can have a larger weight what eventually means the peak flow will be better simulated by the rainfall runoff model. The same can apply for the low flow periods.

After the calibration of the rainfall gage, the rainfall gage and model parameters will be set random at the same time. By calibrating the rainfall gage the water volume is optimized for the given parameter set. The parameter set is best for the average of the rainfall. It is possible that there is a parameter set which is better optimized for a certain rainfall volume or rainfall area.

4. Results

The described research method is to find answers to the research questions in this chapter. The results of the evaluation methods can be read in chapter 4.1, the results of the model calibration can be read in chapter 4.2, the results of the calibration can be read in chapter 4.3 and the results of the calibration of the rainfall gage can be read in chapter 4.4.

4.1 Evaluation method

The scatterplots of the evaluation of the evaluation methods can be seen in appendix 9.3. The Index of agreement evaluates the peak flow badly. The peak flow is not on the y=x line, but is plotted more to the observed runoff. So where there is observed a high runoff, there is simulated a lower runoff, in this case about 40 m^3/s so this method is not chosen for the comparison of the rainfall runoff methods. The best simulation according to these methods have lower peak flows than observed. The coefficient of determination and the NS coefficient evaluate the peak flow better. The NS coefficient is chosen because the NS simulates the volume of the discharge better than the coefficient of determination (0,9% against 1,8% RVE). The Relative volume error seems to have a large error in the scatterplot. The equation used for the calibration and evaluation is:

$$Ev = \mathrm{w}_{NS} \times NS \;+\; w_{RVE} \times \left(1 - \left|\frac{RVE}{100}\right|\right) \;+\; \mathrm{w}_{PF} \times (1 - PF) \;+\; \mathrm{w}_{LF} \times (1 - LF) \qquad \text{(Eq. 4.1-1)}$$

An optimal simulation has an Ev of 1. The weight of every evaluation method has to be determined and set depending on the purpose of the runoff prediction. The largest weight should be given to the NS. The NS evaluates the shape of the simulated runoff compared with the observed runoff. To predict the runoff right, this is the most criterion. The second large weight should be given to the RVE. To be able to use the RVE in the same equation as the NS, the RVE should have a result of 1 for a perfect simulation. This is possible if the absolute value of the RVE is taken and it is divided by 100. The RVE has a result which is between 0 and 1, so it can be compared with the NS. Because 0 is a perfect simulation (no volume difference), the RVE has to be reduced by 1. The perfect simulation has now a result of 1 and the worst simulation has a result of 0. The Peak- and Low flow have a result of 0 for a perfect simulation and > 0 for everything worse than the best simulation. When they also are reduced by 1 the perfect simulation is 1.

For the Florentine Catchment the w_{NS} is 0.75 and the w_{RVE} is 0.25. The Peak flow and Low flow is set 0 because they don’t contribute to a better simulation. In Figure 9.3-1 and Figure 9.3-4 of chapter 9.3, the scatterplots are shown for the low flow and peak flow. Especially the peak flow doesn’t predict the peaks better than the NS or the coefficient of determination does. If the PF does not evaluate the PF will only give a larger error to the evaluation method, because it evaluates the ‘middle’ (between 7 m^3/s and 20 m^3/s) and low flow far worse. If the low flow is included without including the peak flow, the low flow would have more influence in the result so a parameter set better for the low flow will sooner be chosen than a more balanced parameter set for the low and the peak flow.

4.2 Calibration

The theoretical assumptions of the models differ at numerous assumptions as seen in chapter 2. The AWBM, GR4J and SYMHID will be scripted in MatLab. The choice for these models can be explained in its theoretical basis for the stores in the model. In the GR4J model, the runoff based on the store capacities is described by a non-linear equation, while in SYMHID a linear equation is used. The AWBM is a semi-linear semi-nonlinear model. The VIC model is a more data driven model, so this is not easy to calibrate without the right data. To evaluate the calibration results, the evaluation equation will be used. To be sure the NS coefficient is also given for the calibration and evaluation results. The calibration method is done with the Monte Carlo analysis.

The initial values of the stores were treated as parameter in table 4.2-1. The stores were half full for table 4.2-2 and the startup time is used in table 4.2-3. For all models the validation value is lower than by calibration. The difference is the smallest when the starting parameters are used as parameter in the Monte Carlo simulation. The large difference of the calibration and validation result of the SYMHID is notable. This indicates the SYMHID is not able to estimate the parameters for the Florentine Catchment.

TABLE 4.2-1 CALIBRATION WITH STARTING VALUES AS PARAMETER

	EV	NS	RUNS > 0.6	VALIDATION	NS
AWBM	0,844	0,82	332	0,708	0,66
GR4J	0,936	0,926	3849	0,838	0,816
SYMHID	0,809	0,771	129	0,49	0,44

TABLE 4.2-2 CALIBRATION WITH STORES HALF FULL

	EV	RUNS > 0.6	VALIDATION
AWBM	0,841	937	0,635
GR4J	0,918	9505	0,787
SYMHID	0,759	103	0,383

TABLE 4.2-3 CALIBRATION WITH STARTUP TIME

	EV	RUNS > 0.6	VALIDATION
AWBM	0,855	1447	0,686
GR4J	0,919	9362	0,761
SYMHID	0,703	27	0,408

The best parameter set of the GR4J is plotted against the observed runoff in figure 4.2-1.

For the best parameter set of the GR4J, the validation plot is given in figure 4.2-2.

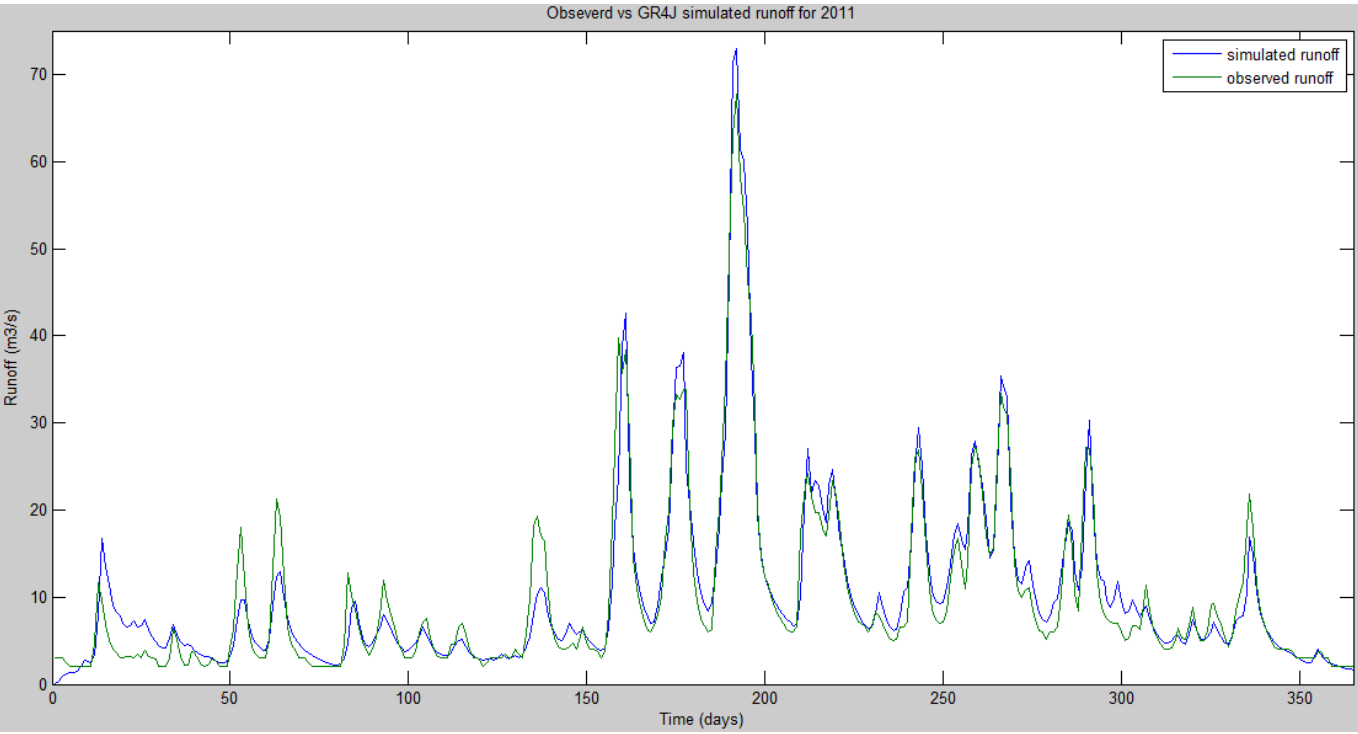


FIGURE 4.2-1 GR4J SIMULATED VS. OBSERVED RUNOFF FOR CALIBRATION (2011)

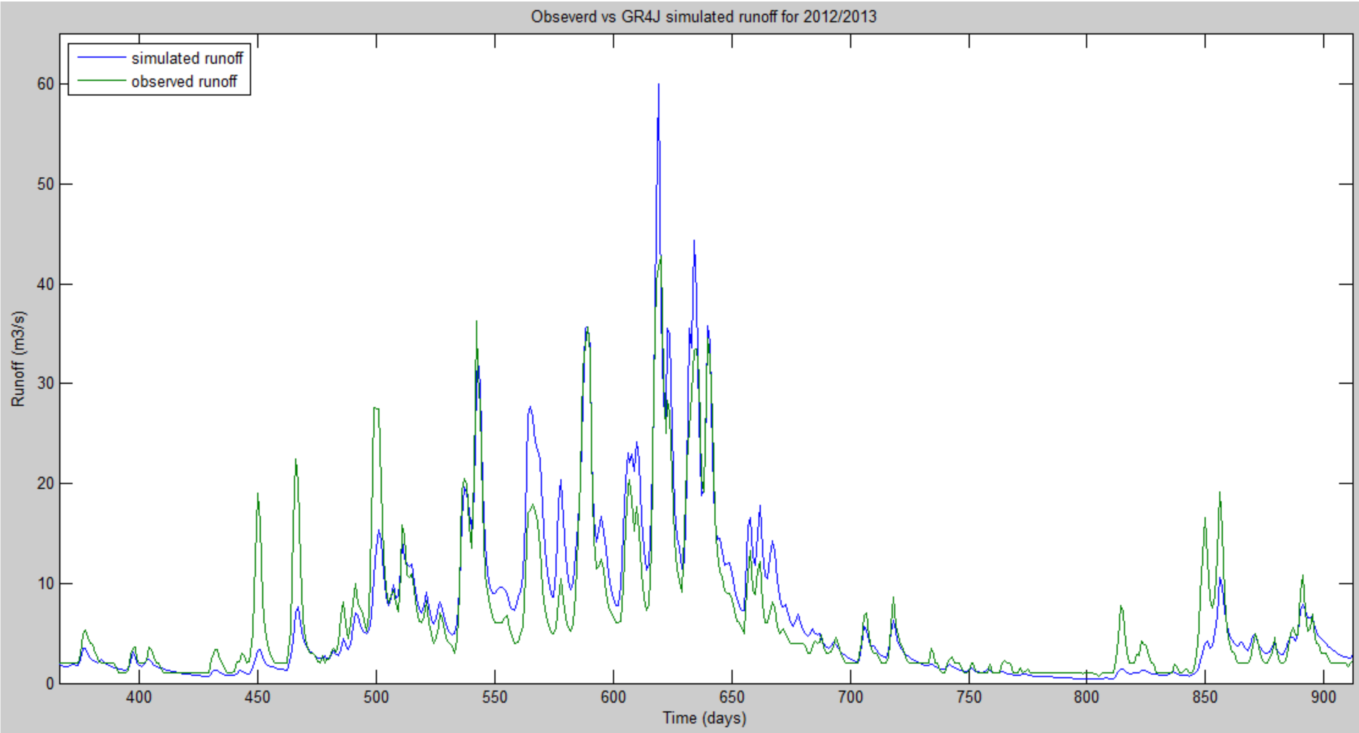


FIGURE 4.2-2 GR4J SIMULATED VS. OBSERVED RUNOFF FOR VALIDATION (1/2012 – 7/2013)

The calibration and validation plots of the AWBM and SYMHID are included in appendix 9.4.

4.3 Comparison of models

The scatterplots give the range of the best simulations plotted against the NS, while the cumulative plots show the sensitivity of the parameter value. The results of the scatterplots and cumulative plots contain only runs with a NS value larger than 0,6. Figure 4.3 -1 till figure 4.3 – 4 show the scatterplots and cumulative plots for two of the four parameters for the GR4J. For the maximum capacity of the production store (x_1) it is likely to have the best result according to the NS between 300 and 550. The probability of the parameter is even for all values, because the cumulative plot is a straight line. For the groundwater exchange (x_2) something can be said about the probability. The value of x_2 will

FIGURE 4.3-2 SCATTERPLOT OF X1 TO NS

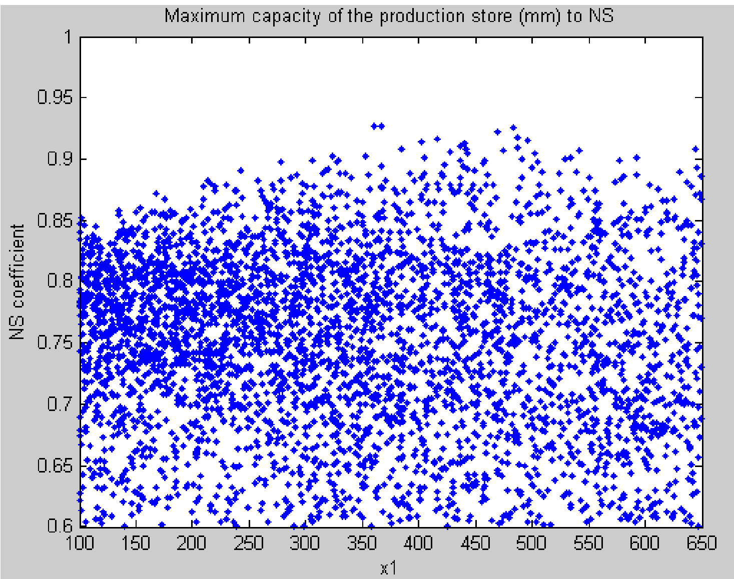
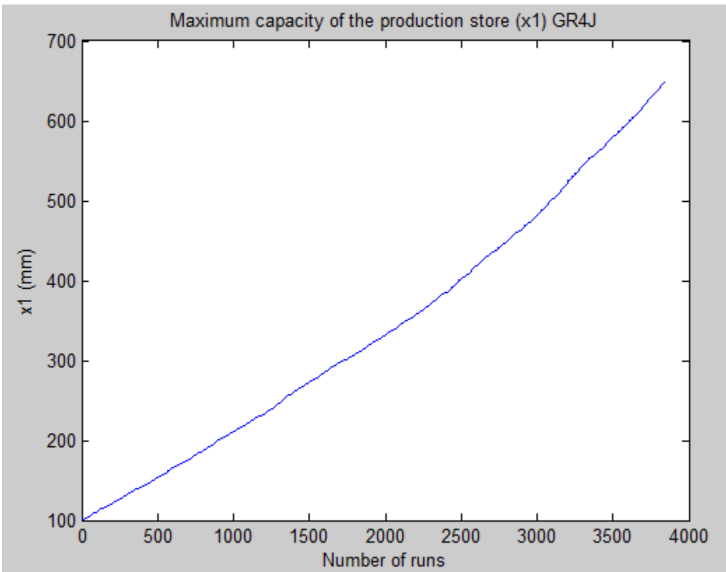


FIGURE 4.3-1 CUMULATIVE PLOT OF X1



probably be larger than -1 according to figure 4.3-3 because the line is steep before -1, so this parameter value is used in a few runs with a NS higher as 0,6. The best simulation will have parameter value larger than -1 and 1 according to figure 4.3-4.

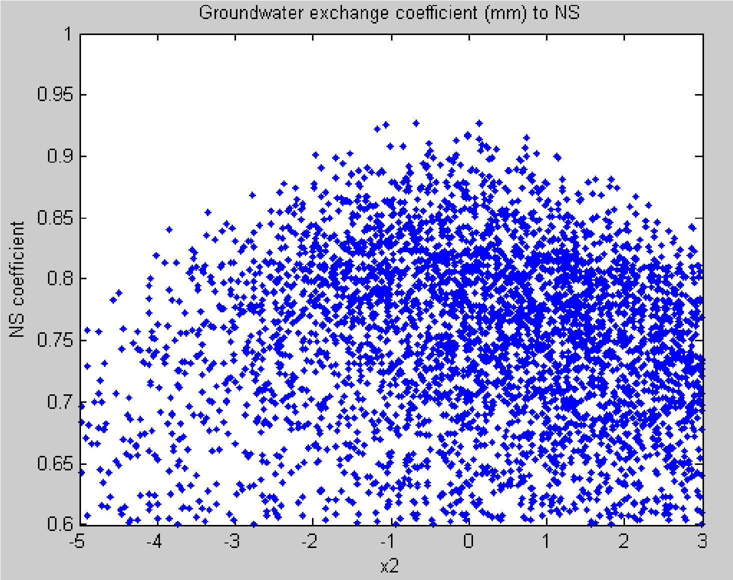


FIGURE 4.3-4 SCATTERPLOT OF X2 TO NS

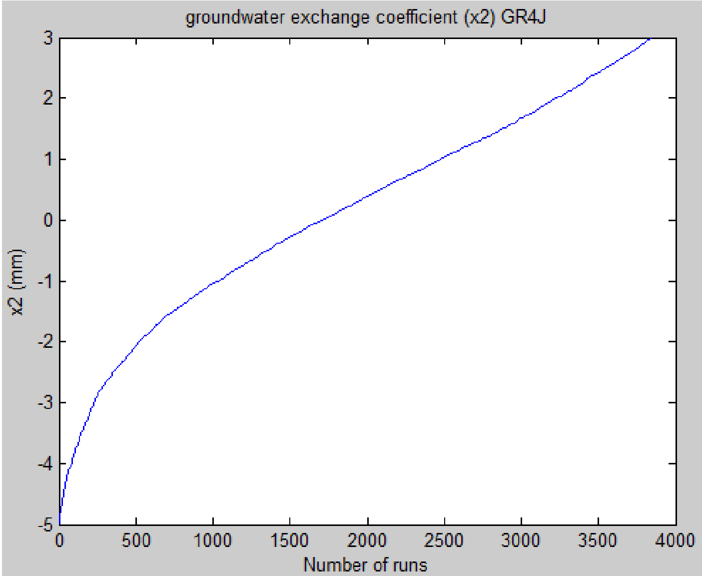


FIGURE 4.3-3 CUMULATIVE PLOT OF X2

If the model is sensitive for a parameter, the parameter will give a high NS value for only a small range of parameter value. When a model is affected this much by a parameter, small changes in circumstances lead to whole different runoff which can affect wrong decisions made by the manager of the dam or lake. All scatterplots and cumulative plots for all three models are included in appendix 9.5. Best examples of sensitive parameters are the SUB parameter in the SYMHID model (Figure 9.5-24 Scatterplot of NS of SUB in appendix 9.5.3) and the base flow, surface flow and average capacity store in the AWBM model (appendix 9.5.1).

Figure 4.3-5 shows the confidence interval of the simulated runoff. The bound of the GR4J model is narrow but the observed runoff is almost everywhere included in the uncertainty bound. The model has low uncertainty because 150 of the best results according to the NS, the results in terms of simulated runoff stay close to each other. If the range would become larger, like the SYMHID model in Figure 9.5-39, the model is sensitive for changes in rainfall or circumstances. Like the scatterplots, if circumstances change or extreme rainfall situations occur, the model can give easily wrong runoff in which the calibrated model becomes less good.

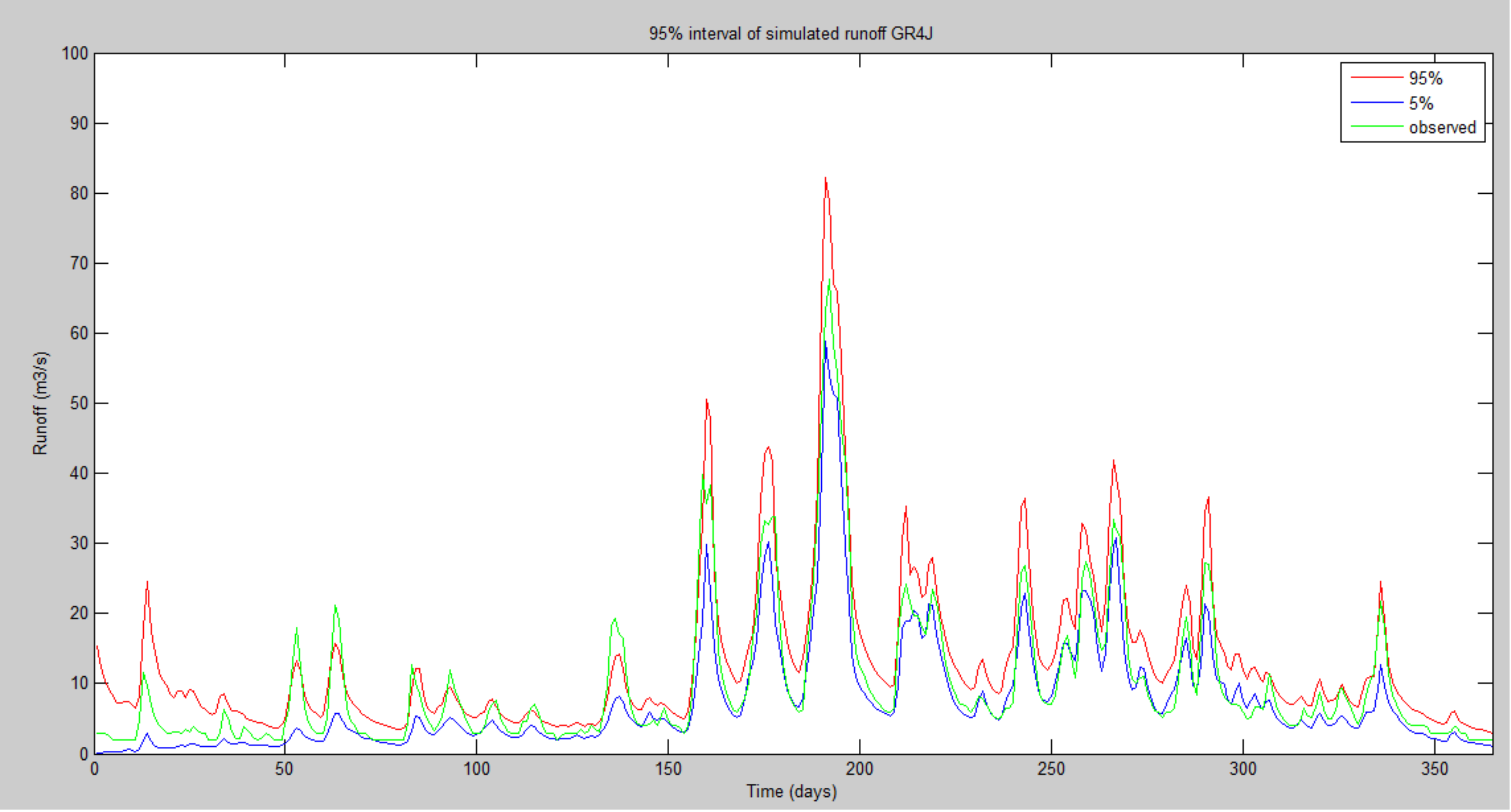


FIGURE 4.3-5 CONFIDENCE INTERVAL OF GR4J

The overall uncertainty bound of the SYMHID seems to be the worse with a large bound (appendix 9.5.4). The uncertainty bound for the peak flow is larger for the AWBM. The overall uncertainty seems to be the largest for the SYMHID but the AWBM is more sensitive for extreme rainfall events.

4.4 Rainfall calibration

In the rainfall calibration, the rainfall areas are calibrated. The rainfall gage is evaluated as something given, but is evaluated as something what can differ even from model to model. The calibration is done with the start values as parameter. This is done because the calibration method gives the best results for SYMHID and GR4J. The AWBM is better with a startup period, but the calibration and evaluation is still good with the start values as parameter. The FV in stands for the Florentine valley, MP stands for Misery Plateau and TS stands for mount Tim Shea in Table 4.4-1.

RUN	*FV*	*MP*	*TS*	*VOL*	*EV*	*NS*
1161	0.55	0.27	0.16	1601	0.846	0,826
1162	0.55	0.22	0.21	1612	0.847	0,825
1163	0.38	0.24	0.39	1681	0.857	0,825
1164	0.34	0.26	0.43	1704	0.862	0,825
1165	0.38	0.21	0.43	1694	0.863	0,827

TABLE 4.4-1 CALIBRATION OF THE RAINFALL GAGE FOR AWBM

RUN	FV	MP	TS	VOL	EV	NS
809	0.36	0.21	0.42	1638	0.935	0,927
810	0.27	0.30	0.42	1630	0.938	0,927
811	0.36	0.21	0.40	1614	0.941	0,927
812	0.31	0.28	0.40	1618	0.942	0,928
813	0.37	0.26	0.35	1608	0.943	0,928

TABLE 4.4-2 CALIBRATION OF THE RAINFALL GAGE FOR GR4J

RUN	FV	MP	TS	VOL	EV	NS
841	0.52	0.40	0.02	1493	0.825	0,779
842	0.51	0.35	0.09	1517	0.830	0,78
843	0.47	0.41	0.08	1525	0.832	0,78
844	0.53	0.33	0.09	1539	0.834	0,78
845	0.53	0.35	0.07	1531	0.834	0,78

TABLE 4.4-3 CALIBRATION OF THE RAINFALL GAGE FOR SYMHID

Instead of calibrating the models in two separate steps, the model can be calibrated in one step. The different areas of the rainfall gages will be set random, as well as the parameter values. The results are given in Table 4.4-4, Table 4.4-5 and Table 4.4-6.

TABLE 4.4-4 CALIBRATION RAINFALL GAGE FOR THE AWBM IN ONE ANALYSIS

RUN	FV	MP	TS	VOL	EV	NS
157	0,75	0,17	0,79	2893	0,842	0,791
158	0,63	0,28	0,11	1658	0,843	0,812
159	0,56	0,45	0,73	2881	0,845	0,807
160	0,38	0,19	0,77	2261	0,852	0,806
161	0,48	0,20	0,42	1851	0,853	0,811

TABLE 4.4-5 CALIBRATION RAINFALL GAGE FOR GR4J IN ONE ANALYSIS

RUN	FV	MP	TS	VOL	EV	NS
630	0,25	0,66	0,09	1639	0,910	0,890
631	0,73	0,14	0,27	1754	0,911	0,908
632	0,41	0,38	0,11	1426	0,918	0,898
633	0,47	0,35	0,66	2429	0,921	0,910
634	0,37	0,28	0,74	2303	0,923	0,880

TABLE 4.4-6 CALIBRATION RAINFALL GAGE FOR SYMHID IN ONE ANALYSIS

RUN	FV	MP	TS	VOL	EV	NS
43	0,48	0,40	0,10	856	0,754	0,70
44	0,58	0,38	0,20	996	0,760	0,71
45	0,65	0,28	0,18	788	0,766	0,70
46	0,09	0,27	0,08	592	0,768	0,73
47	0,37	0,36	0,07	739	0,808	0,76

Table 4.4-7 includes the validation values for the different methods. The best method in calibration described in chapter 4.2 is set in the table in VAL1. The result for the calibration of the rainfall gage is set in VAL 2 in the table and the results for the calibration in one step is set in VAL3.

TABLE 4.4-7 VALIDATION RESULTS FOR THE DIFFERENT MODELS

	VAL1	NS1	VAL 2	NS2	VAL 3	NS3
AWBM	0,708	0,66	0,651	0,61	0,650	0,64
GR4J	0,838	0,816	0,842	0,82	0,438	0,41
SYMHID	0,49	0,44	0,609	0,55	-0,975	-1,24

5. Discussion

Several assumptions were done which doesn't have to be true according to their behavior in nature. The assumption which affects the results the most is the use of a lumped version of the rainfall runoff models. The catchment is seen as one large area with exactly the same characteristics on every small grid, which is not very realistic. The assumption is done to simplify the model and because the characteristics of the different semi-catchments are unknown but also too expensive to investigate for every catchment. For the relatively small Florentine Catchment, the influence of the lumped version is small. For larger catchments like for example the Rhine in Europe, it is advisable to divide the catchment into several semi-catchments and use the semi-catchments for the prediction of runoff.

The evapotranspiration is set as the average for the month of for the region of the Florentine catchment. The average is taken over 30 years of data. The average can affect an extreme dry or wet year but the difference in optimization will probably not be very large. The data is also not specific from the Florentine catchment, but from the middle of Tasmania which contains a larger area than only the Florentine catchment. The differences will not be very large in the area, but still there are probably some differences.

The weight of different elements in the evaluation method is set manual. This leaves space for interpretation. In the proposed evaluation method the NS should always be leading because it evaluates the shape of the simulated runoff best. If there is a small difference in NS coefficient, another evaluation method like the RVE can give solution.

The models have only been calibrated with the Monte Carlo analysis. The Monte Carlo analysis is run 10000 times with random values for the parameters of the model. It is unknown if the number of runs is sufficient to create the right parameter sets. A model with a large number of parameters has a smaller change of finding the optimum parameter set than a model with few parameters. The limits of the parameters also infect the change of finding and selecting the best parameter set in the same procedure. The same can be concluded about the calibration of the parameter set and rainfall gage in one Monte Carlo analysis. By including the rainfall gage in one Monte Carlo analysis, the number of unknown variables is increased and more runs in the Monte Carlo analysis can give more and better results.

The number of results for the SYMHID model for calibration is too small to give conclusions about the sensitivity and the interval bound. Lower the standard for the analysis would only include more pore results, so the quality of the results is lower. More runs with probably more results would give more certainty to set conclusions about the uncertainty and performance of the model.

The improvement of the validation with a variable rainfall gage is minor. The Florentine catchment is a small catchment of 435 km^2. In larger catchments and with more rainfall observation points, the improvement possibly can be larger.

6. Conclusion

Entura selects a model based on experience to predict the runoff for a catchment. To create a simple and stepwise selection method, the Florentine Catchment is used as example to select a model out of the models used by Entura: GR4J, AWBM, SYMHID and the VIC model.

The AWBM, GR4J and SYMHID are conceptual models while the VIC model is a physically based model. The concept of the VIC model is too different and need too much data to be able to compare the VIC model with the other models. So the AWBM, GR4J and SYMHID are evaluated. The difference in the principle of the model is the mathematical relation between the water storage and the runoff. The SYMHID model describes the relation between the storage and the runoff linear, the AWBM describes the runoff semi-linear while GR4J describes the runoff non-linear. The GR4J is the only model to use routing and with it a unit hydrograph. By the other models routing can be included but it doesn't have to be included for the model to give runoff.

To be able to evaluate the rainfall runoff models after calibration, six evaluation methods have been evaluated. The Nash Sutcliffe coefficient (NS), the Index of Agreement and the Coefficient of determination evaluate the shape of the simulated runoff, compared with the observed runoff. The NS coefficient evaluates the shape best compared with the observed runoff. The Nash Sutcliffe coefficient is the basic of the proposed evaluation method. The Nash Sutcliffe method evaluates the shape best for both low flow and peak flow. The three other evaluation methods evaluate different elements of the simulated runoff. The RVE evaluates the difference between the simulated and observed runoff and is the second important element of the evaluation method which should be included. For the Peak flow and the Low flow it depends if they contribute to the evaluation. For the Florentine catchment the peak flow doesn't evaluate the peaks better than the Nash Sutcliffe method and is therefore not included. To not over evaluate the base flow, the base flow evaluation is also excluded from the evaluation. The evaluation used is:

$$Ev = 0{,}75 \times NS \; + \; 0{,}25 \times \left(1 - \left|\frac{RVE}{100}\right|\right)$$

The GR4J model is best calibrated with an Ev of 0,936 and best validated with a Ev of 0,84. This result is very good. The AWBM is second best with a Ev in after calibration of 0,844 and after validation of 0,708. The difference between the calibration and validation result can be the almost linearity of the model with three capacities in the soil moisture. The SYMHID is still good in calibration with a NS of 0,809 but becomes poor in validation with a NS of 0,49. This poor result for validation for SYMHID means the SYMHID is not capable in estimating the conditions of the Florentine catchment. The linearity of the model is probably too rough for the small catchment to estimate the runoff properly.

The calibration is done with treating the initial values with different methods. The initial value is treated as half full as parameter and is decreased with a startup time. The best results were achieved by treating the initial values as a parameter. This is probably the best if there is not enough good data to do a startup time. Startup time is probably still the best method because the problem of over- or under estimation of the initial values is reduced to almost zero. With treating the initial values as a parameter, the initial value can still be overestimated which can give a better RVE but is not realistic.

The GR4J is the less sensitive model for changes in circumstances or extreme events and is therefore the most reliable according to the sensitivity plot. The number of good runs above a Eval of 0,57 value can also give an indication of the reliability of the model. If there are a lot of runs with a good evaluation value, the model is less sensitive and will give good results in calibration and validation. The SYMHID seems to be the worst for overall sensitivity and has a sensitive parameter for circumstances in characteristics in the SUB (interflow to lower stores) and SQ (infiltration loss) parameters. These components give an indication for the volume of water which will infiltrate to lower stores, so it is clear why these parameters are important for the model and have an optimum. The AWBM seems to be the most sensitive for peak flow. This can be the effect for sensitivity of the average capacity C_{ave} of the AWBM. A small value for the capacity results in a high Ev. Because the capacity is small, less water can be intercepted and stored, so the runoff has to be large to lose the water.

The number of sensitive parameters are an indication for the difficulty of the calibration and evaluation of the model. A small change moving from the optimum to the outside gives a large change in output so the model is sensitive. It is an possible explanation for the number of good (Ev > 0,6) calibration runs of the models and can be concluded from the confidence interval. So the SYMHID has the highest sensitivity and the GR4J has the lowest sensitivity.

To decrease the input error of the model and improve the performance of it, the rainfall gage is also calibrated for each model. The volume of rainfall increases from 1290 m^3 to 1500 – 1600 m^3 depending on the model. The Ev also increases for the all the models, but not much. The largest increase in calibration is 0,008 for the SYMHID. The validation also increases for SYMHID and GR4J. The SYMHID validation increases with 0,11 but is with a Ev of 0,609 still a poor result for validation. The NS of the GR4J improves with 0,004 and the AWBM is even decreasing with 0,5. The calibration in rainfall gate doesn't make a poor result good and it doesn't make a good result significant better.

For small catchments with a small number of rainfall gages, it is not be useful to calibrate both the model and the rainfall gage in one or two steps. If the total area of the catchment is larger than the Florentine Catchment and if the measurement points are less equal spread over the area, it maybe is useful to calibrate the rainfall gage.

Overall the GR4J model gives the best output for the Florentine catchment. The Ev values for the calibration and validation period is the best and the model is less sensitive because there is not a single parameter which is most important for the output of the model this can also be seen in the 95% confidence interval. The range of variation in the output of the calibration is small unlike the SYMHID model which has a large variation of the confidence interval and bad results for validation.

7. Recommendations

To select the right model for describing the runoff for a random catchment, a simple stepwise model is tested in this document. It is advised to use the evaluation methods of the Nash - Sutcliffe method and the Relative volume error and combine them with the equation 4.1-1:

$$Ev = 0{,}75 \times NS \; + \; 0{,}25 \times \left(1 - \left|\frac{RVE}{100}\right|\right)$$

The easiest way to calibrate the models for the catchment is to use a Monte Carlo analysis and calibrate the models between the given limits. Because the limits of the VIC model are unknown, this model can only be used if the data of the vegetation is known. The number of runs of the Monte Carlo is not fixed and depends on the number of parameters. About 10000 runs should be sufficient for the AWBM, GR4J and SYMHID even if the initial values are included in the Monte Carlo analysis.

If there is sufficient data, it is advised to use a startup time. If there is not enough data, it is advised to treat the initial values as a parameter in the Monte Carlo analysis. For reliable results, there should be enough data for at least one year of calibration and one year of validation.

If the catchment is large, the catchment should be divided in semi catchments to get reliable results for the catchments. The size of the catchments is not strict, but at a certain point the lumped model is not sufficient enough to predict runoff.

Concluding stepwise:

1. Check if the lumped model is sufficient for the catchment.
2. Create a Monte Carlo analysis for GR4J, AWBM and SYMHID with 10000 runs
    a. Use a startup time if possible
    b. Otherwise treat the initial conditions as parameter
3. Evaluate the parameter sets with equation 4.1-1.
4. Validate the models with a Split sample test.
5. Compare the results of the models and select the best model

The calibration of the rainfall gage turned out not to be useful for the Florentine Catchment. It should be investigated if it is useful for large catchments with measurement points not divided equal to the catchment area.

If there is enough physically based data available, it is recommended to evaluate the VIC model with the conceptual based models. There was not enough data available for the Florentine Catchment to include the VIC model also in the research.

8. Bibliography

Beven, K. J. (2012). Rainfall-Runoff Moddeling, the Primer. Lancaster: Wiley - Blackwell.

Beven, K., & Kirkby, M. (1979). A physically based, variable contributing area model of basin hydrology. *Hydrological Sciences Journal*, 43-69.

Biondi, D., Freni, B., Iacobellis, V., Mascaro, G., & Montanari, A. (2012). Validation of hydrological models: conceptual basis, methodological approaches and a proposal for a code of practice. *Physics and Chemistry of the Earth*, 70 - 76.

Boughton, W. C. (1993). A hydrograph-based model for estimating the water yield of ungauged catchment. *National Conferece publication no. 93/14*, 317 - 324.

Broughton. (2004). The Australian water balance model. *Environmental Modelling & Software* , 943 - 956.

Chapman, T. (1999). A comparison of algorithms for stream flow recession and baseflow separation. *Hydrological Processes*, 701-704.

Chiew, F. H., Peel, M. C., & Westernd, A. (2002). Application and testing of the simple rainfall-runoff model SYMHID. *Water Recources Publications*, 335 - 367.

Crean, D., & Davy, S. (2013, October). *Annual Report.* Retrieved 6 19, 2014, from HydroTasmania: http://www.hydro.com.au/annual-reports/2013/contents/Infrastructure/storage/page01.html

Edijatno, Oliveira Nascimento de, N. D., Yang, X., Makhlouf, Z., & Michel, C. (1999). GR3J: A daily watershed model with three free parametres. *Hydrological Sciences Journal*, 263 - 277.

Gao, H., Tang, Q., Shi, X., Zhu, C., Bohn, T., Su., F., . . . Wood, E. F. (2009). *Water Budget Record from Variable Infiltraion Capacity (VIC) Model, Algorithm Theoretical Basis.* Seattle: University of Washington.

Greenpeace. (2004, January). Retrieved 6 16, 2014, from Greenpeace.org: http://www.greenpeace.org/belgium/Global/belgium/report/2004/1/de-oerbossen-van-tasmani.pdf

HydroTasmania. (n.d.). *About us*. Retrieved 6 16, 2014, from HydroTasmania: http://www.hydro.com.au/about-us

HydroTasmania. (n.d.). *Meadowbank Power Station.* Retrieved 6 16, 2014, from HydroTasmania: file:///C:/Users/User/Downloads/Meadowbank_Power_Station-Fact-Sheets%20.pdf

Jothiprakash, V., & Magar, R. (2009). Soft computing tools in rainfall runoff modeling. *ISH Journal of Hyraulic Engineering*, 84 - 96.

Klemes, V. (1986). Operational testing of hydrological simulation models. *Hydrological sciences*, 13-24.

Madsen, H. (2000). Automatic calibration of a conceptual rainfall-runoff model using multiple objectives. *Journal of Hydrology*, 276 - 288.

Meteorology, B. o. (2005). *Average Evapotranspiration*. Retrieved from Bureau of Meteorology - Australian Government: http://www.bom.gov.au/jsp/ncc/climate_averages/evapotranspiration/index.jsp

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

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

Regulator, A. E. (n.d.). *National Electricity Market consumption*. Retrieved 6 19, 2014, from Australian Energy Regulator: http://www.aer.gov.au/node/9765

Shaw, E., Beven, K., Chappel, N., & Lamb, R. (2011). *Hydrology in Practice.* Oxon: Spoon Press.

Tasmania, E. A. (n.d.). *Department of Treasury and Finance.* Retrieved 6 16, 2014, from Tasmanian Government: http://www.treasury.tas.gov.au/domino/dtf/dtf.nsf/LookupFiles/Population.pdf/$file/Population.pdf

Wood, E. F., Lettenmaier, D. P., & Zartarian, V. G. (1992). A Land-Surface hydrology parameterization with subgrid variability for general circulation models. *Journal of Geophysical Research*, 2717 - 2728.

Xu, L., Lettenmaier, D. P., Wood, E. F., & Burges, S. J. (1994). A simple hydrologically based model of land surface water and energy fluxes for general circulation models. *Journal of Geophysical Research*, 415 - 428.

9. Appendix

9.1 Map of Florentine Catchment

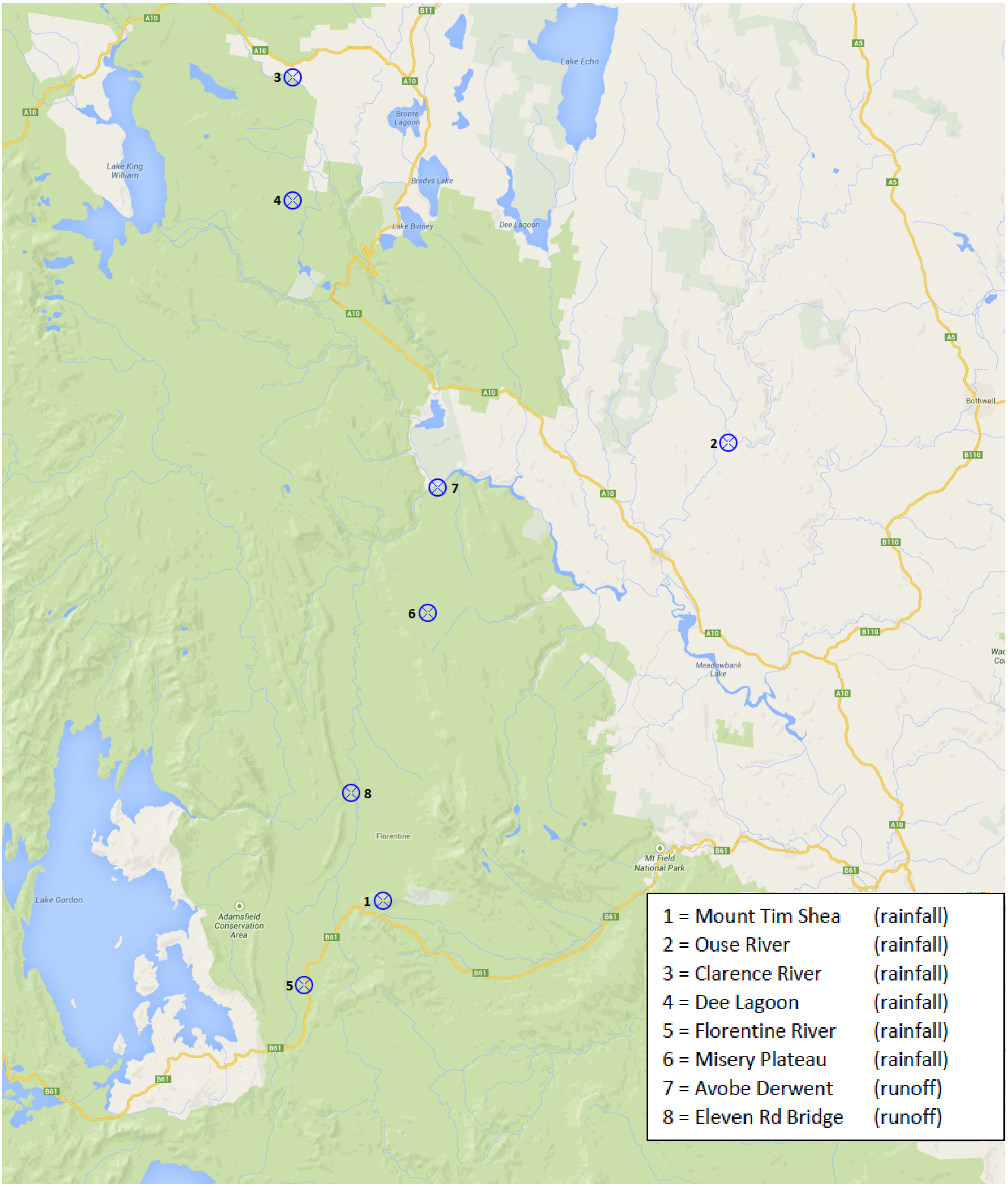


FIGURE 4.4-1 MAP OF FLORENTINE CATCHMENT

9.2.2 GR4J

The mathematical equations for the GR4J are:

If P>E, then $P_n = P - E$ and $E_n = 0$ (Eq. 9.2-16)

Otherwise $P_n = 0$ and $E_n = P\text{-}E$ (Eq. 9.2-17)

In case P_n >E than the rainfall fills the production store P_s. It is determined by a function of the level S in the store by:

$$P_s = \frac{x_1\left(1-\left(\frac{S}{x_1}\right)^2\right)\tanh\frac{P_0}{x_1}}{1+\frac{S}{x_1}\tanh\frac{P_0}{x_1}} \qquad \text{(Eq. 9.2-18)}$$

Where x1 is the maximum capacity of the storage. When E_n >0 the equation is:

$$E_s = \frac{S\left(2-\left(\frac{S}{x_1}\right)^2\right)\tanh\frac{E_0}{x_1}}{1+\left(2-\left(\frac{S}{x_1}\right)\right)\tanh\frac{E_0}{x_1}} \qquad \text{(Eq. 9.2-19)}$$

The new level of storage S will be:

$$S = S - E_s + P_s \qquad \text{(Eq. 9.2-20)}$$

The percolation Perc from the production store is calculated as a power function of the reservoir content:

$$Perc = S\left\{1 - \left[1 + \left(\frac{4}{9}\frac{S}{x_1}\right)^4\right]^{-1/4}\right\} \qquad \text{(Eq. 9.2-21)}$$

$$S = S - Perc \qquad \text{(Eq. 9.2-22)}$$

The percolation occurs if the maximum capacity of the sorage is filled by 4/9. Given the formula given for the percolation, the percolation does not contribute much to the streamflow and is interesting mainly for low flow simulation.

The GR4J is a model which includes routing. The basics of routing is explained in the first chapter. The total amount of water that reaches the routing is:

$$P_r = Perc + (P_n - P_S) \qquad \text{(Eq. 9.2-23)}$$

P_r is divided into two flow components with fixed partial components. 90 % of of the P_r is routed by a unit hydrograph UH1 and then a non-linear routing store. The remaining 10 % are routed by a single unit hyrograph UH2. With UH1 and UH2 the model is able to modelate the time between the rainfall and the runoff peak. UH1 has a time base of x_4 while UH2 has a timebase of $2x_4$. The timebase x_4 is in days and the value is always greater than 0,5 days.

The unit hydrograph is derived from the corresponding S curve, SH1 for UH1 and SH2 for UH2. The is curve is a determined by parameter x4 over time t:

$$For\ t \leq 0, SH1(t) = 0 \qquad \text{(Eq. 9.2-24)}$$

$$For\ 0 < t < x_4, SH1(t) = \left(\frac{t}{x_4}\right)^{5/2} \qquad \text{(Eq. 9.2-25)}$$

$$For\ t \geq x_4, SH1(t) = 1 \qquad \text{(Eq. 9.2-26)}$$

$$For\ t \leq 0, SH2(t) = 0 \qquad \text{(Eq. 9.2-27)}$$

$$For\ 0 < t \leq x_4, SH2(t) = \frac{1}{2}\left(\frac{t}{x_4}\right)^{5/2} \qquad \text{(Eq. 9.2-28)}$$

$$For\ x_4 < t < 2x_4, SH2(t) = 1 - \frac{1}{2}\left(\frac{t}{x_4}\right)^{5/2} \qquad \text{(Eq. 9.2-29)}$$

$$For\ t \geq 2x_4, SH2(t) = 1 \qquad \text{(Eq. 9.2-30)}$$

The value of UH1 and UH2 can then be calculated by:

$$UH1\ (j) = SH1\ (j) - SH1(j-1) \qquad \text{(Eq. 9.2-31)}$$

$$UH2(j) = SH2(j) - SH2(j-1) \qquad \text{(Eq. 9.2-32)}$$

Where j the integer of the timestep (days, hours) of the model for calculating the runoff.

In the model is also groundwater exchange included. The groundwater exchange acts on both flow components and is calculated as:

$$F = \ x_2\left(\frac{R}{x_3}\right)^{7/2} \qquad \text{(Eq. 9.2-33)}$$

Where R is the level in the routing store, x_3 the maximum level possible in the routing store and x_2 the water exchange capacity coefficient. The coefficient can both be positive and negative, depending if the groundwater is extracting water or giving water. The level in the routing store is updated by adding the output Q9 (UH1) and F as follows:

$$R = max(0; R + Q9 + F) \qquad \text{(Eq. 9.2-34)}$$

The outflow Q_r is calculated as follow:

$$Q_r = R\left\{1 - \left[1 + \left(\frac{R}{x_3}\right)^4\right]^{-1/4}\right\} \qquad \text{(Eq. 9.2-35)}$$

Wherein Q_r is always lower as R. The level in the reservoir becomes:

$$R = R - Q_r \qquad \text{(Eq. 9.2-36)}$$

The discharge from the other 10% (UH2) can be calculated in the same way. Also F has to be calculated, the discharge will be:

$$Q_d = \max(0; Q_1 + F) \qquad \text{(Eq. 9.2-37)}$$

The total discharge is finally obtained by:

$$Q = Q_r + Q_d \qquad \text{(Eq. 9.2-38)}$$

9.2.3 SYMHID

Flow to F is:

$$INR = Rain - INT \qquad \text{(Eq. 9.2-39)}$$

Whereby

$$INT = lessor\ of\ \{IMAX, Rain\} \qquad \text{(Eq. 9.2-40)}$$

IMAX is the total storage of the inception store described by

$$IMAX = lesser\ of\ \{INSC, PET\} \qquad \text{(Eq. 9.2-41)}$$

Where INSC is the interception store capacity.

The runoff of the INR will be

$$IRUN = INR - RMO \qquad \text{(Eq. 9.2-42)}$$

$$RMO = lesser\ of\ \{COEFFe^{-SQ\times\frac{SMS}{SMSC}}, INR\} \qquad \text{(Eq. 9.2-43)}$$

COEFF is the maximum infiltration loss, SQ is the infiltration loss exponent, SMS is the soil moisture store and SMSC is the soil moisture store capacity.

The discharge of S who will go to runoff will be:

$$SRUN = SUB \times \frac{SMS}{SMSC} \times RMO \qquad \text{(Eq. 9.2-44)}$$

Where SUB is the constant of proportionality in interflow equation.

The discharge directly to the groundwater store will be:

$$REC = CRAK \times \frac{SMS}{SMSC} \times (RMO - SRUN) \qquad \text{(Eq. 9.2-45)}$$

Where CRAK is the constant of proportionality in groundwater recharge equation.

Het flow to the soil moisture store is described by:

$$SMF = RMO - SRUN - REC \qquad \text{(Eq. 9.2-46)}$$

The discharge of the soil moisture store to the groundwater will be:

$$ET = lesser\ of \left\{10 \times \frac{SMS}{SMSC}, POT\right\} \qquad \text{(Eq. 9.2-47)}$$

$$POT = PET - INT \qquad \text{(Eq. 9.2-48)}$$

The discharge from the groundwater store to the runoff will be:

$$BAS = K \times GW \qquad \text{(Eq. 9.2-49)}$$

Where K is the base flow linear recession parameter. And GW is the groundwater store capacity.

9.2.4 VIC Model

The VIC model can be modelled with the following equations.

The formula of the total evapotranspiration is:

$$E = \sum_{n=1}^{N} C_n \times \left(E_{c,n}+E_{t,n}\right) + C_{N+1} \times E_1 \qquad \text{(Eq. 9.2-50)}$$

Whereby C_n is the vegetation fractional coverage for the n^{th} land cover, C_{N+1} is the bare soil fraction and C_n =1. Ec,n and Et,n are the evapotranspiration of the canopy layer and of the vegetation. E1 is the bare soil evapotranspiration.

$$E_c = \left(\frac{W_i}{W_{im}}\right)^{2/3} E_p \frac{r_w}{r_w+r_0} \qquad \text{(Eq. 9.2-51)}$$

Whereby W_{im} is the maximum infiltration capacity of the canopy intercept, which has a value of 0.2 times LAI. r_0 is the architectural resistance caused by the difference in humidity gradient between the canopy and the air and is dependant of the vegetation layer of the grid. r_0. r_w represents the transfer of heat and water of the vapour from the evaporating grid into the air above. E_p is the potential evapotranspiration calculated by:

$$\lambda_v E_p = \frac{\Delta(R_n-G)+\rho_a c_p(e_s-e_a)/r_a}{\Delta+\gamma} \qquad \text{(Eq. 9.2-52)}$$

Where λ_v is the latent heat of vaporation, R_n is the net radiation, G is the soil heat flux $(e_s\text{-}e_a)$ represents the vapour pressure deficit of the air, ρ_a is the density of air at constant pressure, c_p is the specific heat of the air, Δ represents the slope of the saturation vapour pressure temperature relationship and γ is the psychrometric constant (66). The latent heat vaporation represents the energy flux which in included in the process of evapotranspiration.

The aerodynamic resistance is described:

$$r_w = \frac{1}{C_w u_z} \qquad \text{(Eq. 9.2-53)}$$

Where u_z is the wind speed at height z and C_w is the transfer coefficient for water.

When the rainfall is less than the canopy evapotranspiration the canopy evapotranspiration is:

$$E_c = f \cdot E_c^* \qquad \text{(Eq. 9.2-54)}$$

Where f is the fraction for canopy evapotranspiration given by:

$$f = min\left(1, \frac{W_i+P\cdot\Delta t}{E_c^*\cdot\Delta t}\right) \qquad \text{(Eq. 9.2-55)}$$

The vegetation evapotranspiration is given by:

$$E_t = \left(1-\left(\frac{W_i}{W_{im}}\right)^{2/3}\right) E_p \frac{r_w}{r_w+r_0+r_c} \qquad \text{(Eq. 9.2-56)}$$

Where r_c is the canopy resistance given by:

$$r_c = \frac{r_{0c} g_T g_{vpd} g_{PAR} g_{sm}}{LAI} \qquad \text{(Eq. 9.2-57)}$$

Where r_{0c} is the minimum canopy resitance, g_t, g_{PAR} and g_{sm} are the temperature factor, vapour pressure deficit factor, photo synthetically active radiation flux (PAR) factor and soil moisture factor. When the rainfall is less than the canopy evaporation the vegetation evapotranspiration is:

$$E_t = (1-f)E_p \frac{r_w}{r_w+r_0+r_c} + f \cdot \left(1 - \left(\frac{W_i}{W_{im}}\right)^{2/3}\right) E_p \frac{r_w}{r_w+r_0+r_c} \qquad \text{(Eq. 9.2-58)}$$

The bare soil capacity is partly dependant of the infiltration on the grid where the rain falls. The infiltration capacity is variable so this has to be determined first by:

$$i = i_m \left[1 - (1-A)^{1/B}\right] \qquad \text{(Eq. 9.2-59)}$$

Where A is the fraction of a grid cell (or catchment) for which the infiltration capacity is less than i (thus $0 \leq A \leq 1$), i_m is the infiltration capacity of the fraction A_s and B is the shape parameter.

The bare soil evapotranspiration is then given by:

$$E_1 = E_p \left(\int_0^{A_s} dA + \int_{A_s}^{1} \frac{i_0}{i_m\left(1-(1-A)^{1/b_i}\right)} \, dA \right) \qquad \text{(Eq. 9.2-60)}$$

The amount of water what infiltrates in the soil can be determined by:

$$\int_{i_0}^{i_0+P} A(i)di \qquad \text{(Eq. 9.2-61)}$$

Where i_0 is the infiltration capacity of the saturated fraction A. P + i_0 is the total amount of net rain falling on the grid.

The water which can't infiltrate is given by:

$$P - \int_{i_0}^{i_0+P} A(i)di \qquad \text{(Eq. 9.2-62)}$$

The total soil moisture storage (W_c) over the area is given by:

$$W_c = \frac{i_m}{(1+B)} \qquad \text{(Eq. 9.2-63)}$$

The generated runoff (Q_d) from the area with the 'free flowing' water P is given by:

$$Q_d = P - W_c + W_0 \qquad i_0 + P \geq i_m \qquad \text{(Eq. 9.2-64)}$$

$$Q_d = P - W_c + W_0 + W_c\left[1 - \frac{(i_0+P)}{i_m}\right] 1 + B, \qquad i_0 + P \leq i_m \qquad \text{(Eq. 9.2-65)}$$

The infiltration runoff equations, the ratio W_o/W_c, represents the initial conditions on the timestep:

$$\beta = \frac{e}{e_p} = 1 - \left[1 - \frac{W_0}{W_c}\right]^{1/B_e} \qquad \text{(Eq. 9.2-66)}$$

Whereby B_e is the evaporation rate and is empirical determined by ≅ 0.6. The parameters e and ep

The runoff from the base capacity is given by:

$$Q_b = k_b W_0 \qquad 0 \leq k_b \leq 1 \qquad \text{(Eq. 9.2-67)}$$

Where Q_b is the baseflow, kb is the baseflowcoefficient which value is between 0 and 1.

$$W_0{}^{+} = W_0{}^{-} - Q_b - e \qquad \text{(Eq. 9.2-68)}$$

$$Q = Q_b + Q_d \qquad \text{(Eq. 9.2-69)}$$

9.3 Evaluation method scatterplots

This appendix contains the scatterplots of the different evaluation methods.

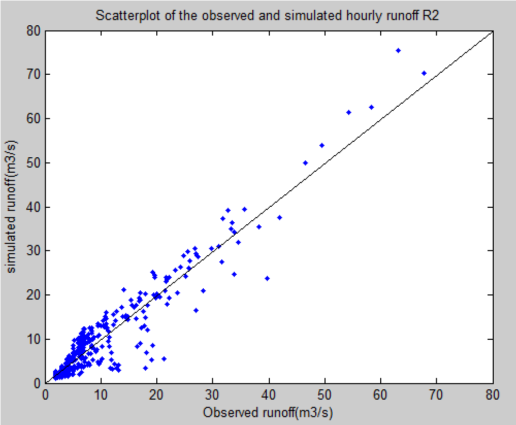


FIGURE 9.3-6 SCATTERPLOT OF THE CORRELATION

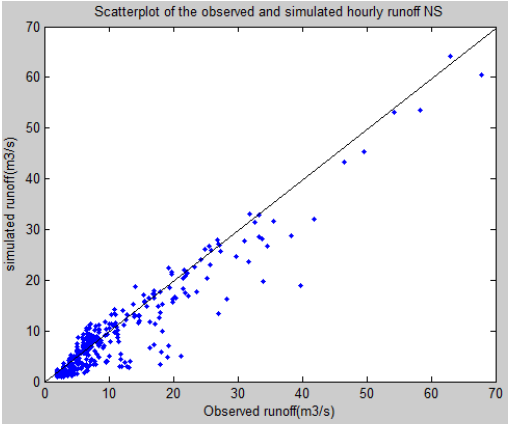


FIGURE 9.3-5 SCATTERPLOT OF THE NASH SUTCLIFFE

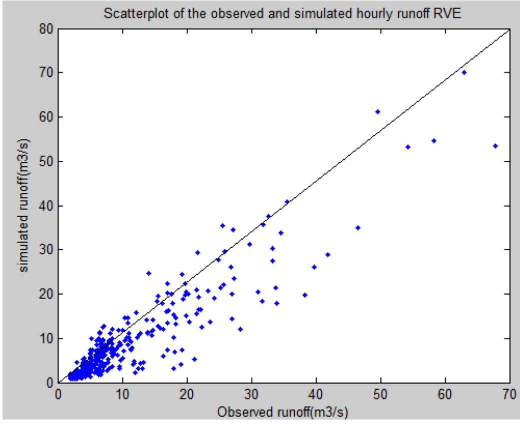


FIGURE 9.3-2SCATTERPLOT OF THE RVE

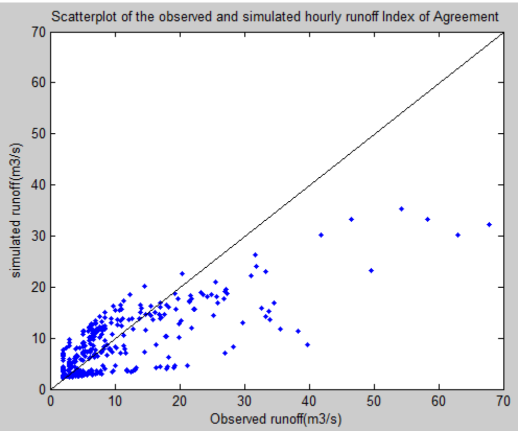


FIGURE 9.3-3 SCATTERPLOT OF THE INDEX OF AGREEMENT

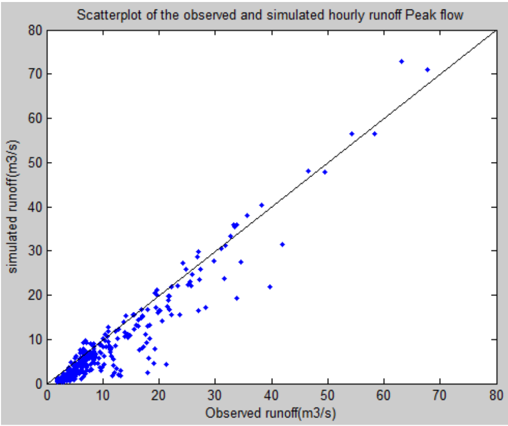


FIGURE 9.3-1 SCATTERPLOT OF THE PEAK FLOW

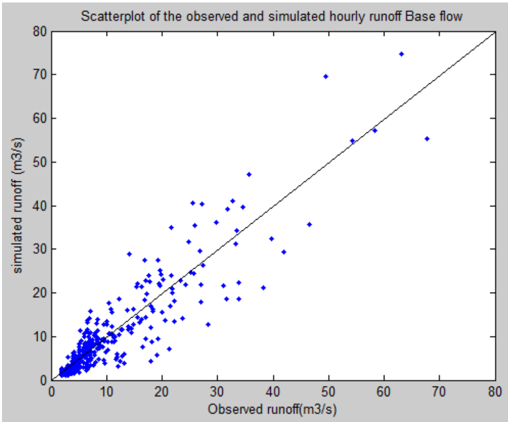


FIGURE 9.3-4 SCATTERPLOT OF THE BASEFLOW

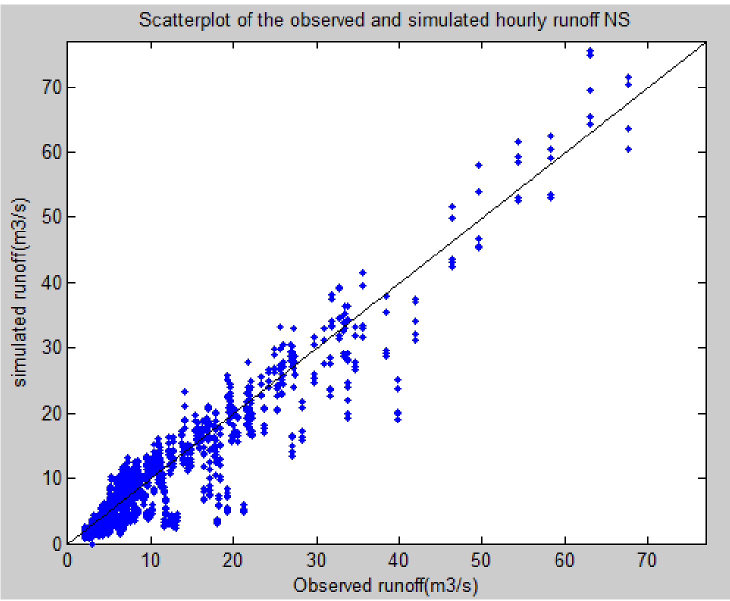


FIGURE 9.3-8 SCATTERPLOT OF THE NS

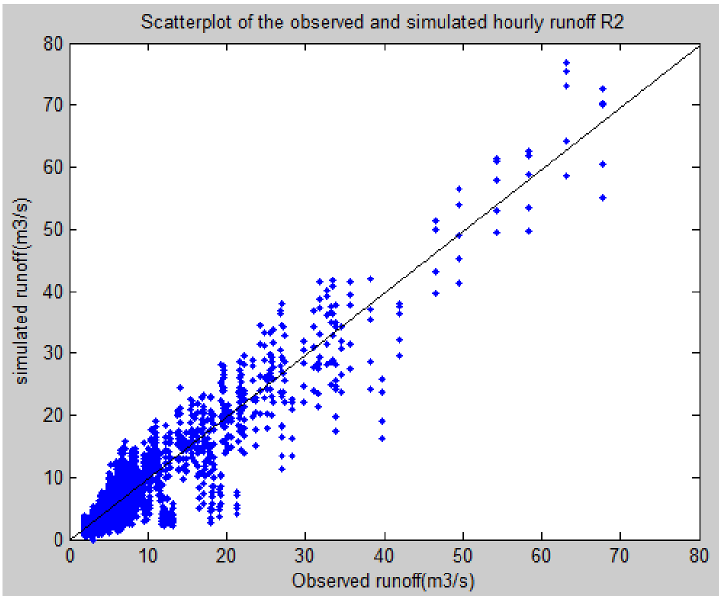


FIGURE 9.3-7 SCATTERPLOT OF THE CORRELATION

9.4 Calibration and validation plots

This appendix contains the validation plots for the AWBM, GR4J and the SYMHID for the period 2011 and 1/2012 – 7/2013.

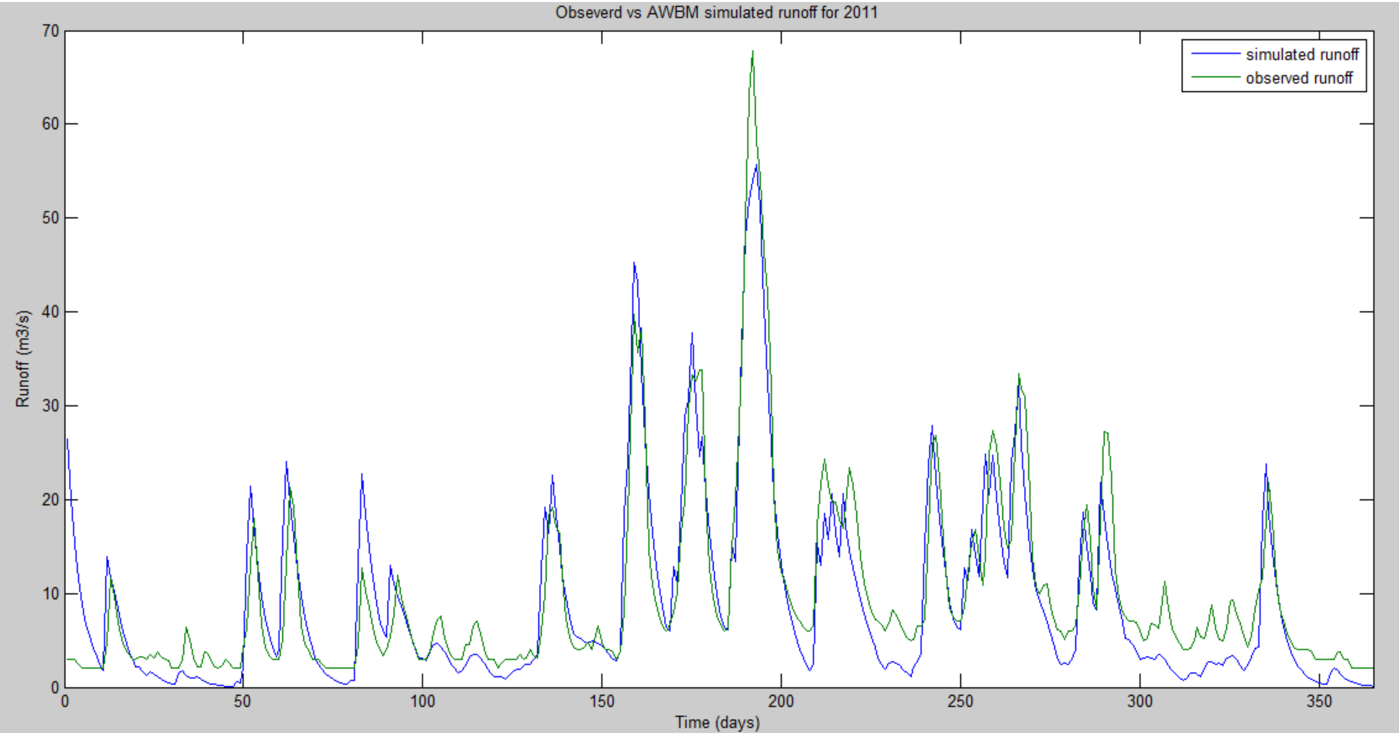


FIGURE 9.4-1 CALIBRATION PLOT OF AWBM FOR 2011

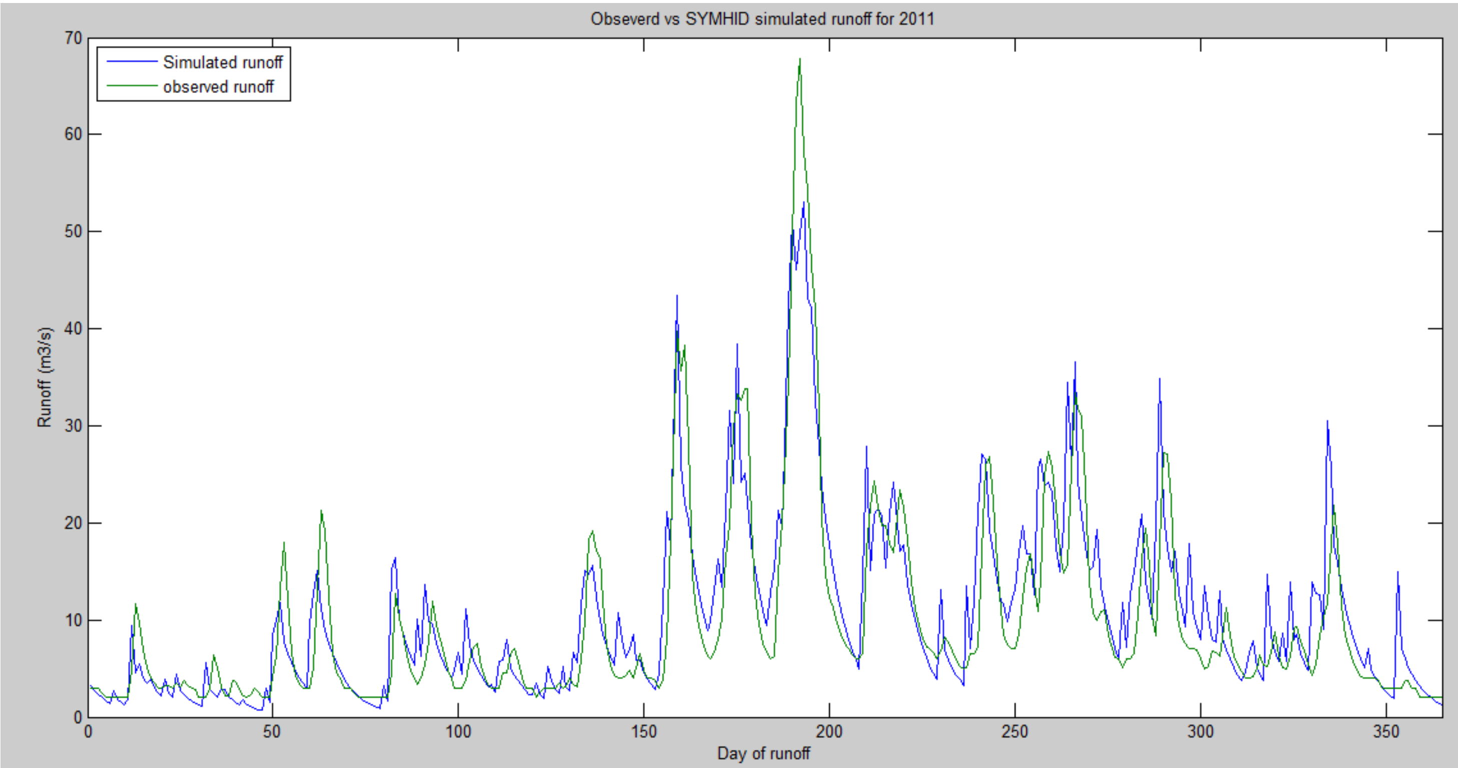


FIGURE 9.4-2 CALIBRATION PLOT OF SYMHID FOR 2011

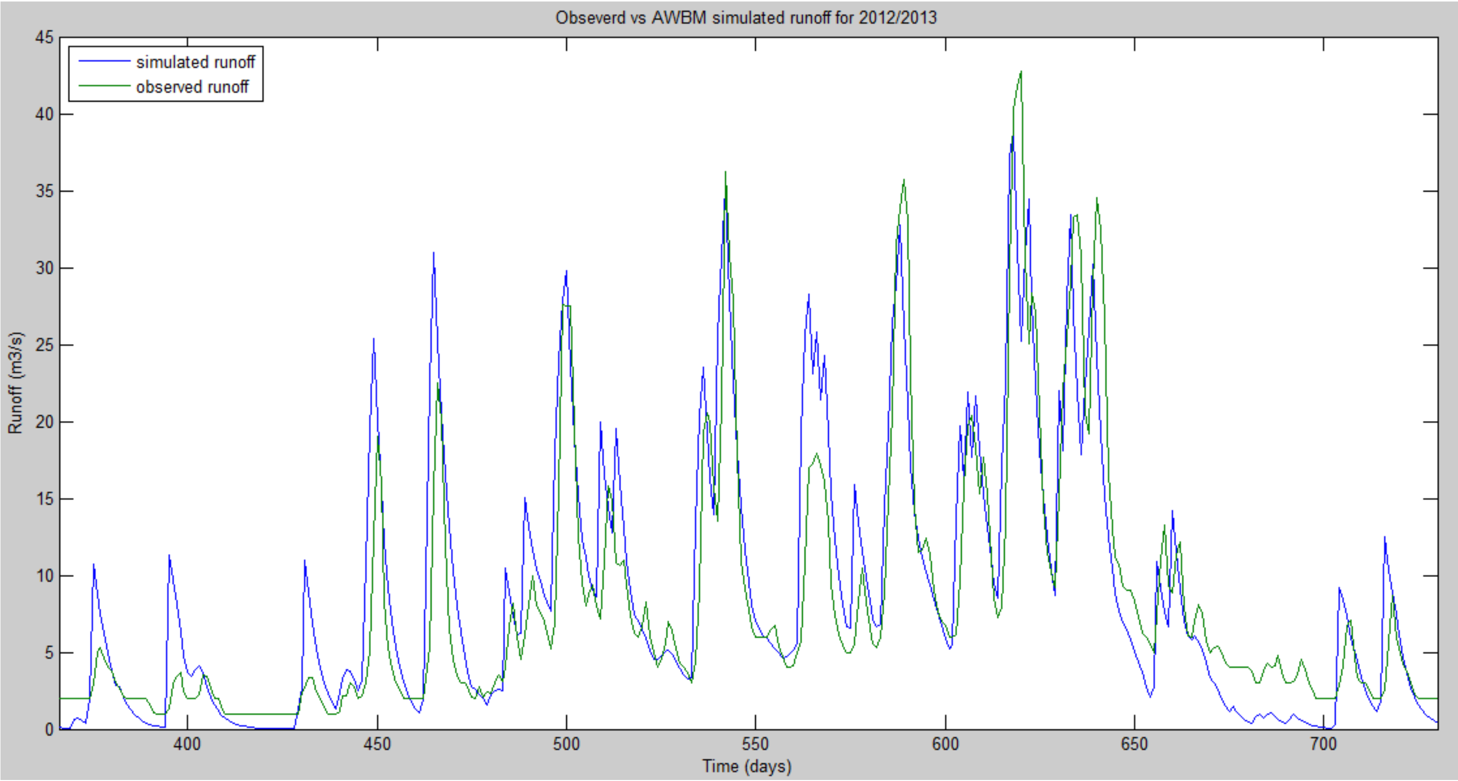


FIGURE 9.4-4 VALIDATION PLOT OF AWBM FOR 1/2012 – 7/2013

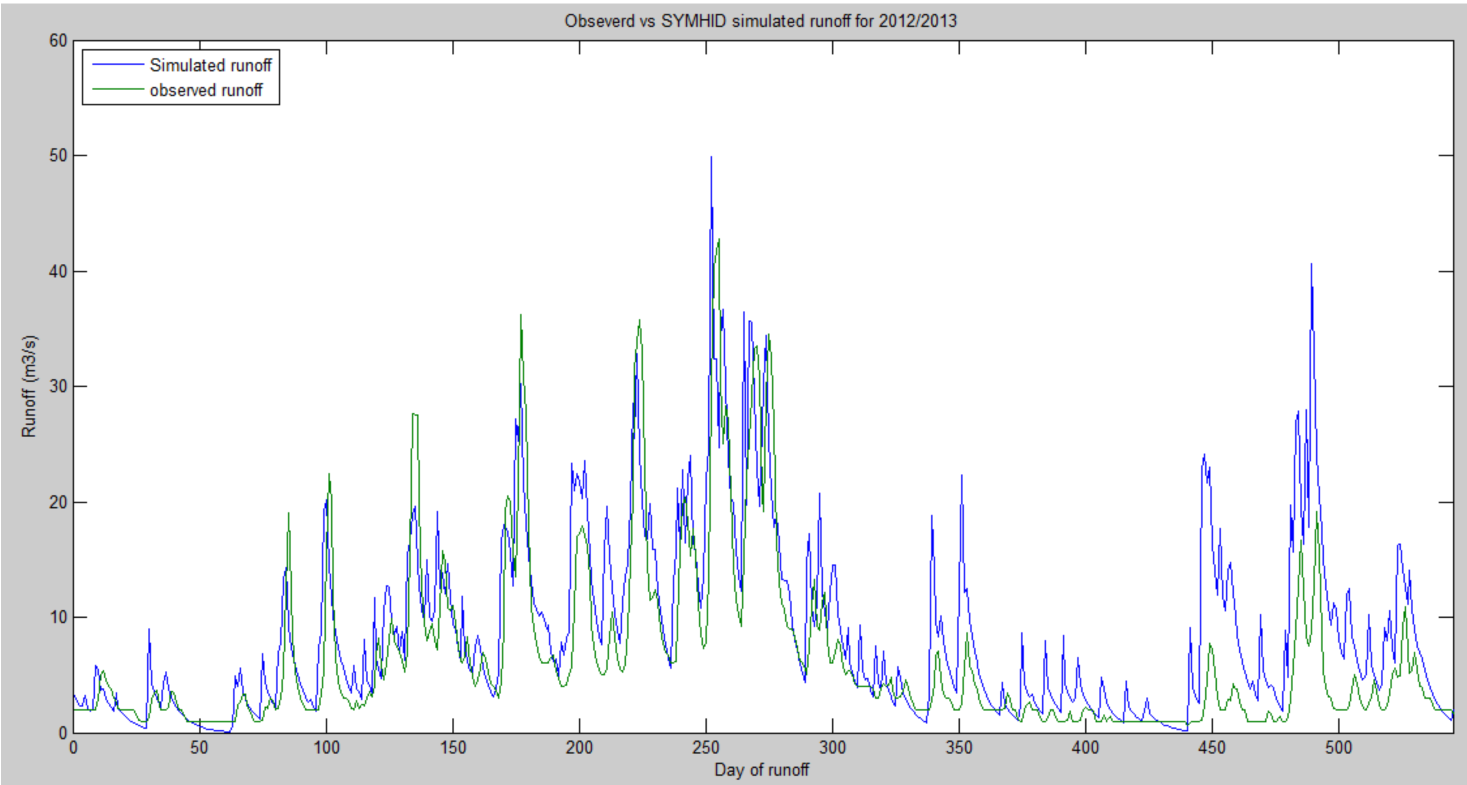


FIGURE 9.4-3 VALIDATION PLOT OF SYMHID FOR 1/2012 – 7/2013

9.5 Sensitivity plots

The sensitivity plots help to understand the importance of the different parameters for calibration.

9.5.1 Parameter evaluation AWBM

The next graphs will show scatter plots on the left side and cummalitive value plots on the right side. The parameters evaluated are the parameters average capacity (C ave), baseflowindex (BFI), baseflow recession constant (kb) and surfaceflow recession constant (ks).

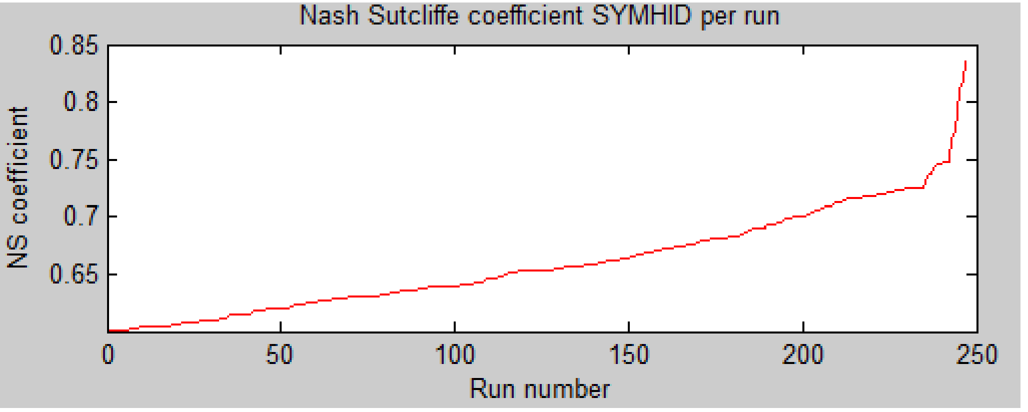


FIGURE 9.5-1 NS COEFFICIENTS FOR THE RUNS

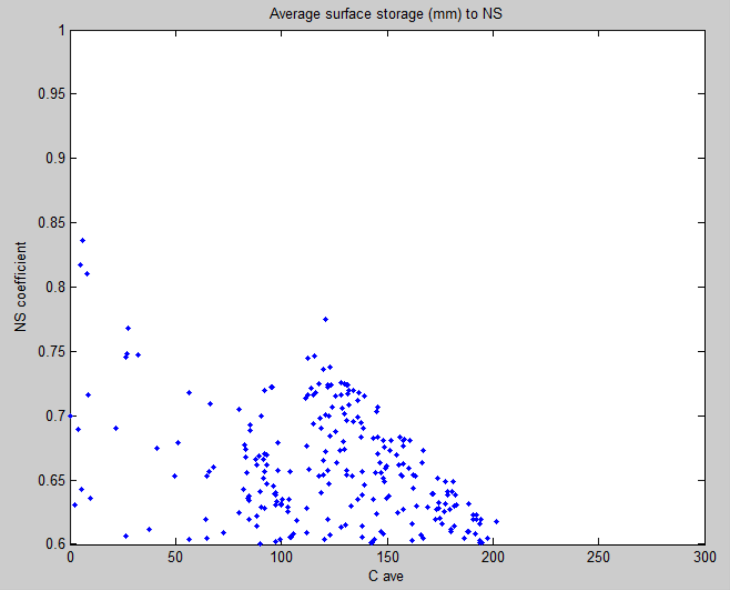


FIGURE 9.5-3 SCATTERPLOT OF C AVE TO NS

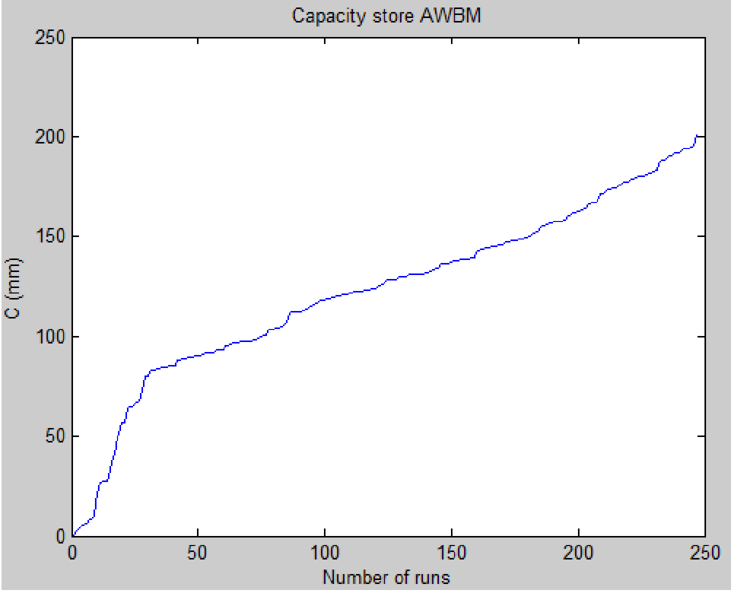


FIGURE 9.5-2 CUMULATIVE GRAPH OF C AVE

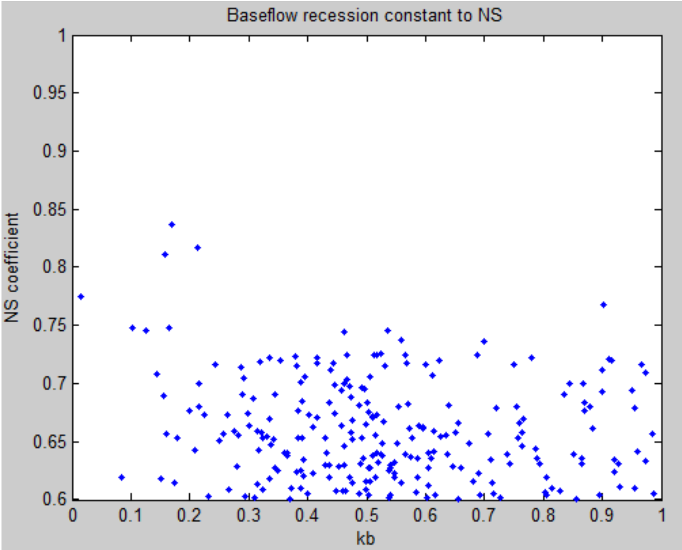


FIGURE 9.5-9 SCATTERPLOT OF KB TO NS

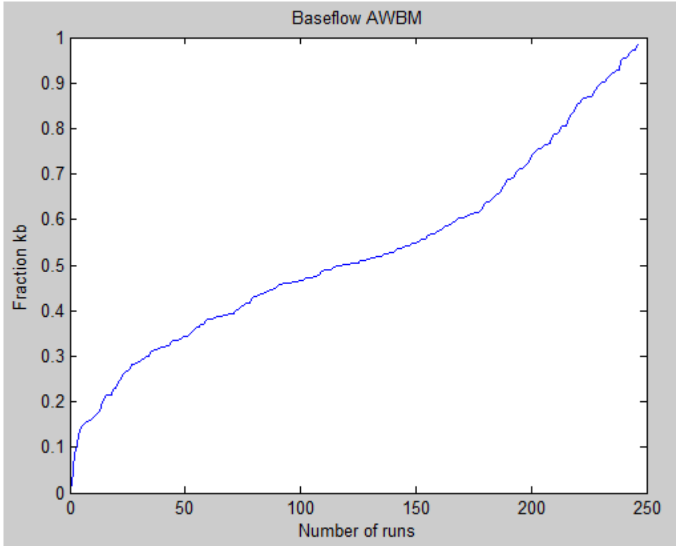


FIGURE 9.5-8 CUMULATIVE GRAPH OF KB

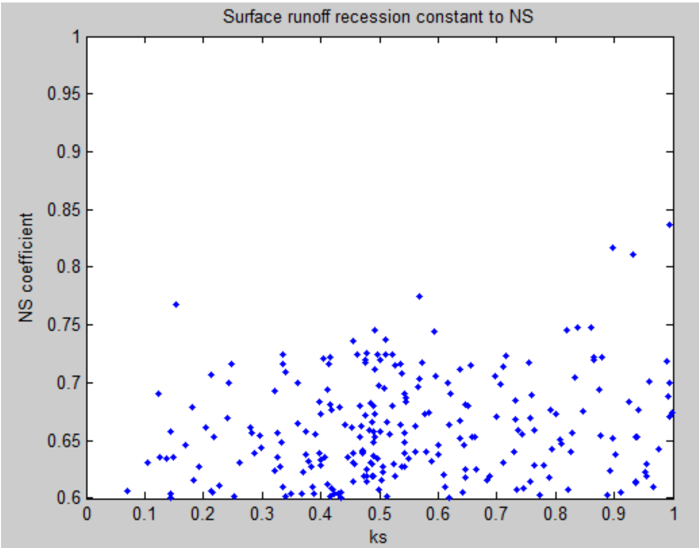


FIGURE 9.5-7 SCATTERPLOT OF KS TO NS

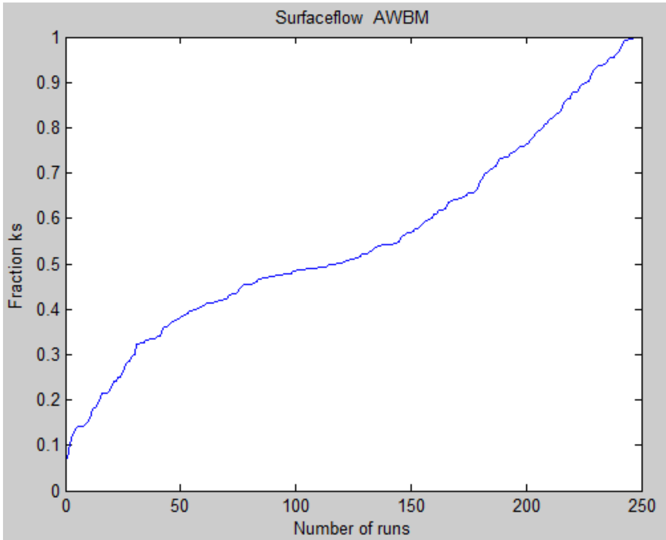


FIGURE 9.5-6 CUMULATIVE GRAPH OF KS

FIGURE 9.5-4 SCATTERPLOT OF BFI TO NS

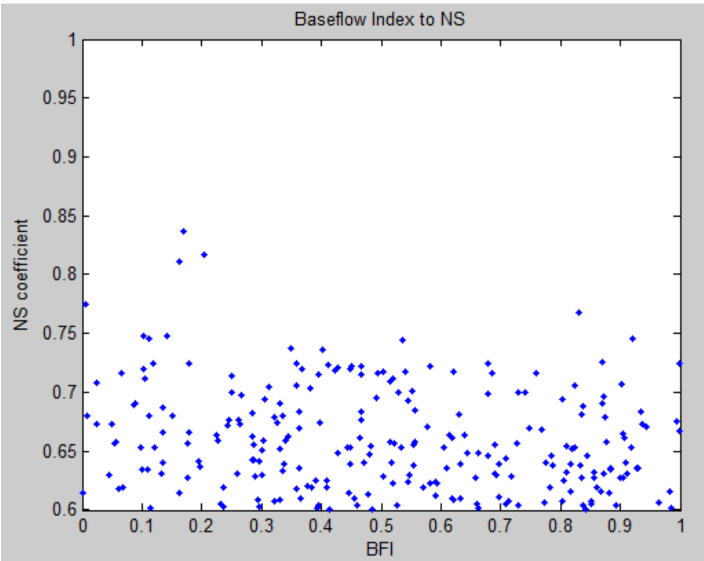
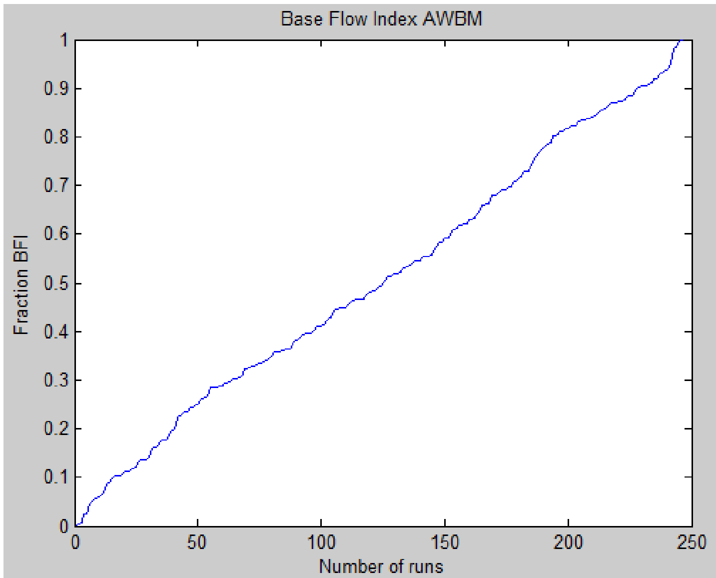


FIGURE 9.5-5 CUMULATIVE GRAPH OF BFI



The cumulative plot gives the probability of a certain value for the parameter. If the line is steep, the value changes quick per run so the probability of the value is small. If the line is relatively flat, a lot of runs have this (range of) values, so the probability of the parameter is large. With straight line (like f=x) there is nothing to say about the probability of a certain value of the parameter. For the capacity it is not likely the value of the parameter is below 50. For the recession constants (base flow and surface flow) it is likely the parameter is between 0.4 and 0.6, and the base flow index can have each value between 0 and 1 with the same probability.

The dotty plots in the left side show the value of the parameter with respect to the NS. Is shows in figure 4 that the lower values of the capacity store give a higher NS coefficient. This is in contrast with figure 5 which shows a higher probability of a value higher than 50.

9.5.2 Parameter evaluation GR4J

The parameters for the GR4J model are x_1: maximum capacity of the production store (mm), x_2: groundwater exchange coefficient (mm), x_3: one day ahead maximum capacity of the routing store (mm) and x_4: time base of unit hydrograph (days).

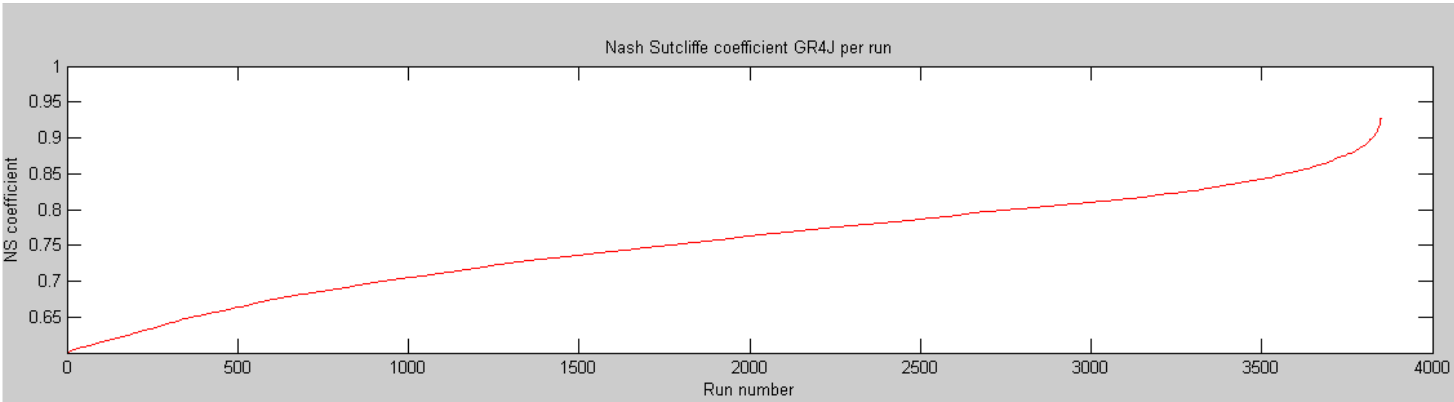


FIGURE 9.5-10 NS FOR THE DIFFERENT RUNS FOR GR4J

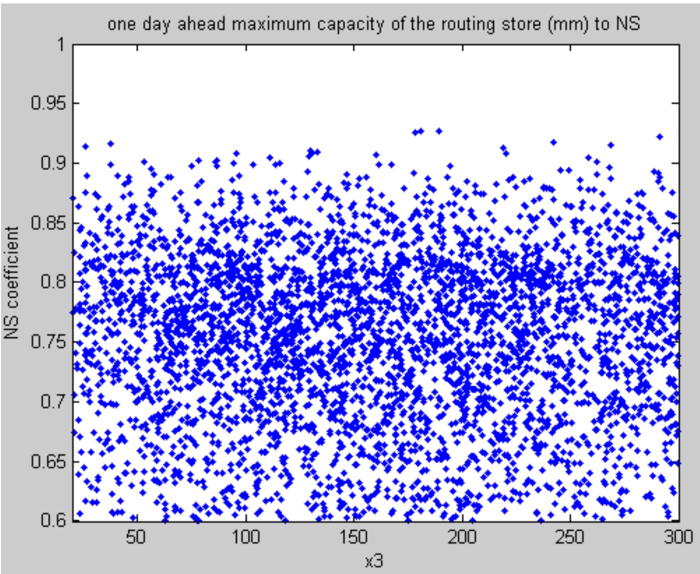


FIGURE 9.5-12 SCATTERPLOT NS AGIANST X3

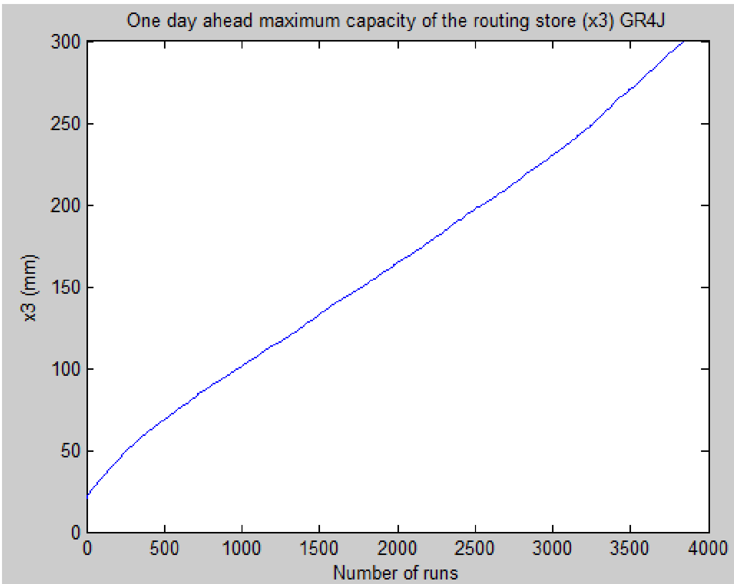


FIGURE 9.5-11 CUMELATIVE PLOT X3

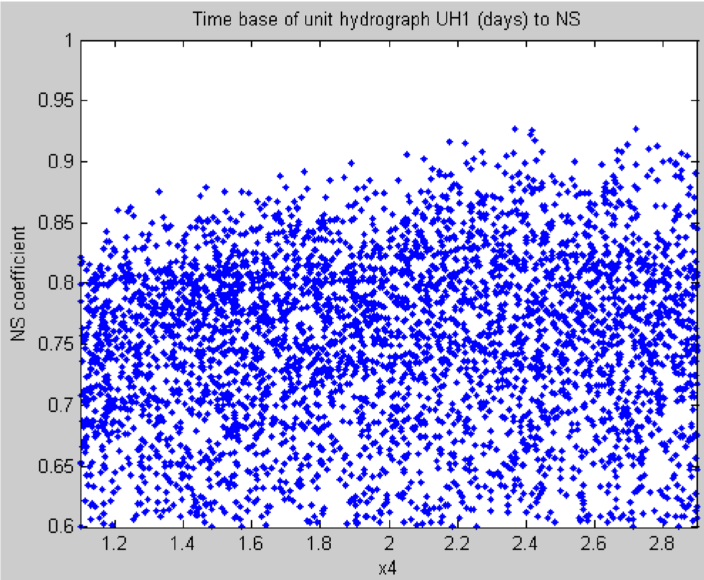


FIGURE 9.5-14 SCATTERPLOT NS AGIANST X4

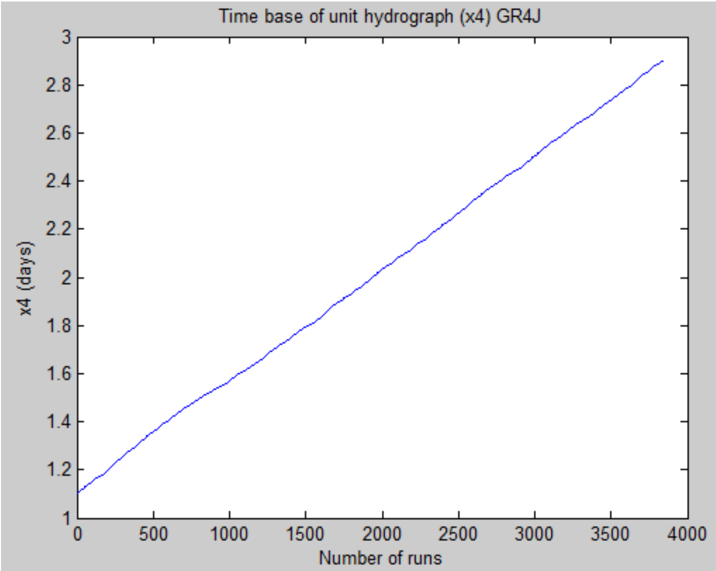


FIGURE 9.5-13 CUMULATIVE PLOT X4

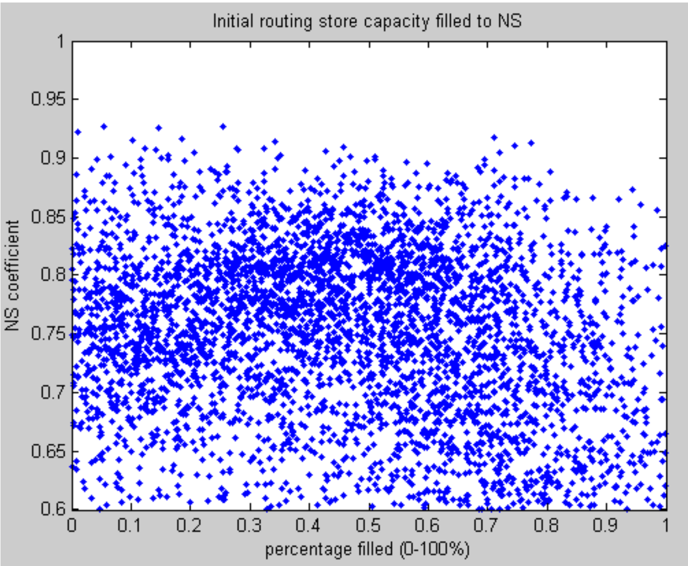


FIGURE 9.5-16 NS AGAINST THE INITIAL ROUTING STORE CAPACITY

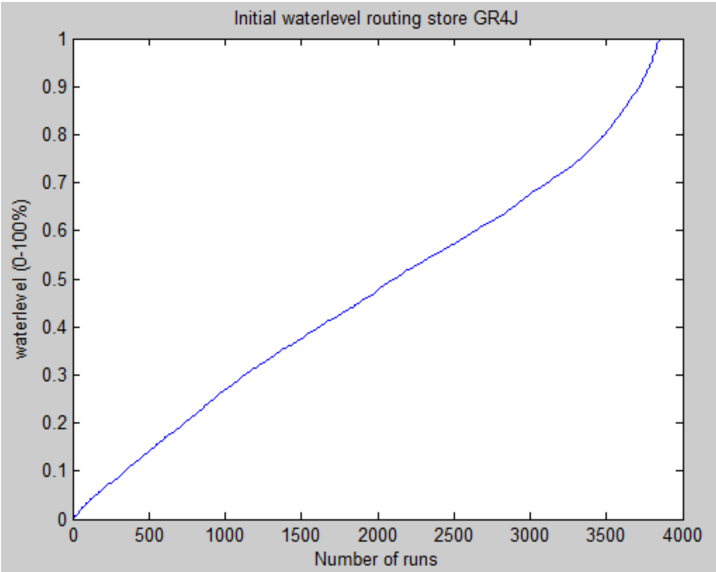


FIGURE 9.5-15 CUMULATIVE PLOT OF THE INITIAL ROUTING STORE CAPACITY

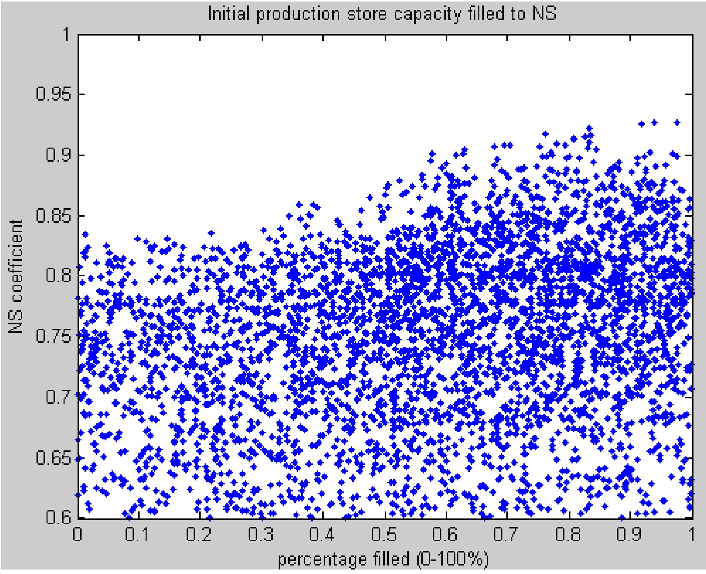


FIGURE 9.5-17 SCATTERPLOT OF THE INITIAL PRODUCTION STORE

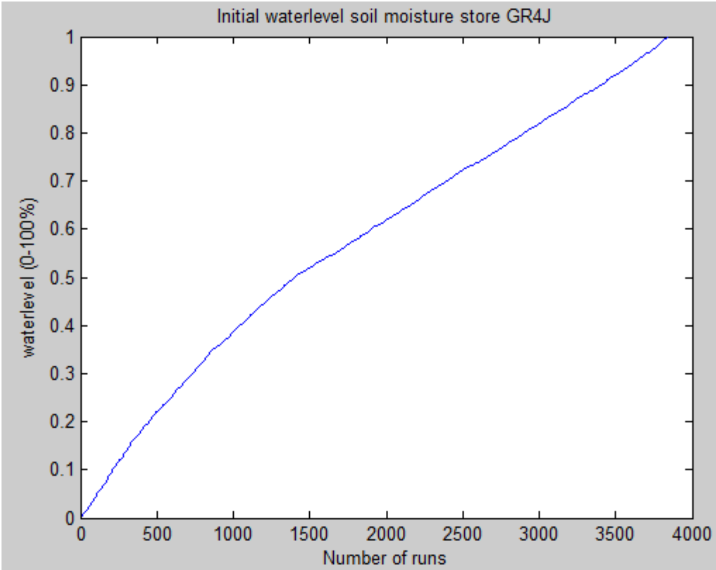


FIGURE 9.5-18 CUMULATIVE PLOT OF THE INITIAL PRODUCTION STORE

9.5.3 Parameter evaluation SYMHID

The parameters evaluated for the SYMHID are Interception store (INSC (mm)), soil moisture capacity (SMSC (mm)), proportionality in interflow (SUB), base flow recession (K), groundwater recharge parameter (CRAK), the infiltration loss parameter (SQ) and the maximum infiltration loss (COEFF (mm)).

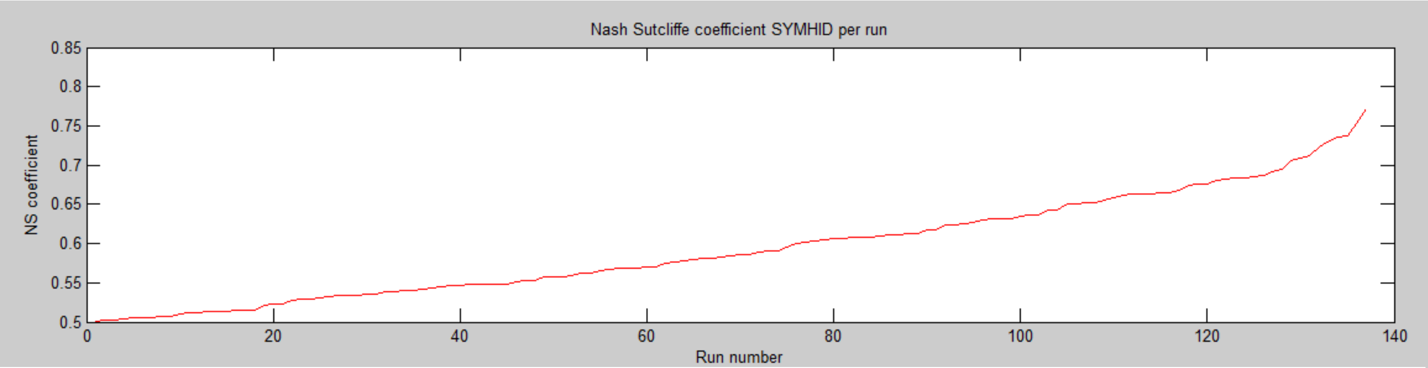


FIGURE 9.5-19 THE NS FOR THE BEST RUNS

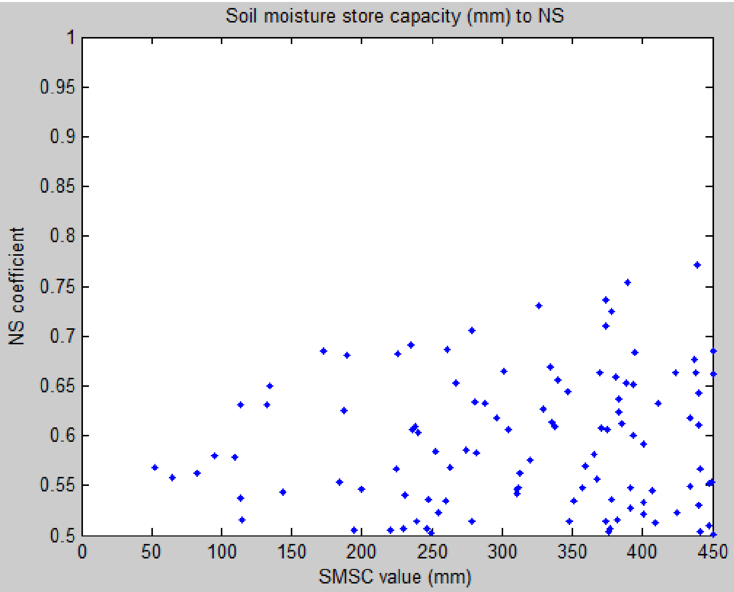


FIGURE 9.5-21 SCATTERPLOT OF NS OF INITIAL SMS

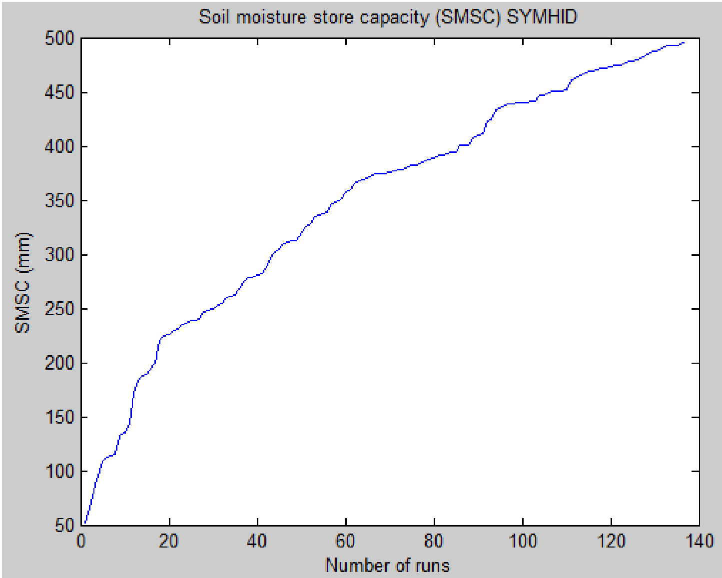


FIGURE 9.5-20 CUMULATIVE OF THE INITIAL SMS

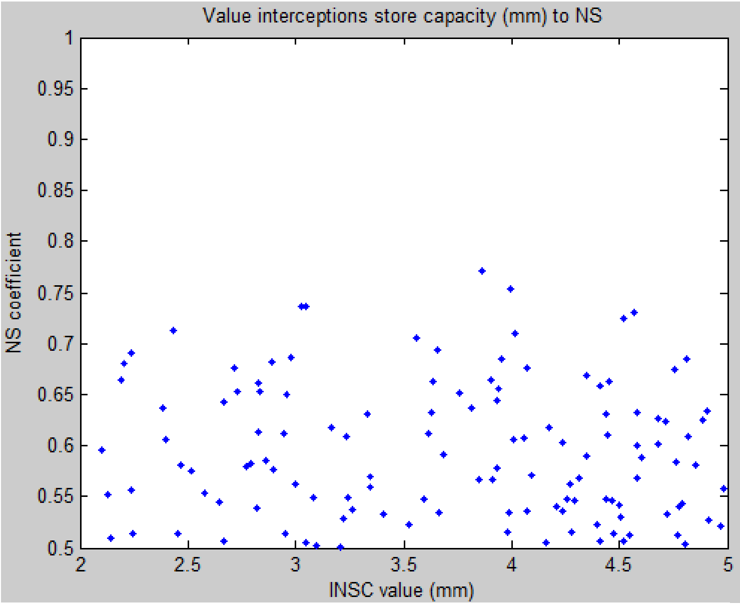


FIGURE 9.5-22 SCATTERPLOT OF SMSC

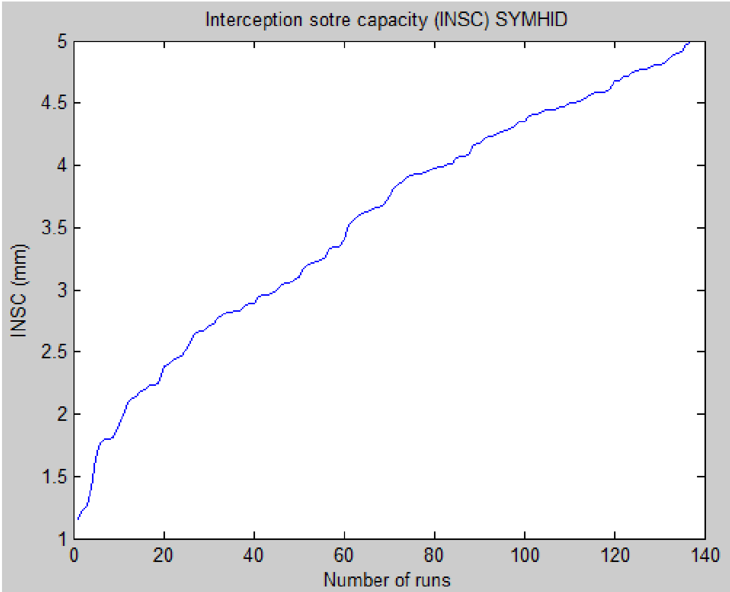


FIGURE 9.5-25 CUMULATIVE PLOT OF SMSC

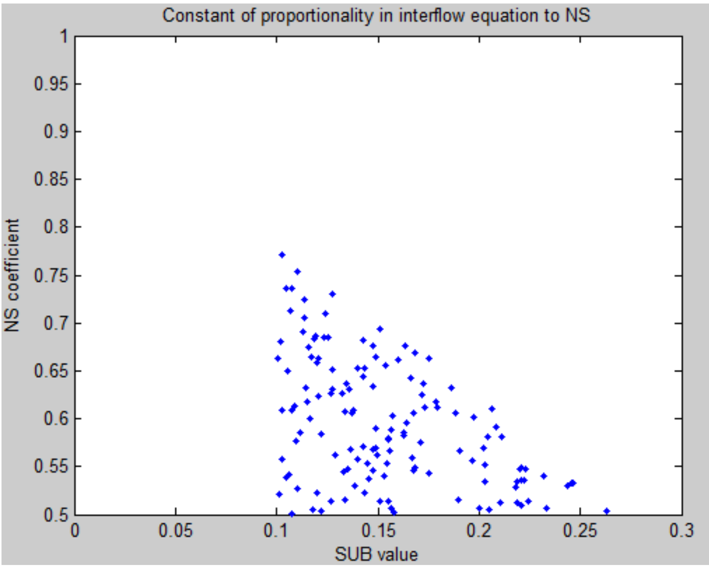


FIGURE 9.5-24 SCATTERPLOT OF NS OF SUB

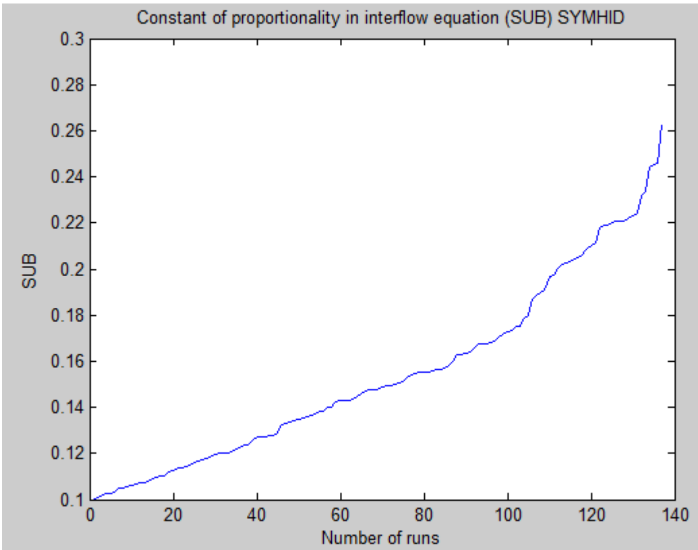


FIGURE 9.5-23 CUMULATIVE PLOT OF SUB

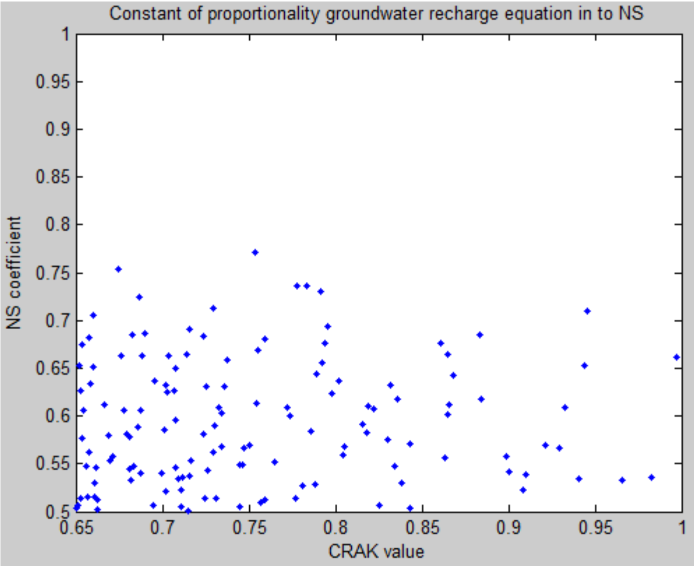


FIGURE 9.5-30 SCATTERPLOT OF NS AGAINST CRAK

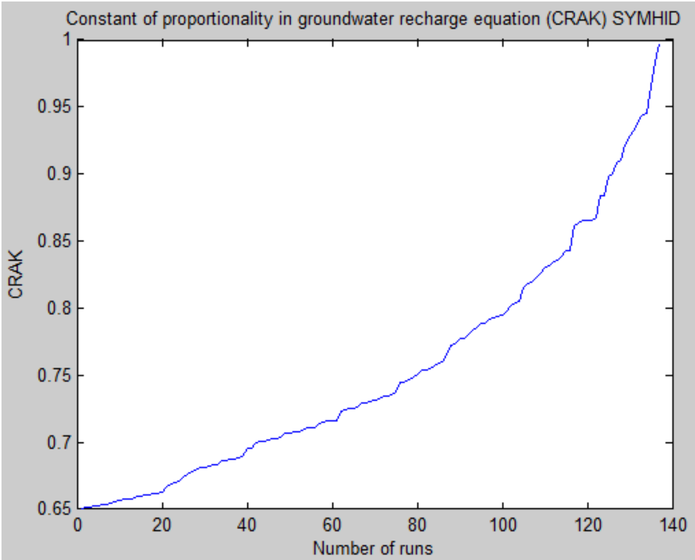


FIGURE 9.5-29 CUMULATIVE PLOT OF CRAK

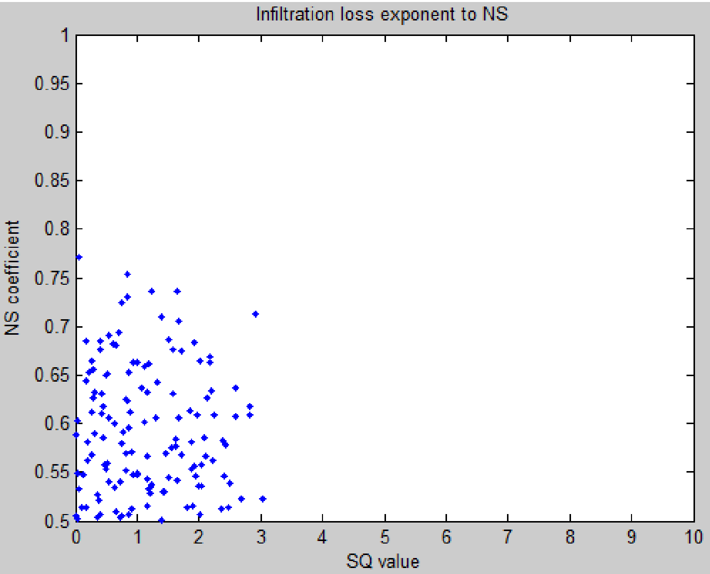


FIGURE 9.5-28 SCATTERPLOT OF NS OF SQ

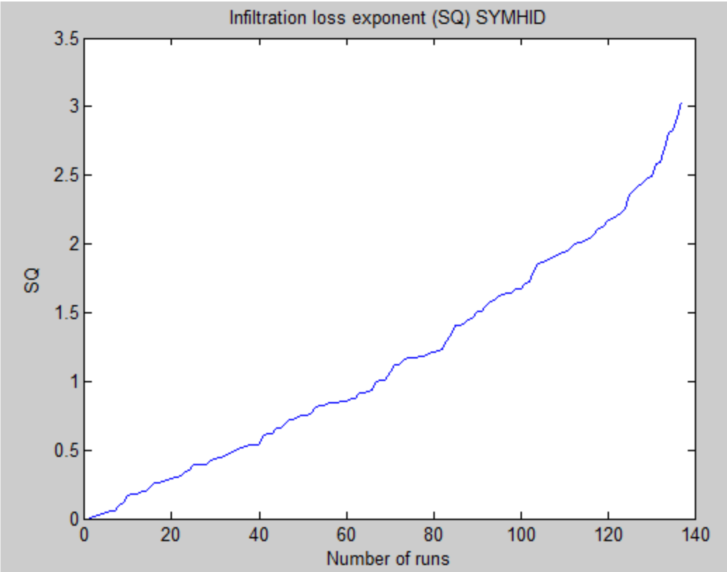


FIGURE 9.5-27 CUMULATIVE PLOT SQ

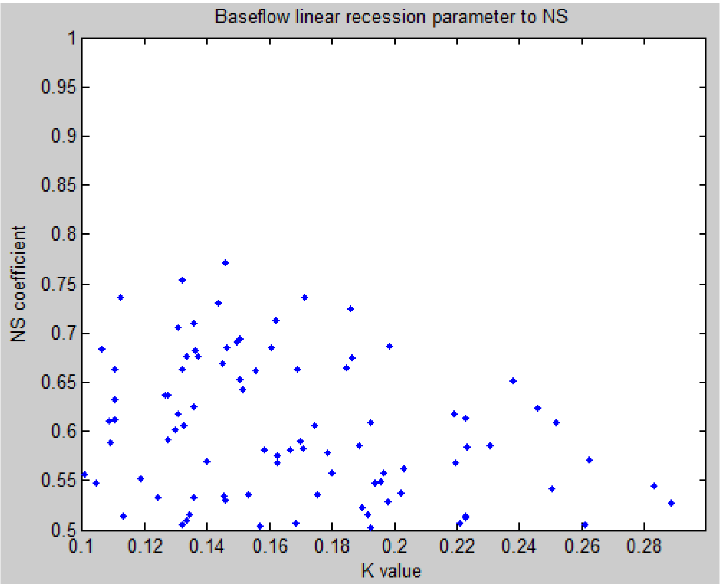


FIGURE 9.5-26 SCATTERPLOT OF NS TO K

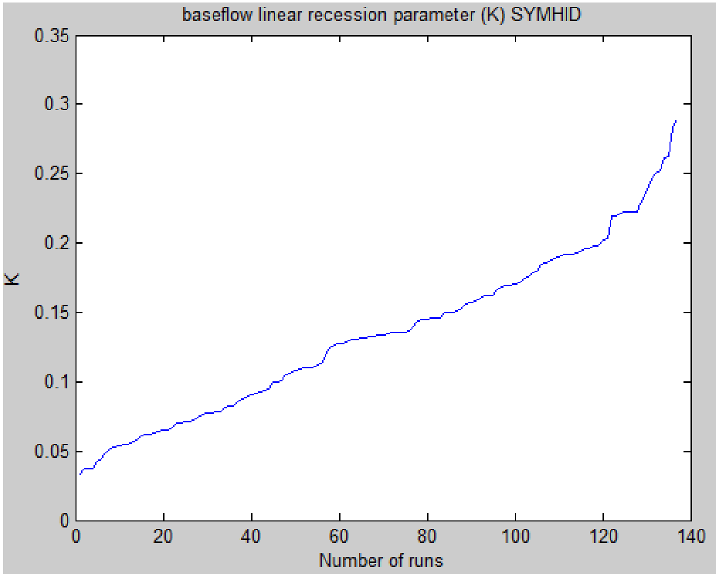


FIGURE 9.5-31 CUMULATIVE PLOT OF K

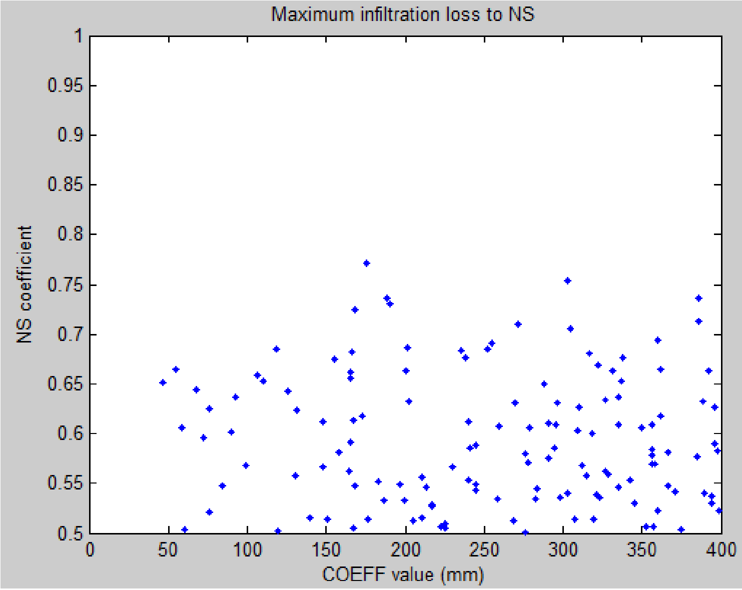


FIGURE 9.5-36 SCATTERPLOT OF NS TO COEFF

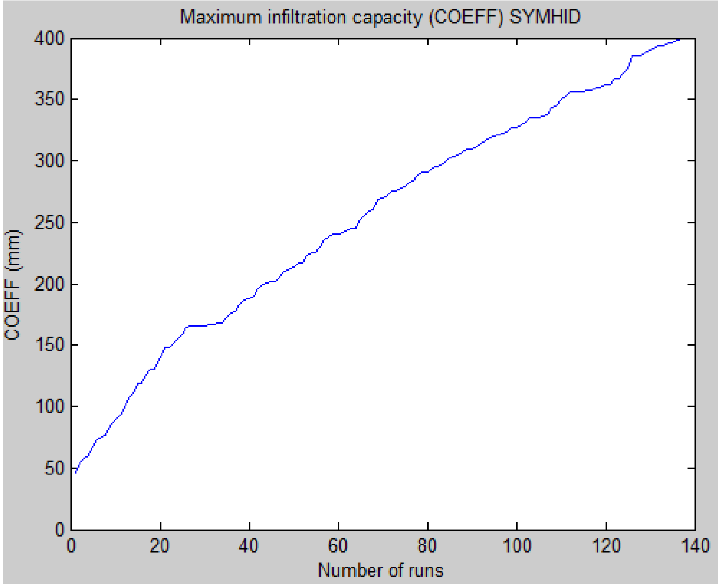


FIGURE 9.5-37 CUMULATIVE PLOT OF COEFF

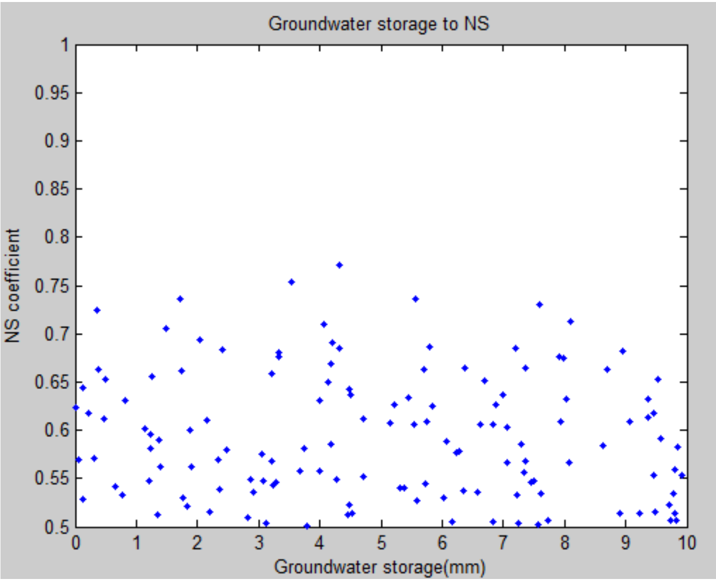


FIGURE 9.5-35 SCATTERPLOT OF NS TO GROUNDWATER STORAGE

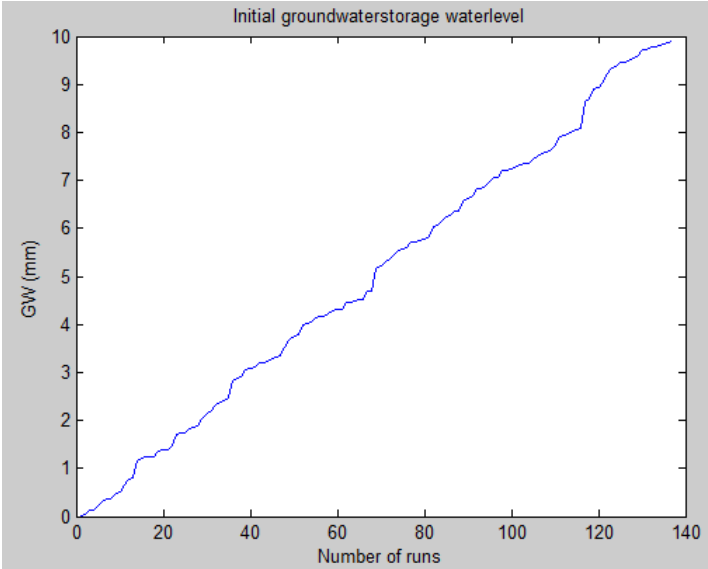


FIGURE 9.5-34 CUMULATIVE PLOT OF GROUNDWATER STORAGE

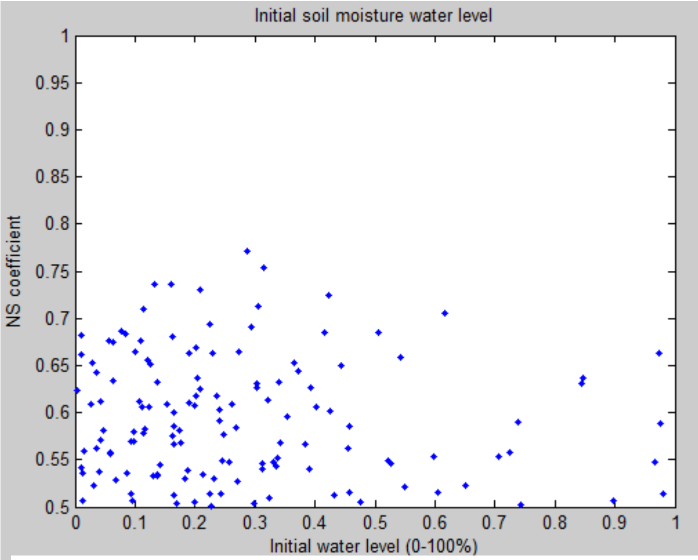


FIGURE 9.5-33 SCATTERPLOT OF THE INITIAL WATER LEVEL

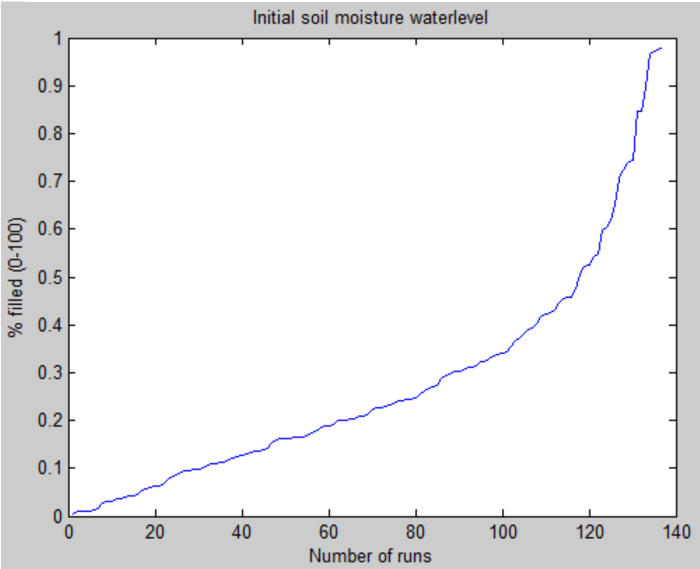


FIGURE 9.5-32 CUMULATIVE PLOT OF INITIAL WATER LEVEL

9.5.4 Confidence interval

Figure 9.5-38 and Figure 9.5-39 show the 95% interval of the simulated runoff for the year 2011. The interval is created to take the best 150 results for the simulated runoff. The lowest 5 percent and the highest five percent are taken off, then the runoff is sorted for each time step from low to high.

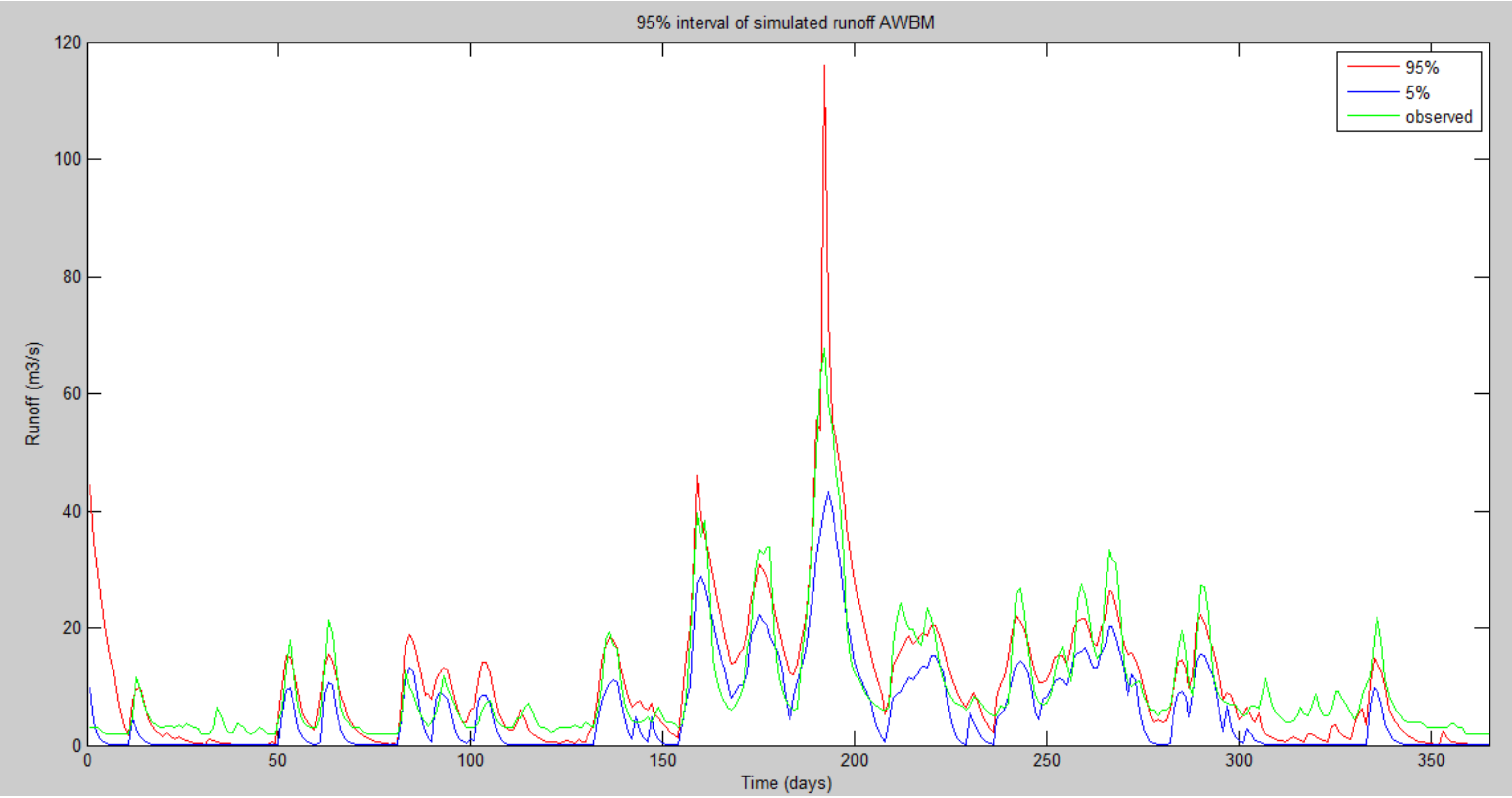


FIGURE 9.5-38 95% INTERVAL OF SIMULATED RUNOFF OF THE AWBM

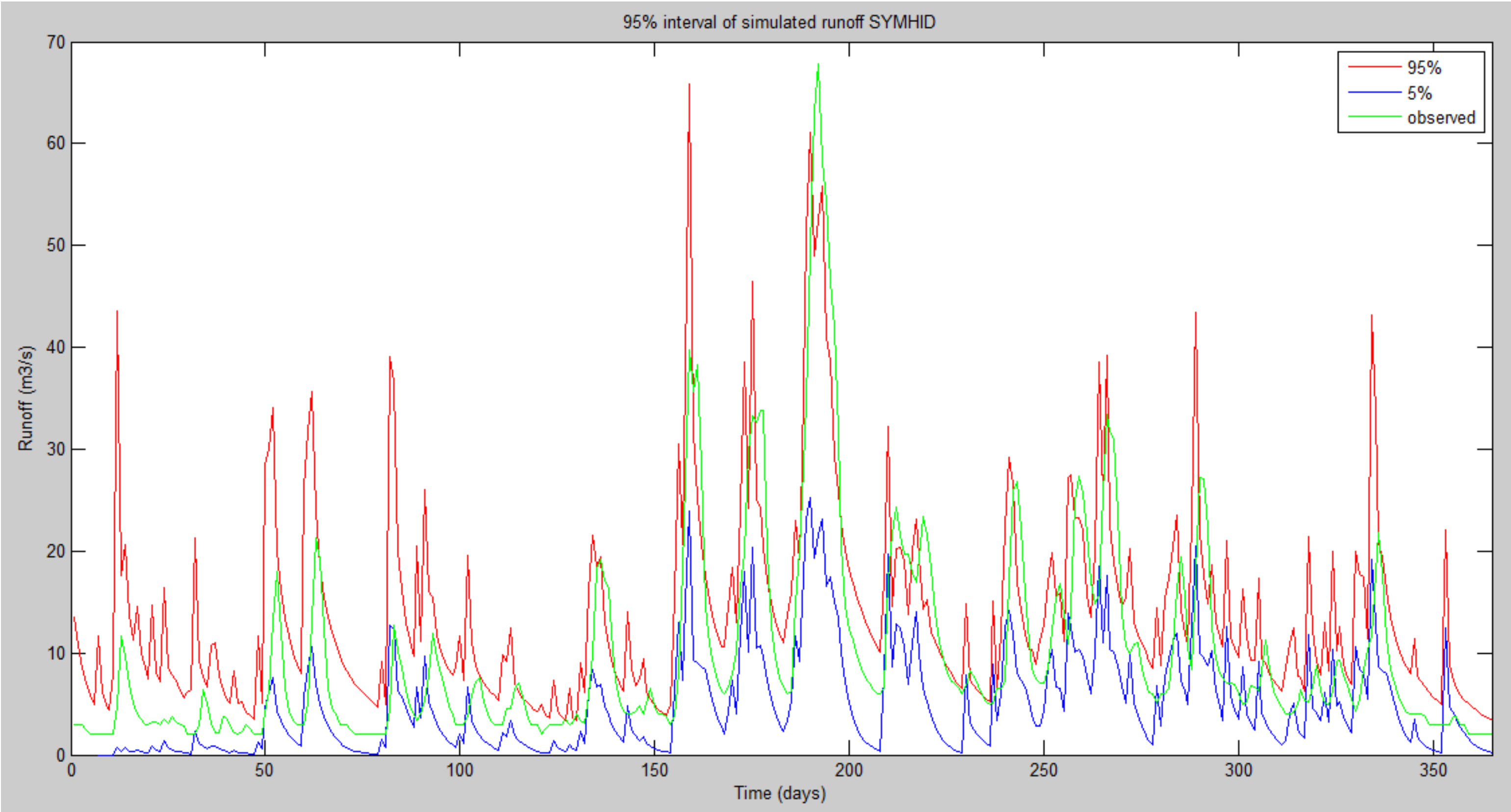


FIGURE 9.5-39 95% INTERVAL OF THE SIMULATED RUNOFF SYMHID